第 0 章「数 理 統 計 学」Introduction

(i) 講義内容

中級レベルの数理統計学について，推測統計の観点からの講義を行う．統計学は，データに基づいてデータの背後にある母集団の構造を推論する科学である．データは，あくまでも母集団の一部であり，データ（部分）から母集団（全体）を推論する際には，何らかの誤りがある．数理統計学は，その誤りを確率的な概念で捉えて，統計的推論を進める．講義内容は以下の通り．

第	0	章	「数理統計学」Introduction
第	1	章	確率変数と確率分布
第	2	章	確率分布の代表的モデル
第	3	章	2 次元確率ベクトルの分布
第	4	章	多変量確率ベクトルの分布
第	5	章	確率過程
第	6	章	標本分布
第	7	章	統計学における情報量
第	8	章	統計的推定問題
第	9	章	仮説検定問題
第	10	章	最尤推定

(ii) テキストと参考書

テキスト：「数理統計学（改訂版）」稲垣宣生 著，裳華房

参考書：「確率と統計演習」鈴木・安岡・志村 共編，共立出版
「入門・演習 数理統計」野田・宮岡 共著，共立出版

(iii) レジュメ

私の HP

`http://www.econ.hit-u.ac.jp/~tanaka/org.html`

の「数理統計学」の項目をクリックすれば，上記 (i) のレジュメにアクセスすることができる．ダウンロードも可能である．

(iv) 単位の評価方法

数回の宿題（教科書の演習問題を解くこと）と 2 回の試験（持ち込み不可）で総合的に評価する．

１．事象と確率

試行：同じ条件のもとで繰り返すことのできる実験，観測，調査などの総称

全事象：ある試行により起こりうるすべての結果の集まり（集合）で，以下では総称的に Ω で表す.

$\sigma-$ **集合族：**全事象 Ω の部分集合から作られる集合の集まりで，次の条件をみたす $\mathcal{F}$ のことをいう．$\sigma-$ 代数，$\sigma-$ 集合体，可算加法族などとも呼ばれる.

(C1) $\Omega \in \mathcal{F}$
(C2) $A \in \mathcal{F} \Longrightarrow A^c \in \mathcal{F}$ (A^c は A の補集合を表す)
(C3) $A_n \in \mathcal{F}$ $(n = 1, 2, \cdots) \Longrightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$

事象：$\sigma-$ 集合族の要素（集合）のことであり，確率が議論できる対象である.（全事象 Ω が可算集合ならば，その任意の部分集合は事象である．しかし，Ω が非可算無限集合ならば，その任意の部分集合は必ずしも事象とはならない.）

さまざまな事象：全事象 Ω，余事象 A^c，空事象 ϕ（$=\Omega^c$），和事象 $A \cup B$，積事象 $A \cap B$，
差事象 $A - B = A \cap B^c$，互いに排反な事象 $A \cap B = \phi$

確率は，$\mathcal{F}$ の要素，すなわち事象に対して定義される（全事象 Ω の任意の部分集合を扱うことは一般に不可能）.

$(\Omega, \mathcal{F})$ の２つの組を**可測空間（Measurable Space）**という.

確率の公理：次の３つの条件をみたす集合関数 P を確率という．**確率測度（Probability Measure）**ともいう.

(P1) 任意の $A \in \mathcal{F}$ に対して $0 \leq P(A) \leq 1$
(P2) $P(\Omega) = 1$
(P3) $A_i \cap A_j = \phi$ $(i \neq j)$ ならば

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n) \qquad \textbf{（完全加法性）}$$

このとき，$(\Omega, \mathcal{F}, P)$ の３つの組を**確率空間（Probability Space）**という.

