確率統計
Probability and Statistics

第1回講義資料
Lecture notes 1

確率変数と離散分布
Random Variables and Discrete 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

確率論の目標(Goal of probability theory)

例：量子化(Example: Quantization)

非一様な頻度で発生する連続値信号$X \in \mathbb{R}$($\mathbb{R}$:実数全体)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 $X \in \mathbb{R}$ ($\mathbb{R}$:real field) 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.

困難(Challenges)

- 非一様な確率をどのように取り扱うべきか？

  How should we treat non-uniform probability?

- 連続値を取る事象の確率をどのように定義すべきか？

  How should we define the probability of events that take continuous values?

- 理論が実験と合致することをどのように立証すべきか？

  How should we justify the consistency between theory and experiments?

目標(Goal)

これらの困難を解決する確率論を樹立すること。

Establishment of probability theory to solve these challenges.

講義の流れ(Outline of lectures)

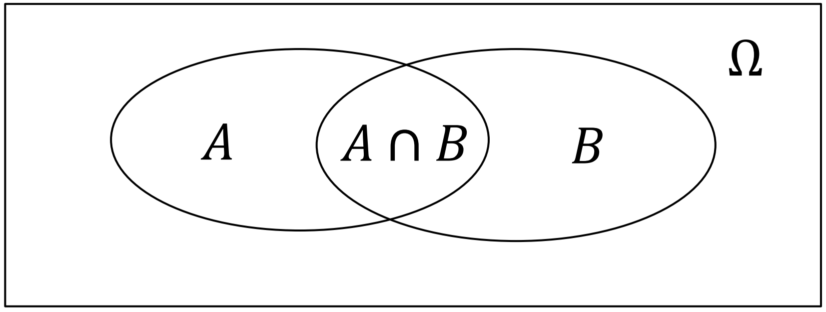
1. 古典的確率論(Classical probability theory)
   - 離散確率分布(事象の数が有限個の場合)
     Discrete probability distribution (Case of a finite number of events)
   - 確率構造の抽出(Extraction of probabilistic structure)
2. 公理的確率論(Axiomatic probability theory)
   - 公理系の樹立(Establishment of an axiomatic system)
   - 連続確率分布の定義(Definition of continuous probability distributions)
3. 理論の検証(Verification of the theory)
   - 大数の法則 (Law of large numbers)

     公理的確率論は実験と合致するか？

     Is the axiomatic probability theory consistent with experiments?

集合と演算(Sets and operations)

全体集合(Universal set)	Ω
部分集合(Subset)	$A \subset B \overset{\text{def}}{=} x \in A \Longrightarrow x \in B$
和集合(Union)	$A \cup B \overset{\text{def}}{=} \{x \in \Omega : x \in A \text{ or } x \in B\}$
共通部分(Intersection)	$A \cap B \overset{\text{def}}{=} \{x \in \Omega : x \in A \text{ and } x \in B\}$
補集合(Complement)	$A^{\mathrm{c}} \overset{\text{def}}{=} \{x \in \Omega : x \notin A\}$
差集合(Relative complement)	$A \setminus B \overset{\text{def}}{=} A \cap B^{\mathrm{c}}$
空集合(Empty set)	$\emptyset \overset{\text{def}}{=} \Omega^{\mathrm{c}}$



補集合の補集合(Complement of complement) $(A^{\mathrm{c}})^{\mathrm{c}} = A$

ド・モルガンの法則(De-Morgan’s law) $(A \cap B)^{\mathrm{c}} = A^{\mathrm{c}} \cup B^{\mathrm{c}}$

TOYOHASHI UNIVERSITY OF TECHNOLOGY

可算(無限)個の集合演算(Operations for a countable (=countably infinite) number of sets)

可算個の集合列(Countable number of sets) $\{A_i \in \Omega\}_{i=1}^{\infty}$

和集合(Union)

$$\bigcup_{i=1}^{\infty} A_i \stackrel{\text{def}}{=} \{x \in \Omega :\ \exists n \in \mathbb{N} \text{ s.t. } x \in A_n\}$$

ある自然数nが存在して、$x \in A_n$が成り立つ。

There exists some natural number n such that $x \in A_n$ holds.

共通部分(Intersection)

$$\bigcap_{i=1}^{\infty} A_i \stackrel{\text{def}}{=} \{x \in \Omega : x \in A_n \text{ for } \forall n \in \mathbb{N}\}$$

すべての自然数nに対して、$x \in A_n$が成り立つ。

$x \in A_n$ holds for all natural number n.

ド・モルガンの法則(De-Morgan's law)

$$\left(\bigcap_{i=1}^{\infty} A_i\right)^{\mathrm{c}} = \bigcup_{i=1}^{\infty} A_i^{\mathrm{c}}$$

$$\because \quad \left(\bigcap_{i=1}^{\infty} A_i\right)^{\mathrm{c}} = \left\{x \in \Omega : x \notin \bigcap_{i=1}^{\infty} A_i\right\} = \{x \in \Omega :\ \exists n \in \mathbb{N} \text{ s.t. } x \notin A_n\}$$

$$= \{x \in \Omega : \exists n \in \mathbb{N} \text{ s.t. } x \in A_n^{\mathrm{c}}\} = \bigcup_{i=1}^{\infty} A_i^{\mathrm{c}}.$$

例：コイン投げ(Example: Coin tossing)

コインを1回投げたときに、表が出る事象をω_0、裏が出る事象をω_1とする。

Let ω_0 and ω_1 denote the occurrence events of a head and tail in one coin-tossing, respectively.

二つの事象は等確率で起きる。

The two events occur with an equal probability.

$$P(\omega_0) = P(\omega_1) = \frac{1}{2}.$$

確率構造の抽出(Extraction of probabilistic structure)

- 表か裏のどちらかは確率1で出る。(Either a head or a tail occurs with probability 1.)

$$P(\Omega) = P(\omega_0) + P(\omega_1) = 1, \qquad \Omega = \omega_0 \cup \omega_1.$$

- 表が出ない確率は1/2である。(A head does not occur with probability 1/2)

$$P(\omega_0^{\mathrm{c}}) = P(\Omega \setminus \omega_0) = P(\omega_1) = \frac{1}{2}.$$

- 何も出ない確率は0である。(Nothing occurs with zero probability.)

$$P(\emptyset) = 0.$$

事象は集合である(An event is a set.)

事象を集合とみなすと、数学的構造が単純になる。

Regarding an event as a set simplifies the mathematical structure.

確率空間(Probability space): $(\Omega, \mathcal{F}, P)$

$\Omega = \{\omega_0, \omega_1\}$ 標本空間(Sample space)

特に、標本空間Ωの各元ω_iを標本点と呼ぶ。

In particular, each element of the sample space is called an outcome.

$\mathcal{F} = \{\emptyset, \{\omega_0\}, \{\omega_1\}, \Omega\}$ 集合体(Field of sets)

集合体$\mathcal{F}$の各元(Ωの部分集合)を事象と呼ぶ。

Each element (subset of Ω) of $\mathcal{F}$ is called an event.

$P: \mathcal{F} \to [0, 1]$ 確率分布(Probability distribution)

$P(\Omega) = 1$を満たす集合関数(入力が集合)である。

A set function (Input is a set) satisfying $P(\Omega) = 1$.

確率構造の抽出(Extraction of probabilistic structure)

集合体$\mathcal{F}$の性質(Properties of the field $\mathcal{F}$ of sets)

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. 事象AとBが$\mathcal{F}$に含まれるならば、和集合$A \cup B$も$\mathcal{F}$に含まれる。

   If $\mathcal{F}$ contains events A and B, $\mathcal{F}$ also contains the union $A \cup B$.

確率分布Pの性質(Properties of the 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. 二つの排反事象AとB($A \cap B = \emptyset$)に対して、

   For two exclusive events A and B ($A \cap B = \emptyset$),

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

離散分布の確率構造(Probabilistic structure of discrete distributions)

標本空間(Sample space) Ω　有限集合(Finite set)

集合体(Field of sets) $\mathcal{F}$

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. 事象列$\{A_i\}_{i=1}^k$が$\mathcal{F}$に含まれるならば、有限和集合$\cup_{i=1}^k A_i$も$\mathcal{F}$に含まれる。

   If $\mathcal{F}$ contains events $\{A_i\}_{i=1}^k$, $\mathcal{F}$ also contains the finite union $\cup_{i=1}^k A_i$.