確率の性質：

(1) $P(A^c) = 1 - P(A)$
(2) $A \subset B$ ならば $P(A) \leq P(B)$
(3) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
(4) $A_i \in \mathcal{F}$ $(i = 1, 2, \cdots)$ ならば

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} P(A_n)$$

(5) $A_1 \supset A_2 \supset \ \cdots \in \mathcal{F}$ ならば

$$P\left(\bigcap_{n=1}^{\infty} A_n\right) = P\left(\lim_{n\to\infty} A_n\right) = \lim_{n\to\infty} P(A_n)$$

(6) $A_1 \subset A_2 \subset \ \cdots \in \mathcal{F}$ ならば

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

(7) 単調でなくても，$\lim A_n = A$ ならば，

$$\lim_{n\to\infty} P(A_n) = P(A)$$

（注）(4) の性質を**確率の劣加法性**という．また，(5), (6), (7) の性質を**確率の連続性**という．なお，集合列 $\{A_n\}$ の極限 $\lim A_n$ の定義については，次節を参照のこと．

(4) の証明：$B_1 = A_1,\ B_k = A_k \cap A_1^c \cap \cdots \cap A_{k-1}^c = A_k - \bigcup_{i=1}^{k-1} A_i\ (k = 2, 3, \cdots)$ とすれば，$B_1, B_2, \cdots$ は互いに排反であり，また，$B_k \subset A_k$ であるから，次のことが成り立つ．

$$\bigcup_{k=1}^{\infty} A_k = \bigcup_{k=1}^{\infty} B_k, \qquad P\left(\bigcup_{k=1}^{\infty} A_k\right) = P\left(\bigcup_{k=1}^{\infty} B_k\right) = \sum_{k=1}^{\infty} P(B_k) \leq \sum_{k=1}^{\infty} P(A_k)$$

(6) の証明：集合列 $\{A_n\}$ は単調増加であるから，この集合列の極限が存在して

$$\lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n$$

となる（次節参照）．また，

$$B_1 = A_1, \qquad B_i = A_i - A_{i-1} \quad (i \geq 2)$$

とおくと，$B_1,\ B_2,\ \cdots$ は互いに排反であり，

$$\bigcup_{i=1}^{n} B_i = \bigcup_{i=1}^{n} A_i = A_n, \quad \bigcup_{i=1}^{\infty} B_i = \bigcup_{i=1}^{\infty} A_i = \lim_{n\to\infty} A_n$$

となる．したがって，

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

(5) の証明： $C_n = A_n^c$ とおくと，集合列 $\{C_n\}$ は単調増加であるから，(6) より

$$P\left(\bigcup_{n=1}^{\infty} C_n\right) = P\left(\lim_{n\to\infty} C_n\right) = \lim_{n\to\infty} P(C_n)$$

となる．また，

$$\bigcup_{n=1}^{\infty} C_n = \bigcup_{n=1}^{\infty} A_n^c = \left(\bigcap_{n=1}^{\infty} A_n\right)^c \qquad \text{（de Morgan の法則）}$$

であるから，

$$\begin{aligned} P\left(\bigcup_{n=1}^{\infty} C_n\right) &= 1 - P\left(\bigcap_{n=1}^{\infty} A_n\right) \\ &= \lim_{n\to\infty} P(C_n) = \lim_{n\to\infty} P(A_n^c) = \lim_{n\to\infty}(1 - P(A_n)) \\ &= 1 - \lim_{n\to\infty} P(A_n) \end{aligned}$$

となる．

(7) の証明： $B_n = \cap_{k=n}^{\infty} A_k, C_n = \cup_{k=n}^{\infty} A_k$ とおくと，$\{B_n\}$ は単調増加，$\{C_n\}$ は単調減少であるから，(5), (6) により，次のことが成り立つ．

$$P(A) = P(\lim_{n\to\infty} B_n) = \lim_{n\to\infty} P(B_n) = P(\lim_{n\to\infty} C_n) = \lim_{n\to\infty} P(C_n)$$