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. 排反事象列$\{A_i\}_{i=1}^k(A_i \cap A_j = \emptyset \text{ for } i \neq j)$に対して、

   For exclusive events $\{A_i\}_{i=1}^k(A_i \cap A_j = \emptyset \text{ for } i \neq j)$,

$$P\left(\cup_{i=1}^k A_i\right) = \sum_{i=1}^k P(A_i).$$

有限加法性(Finite additivity)

確率空間$(\Omega, \mathcal{F}, P)$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) $\mathcal{F}$ ($\neq$ Ω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$\cup_{i=1}^{\infty} A_i$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 $\cup_{i=1}^{\infty} A_i$.

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 $i \neq j$),

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

確率変数の例(Example of random variables)

サイコロ(Dice)

サイコロを一回投げて、3以上が出たら100円、2以下であれば10円が得られるゲームを考える。このゲームを行ったときに獲得賞金としてどのような値が期待できるか？

Consider a dice game in which one gets 100 yen if 3 to 6 occur, otherwise 10 yen. How much can he/she expect as a cash prize in this game?

標本空間(Sample space) $\Omega = \{\omega_0, \omega_1\}$

ω_0: 2以下が出る。(outcomes of 1 and 2.)　　ω_1: 3以上が出る。(outcomes of 3 to 6.)

確率分布(Probability distribution)

$$P(\omega_0) = \frac{1}{3}, \qquad P(\omega_1) = \frac{2}{3}.$$

獲得賞金を表す確率変数(Random variable to represent the cash prize)

$$X(\omega_0) = 10, \qquad X(\omega_1) = 100.$$

賞金の期待値(Expectation of the cash prize)

$$\mathbb{E}[X] = 10 \cdot P(\omega_0) + 100 \cdot P(\omega_1) = 70.$$

(実)確率変数の定義(Definition of (real) random variables)

標本空間から実数体への関数$X\colon\Omega\to\mathbb{R}$が以下を満たすとき、
関数Xを(実)確率変数と呼ぶ。

A mapping $X\colon\Omega\to\mathbb{R}$ from the sample space to the real field is called a (real) random variable if the following holds:

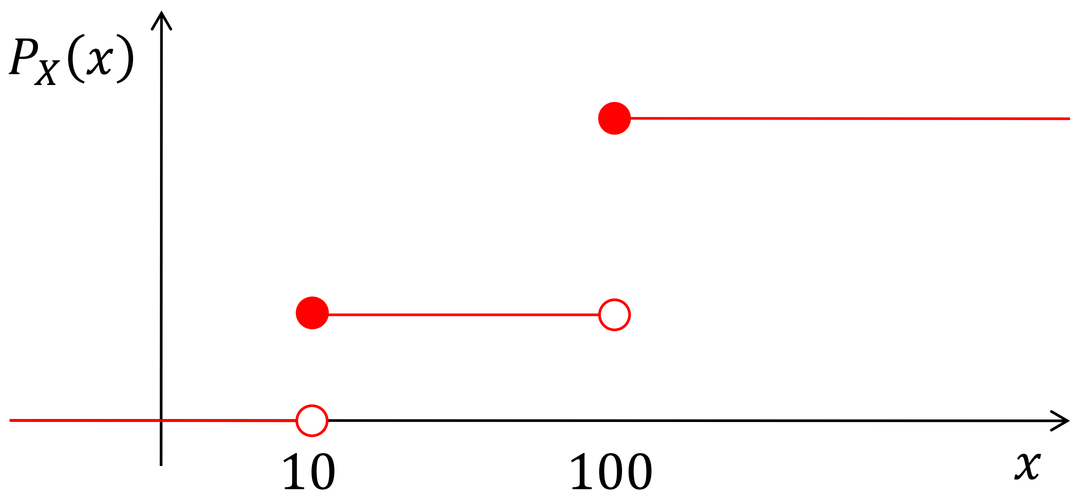
For all $x\in\mathbb{R}$, $\{\omega\in\Omega : X(\omega)\le x\}\in\mathcal{F}$.

任意の$x\in\mathbb{R}$に対して、$X(\omega)\le x$となる事象の発生確率を定義できる。

For any $x\in\mathbb{R}$, the occurrence probability of the event $\{X(\omega)\le x\}$ is well defined.

サイコロの例(The dice example)

$$P_X(x)=P(\{\omega\in\Omega : X(\omega)\le x\})=\begin{cases}0 & \text{for } x<10.\\ 1/3 & \text{for } x\in[10,100),\\ 1 & \text{for } x\ge 100.\end{cases}$$



確率分布関数の定義(Definition 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:

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

注意1(Remark 1)

確率分布関数は、分布関数あるいは累積分布関数とも呼ばれる。

The probability distribution function is also called distribution function or cumulative distribution function.