ここで，B_n, C_n の定義から，$B_n \subset A_n \subset C_n$ が成り立つから，

$$P(B_n) \le P(A_n) \le P(C_n), \qquad \lim_{n\to\infty} P(B_n) = P(A) \le \lim_{n\to\infty} P(A_n) \le \lim_{n\to\infty} P(C_n) = P(A)$$

となる．このことから，結果が得られる．

2．集合列の極限

以下，$A_1, A_2, \cdots, A_n, \cdots$ を集合の無限列とする．

(1) 単調増加列の場合　$A_1 \subset A_2 \subset \cdots \subset A_n \subset \cdots$ のとき，この集合列の極限を

$$\lim_{n\to\infty} A_n = \cup_{k=1}^{\infty} A_k$$

で定義する．

(2) 単調減少列の場合　$A_1 \supset A_2 \supset \cdots \supset A_n \supset \cdots$ のとき，この集合列の極限を

$$\lim_{n\to\infty} A_n = \cap_{k=1}^{\infty} A_k$$

で定義する．

(3) 単調列でない場合

・下極限　$A_n, A_{n+1}, \cdots$ の共通部分を

$$B_n = \cap_{k=n}^{\infty} A_k$$

で定義すると，$\{B_n\}$ は単調増加，すなわち，$B_1 \subset B_2 \subset \cdots \subset B_n \subset \cdots$ となり，この $\{B_n\}$ の極限集合を

$$\lim_{n\to\infty} \inf A_n = \lim_{n\to\infty} B_n = \lim_{n\to\infty} (\cap_{k=n}^{\infty} A_k) = \cup_{n=1}^{\infty} (\cap_{k=n}^{\infty} A_k)$$

で定義して，$\{A_n\}$ の**下極限（最小極限）集合**という．

・**上極限**　$A_n, A_{n+1}, \cdots$ の和集合を

$$C_n = \cup_{k=n}^{\infty} A_k$$

で定義すると，$\{C_n\}$ は単調減少，すなわち，$C_1 \supset C_2 \supset \cdots \supset C_n \supset \cdots$ となり，この $\{C_n\}$ の極限集合を

$$\lim_{n\to\infty} \sup A_n = \lim_{n\to\infty} C_n = \lim_{n\to\infty} (\cup_{k=n}^{\infty} A_k) = \cap_{n=1}^{\infty} (\cup_{k=n}^{\infty} A_k)$$

で定義して，$\{A_n\}$ の**上極限（最大極限）集合**という．

・**一般の集合列の極限**　すべての n について $B_n \subset C_n$ であるから，

$$\lim_{n\to\infty} B_n = \lim_{n\to\infty} \inf A_n \subset \lim_{n\to\infty} C_n = \lim_{n\to\infty} \sup A_n$$

が成立する．特に，

$$\lim_{n\to\infty} \inf A_n = \lim_{n\to\infty} \sup A_n$$

のとき，$\{A_n\}$ は収束するといい，極限集合を $\lim A_n$ とかく．

（注 1）　一般に，$\liminf A_n$ と $\limsup A_n$ は異なる．例えば，

$$A_{2n-1} = [0, 1/2), \qquad A_{2n} = [1/2, 1] \qquad (n = 1, 2, \cdots)$$

と定義すれば，すべての n に対して，

$$\cap_{k=n}^{\infty} A_k = \phi, \qquad \cup_{k=n}^{\infty} A_k = [0, 1]$$

であるから，

$$\liminf A_n = \cup_{n=1}^{\infty} (\cap_{k=n}^{\infty} A_k) = \phi, \qquad \limsup A_n = \cap_{n=1}^{\infty} (\cup_{k=n}^{\infty} A_k) = [0, 1]$$

（注 2）

$$(a) \quad \liminf A_n = \cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_k$$