注意2(Remark 2)

確率分布か分布関数のどちらかを明示することなく、「分布」という言葉を確率変数の統計的性質を定義するものという意味で使用する場合がある。

The terminology *distribution* may be used to indicate something to define the statistical properties of a random variable, without specifying probability distribution or distribution function.

注意3(Remark 3)

確率分布関数は、以下のように略記される場合がある。

The probability distribution function may be abbreviated as follows:

$$P_X(x) = P(X(\omega) \leq x) = P(X \leq x) = P(x).$$

離散分布(Discrete distributions)

ベルヌーイ分布(Bernoulli distribution)

ベルヌーイ確率変数Xは確率pで1を取り、確率$1-p$で0を取る。

A Bernoulli random variable X takes 1 and 0 with probabilities p and $1-p$, respectively.

二項分布(Binomial distribution) : $\mathcal{B}(n,p)$

二つのパラメータ$n \in \mathbb{N}$と$p \in [0,1]$を持つ二項確率変数Xは、整数$k \in \{0,\ldots,n\}$を確率p_kで取る。

A binomial random variable X with two parameters n and p takes an integer $k \in \{0,\ldots,n\}$ with probability p_k.

$$p_k = \binom{n}{k} p^k (1-p)^{n-k}, \qquad \binom{n}{k} = \frac{n!}{(n-k)!\,k!}.$$

ポアソン分布(Poisson distribution)

パラメータ$\lambda > 0$を持つポアソン確率変数Xは、$n = 0,1,\ldots$を確率p_nで取る。

A Poisson random variable X with a parameter $\lambda > 0$ takes $n = 0,1,\ldots$ with probability p_n.

$$p_n = \frac{\lambda^n e^{-\lambda}}{n!}.$$

離散分布の期待値(Expectation of discrete distributions)

離散集合$\{x_i\}_{i=1}^k$上の分布P_Xから多数の実現値を得たときに、期待値$\mathbb{E}[X]$を実現値の算術平均に等しくなるように定義したい。

Suppose that we obtained many realizations sampled from a distribution P_X on a discrete set $\{x_i\}_{i=1}^k$. Define the expectation $\mathbb{E}[X]$ such that it is equal to the arithmetic mean of the realizations.

確率$P(X = x_i) = p_i$*をx_iの発生頻度とみなすと、N個の実現値の並び替え後に、以下のようになるはずである。

Regarding the probability $P(X = x_i) = p_i$ as the occurrence ratio of x_i, after a permutation of N realizations, we should obtain the following sequence:

$$\overbrace{(\underbrace{x_1,\dots,x_1}_{Np_1},\underbrace{x_2,\dots,x_2}_{Np_2},\dots,\underbrace{x_k,\dots,x_k}_{Np_k})}^{N\to\infty}$$

$$\text{Arithmetic mean} = \frac{\sum_{i=1}^k x_i \cdot Np_i}{N} = \sum_{i=1}^k x_i p_i \stackrel{\text{def}}{=} \mathbb{E}[X].$$

*:確率分布関数と区別するために確率関数と呼ぶことがあるが、本講義では使わない。離散分布の場合に確率分布を確率関数の意味で使用することがある。

While it may be called probability mass function (pmf) to distinguish it from the probability distribution function, this lecture does not use pmf. The probability distribution may mean pmf in the discrete case.

離散分布の期待値の定義(Definition of the expectation of discrete distributions)

$\sum_{i=1}^{\infty}|x_i|P(X=x_i)<\infty$のとき、離散確率変数$X$の期待値を以下で定義する。

If $\sum_{i=1}^{\infty}|x_i|P(X=x_i)<\infty$ holds, the expectation of a discrete random variable X is defined as

$$\mathbb{E}[X] \stackrel{\text{def}}{=} \sum_{i=1}^{\infty} x_i P(X=x_i).$$

一般化(Generalization)

ある決定論的な関数fに対して$\sum_{i=1}^{\infty}|f(x_i)|P(X=x_i)<\infty$のとき、$f(X)$の期待値を以下で定義する。

If $\sum_{i=1}^{\infty}|f(x_i)|P(X=x_i)<\infty$ holds for some deterministic function f, the expectation of $f(X)$ is defined as

$$\mathbb{E}[f(X)] \stackrel{\text{def}}{=} \sum_{i=1}^{\infty} f(x_i) P(X=x_i).$$

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.