$a \in \liminf A_n$ ならば，ある m に対して $a \in \cap_{k=m}^{\infty} A_k$ である．したがって，$m \leq n$ ならば，$a \in A_n$ であるから，a は有限個の A_n を除けばすべての A_n の要素である．逆に，a が有限個の A_n を除いてすべての A_n の要素であるとする．除外される A_n のうち，番号の最も大きいものを A_{m-1} とすると，$n \geq m$ ならば $a \in A_n$．ゆえに，$a \in \cap_{k=m}^{\infty} A_k$．ゆえに，$a \in \liminf A_n$．すなわち，$\liminf A_n$ は，ある番号から先のすべての A_n に含まれるもの全体からなる集合である．

$$(b) \quad \limsup A_n = \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k$$

$a \in \limsup A_n$ ならば，いかなる m に対しても，$a \in \cup_{k=m}^{\infty} A_k$ である．したがって，$m \le n$ に対して，$a \in A_n$ なる A_n が必ずあるから，a は無限に多くの A_n に含まれる．逆に，a が無限に多くの A_n の要素ならば，m がいかなる値でも，$m \le n$ に対して，$a \in A_n$ なる A_n があるから，$a \in \cup_{k=m}^{\infty} A_k$．ゆえに，$a \in \limsup A_n$．すなわち，$\limsup A_n$ は無限に多くの A_n の要素全体からなる集合である．

$\liminf A_n$ に含まれる要素は，必ず $\limsup A_n$ に含まれる．

問 1　集合列を

$$A_{2n-1} = [-1/2 - 1/n, 1 + 1/n], \qquad A_{2n} = [-1 - 1/n, 1/2 + 1/n]$$

で定義するとき，次のことを示せ．

$$\liminf A_n = [-1/2, 1/2], \qquad \limsup A_n = [-1, 1]$$

問 2　集合列 $\{A_n\}$ が単調増加あるいは単調減少のとき，$\liminf A_n = \limsup A_n$ となることを示せ．

問 3　集合列 $\{A_n\}$ が互いに素（disjoint）ならば，$\lim A_n = \phi$ となることを示せ．

３．ベイズの定理：原因の確率，事後確率の計算方式を与える．

全事象 Ω が n 個の互いに排反な事象により分割されているとき，すなわち，

$$\Omega = \bigcup_{i=1}^{n} A_i, \qquad A_i \cap A_j = \phi \quad (i \ne j)$$

の場合，任意の事象 B に対して，

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

が成り立つ．これを**全確率の公式という．**

このとき，ある事象 B が起きたという情報が与えられたときに A_i が起きる**条件付き確率**は

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

で計算される．これを**ベイズの定理**という．

(応用例 1) ある病気の検査法は完全でなく，陰性でも感染している可能性が若干あり，陽性でも感染していないことがあり得る．このとき，陽性の結果が出た被検者が本当に感染している可能性は？

(応用例 2) 入試合格者の中には，実力通りに合格した者の他に，実力がなくても幸運にも合格する者もいると言われている．このとき，合格者の中で本当に実力があって合格した者は何割？

４．事象の従属性と独立性

事象 A と B は,

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

が成り立つとき，互いに独立であるという．独立性の条件は

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

と同等である．ただし，$P(A) > 0$.

独立性の意味： A が起きたという情報が B の生起に何ら影響しない（情報としての価値がない）.

(例) 52 枚のトランプのカードから 1 枚をランダムに抜くとき，スペードを抜くという事象と絵札を抜くという事象は互いに独立である．

一般に，事象の無限列 $A_1, \cdots, A_n, \cdots$ が互いに独立とは，この中の任意の有限個の事象 $A_{i_1}, \cdots, A_{i_k}$ に対して,

$$P\left(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}\right) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_k})$$

（注） 有限個の事象の独立性についても同様に定義される．ペアごとに独立でも，互いに独立になるとは限らない．

５．Borel-Cantelli の補題

事象 A を

$$A = \limsup A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$$

で定義する.（事象 A は，A_n が無限回起きる事象であり，$P(A) = P(\limsup A_n) = P(A_n \text{ i.o.})$ とも表わされる）．このとき，次のことが成り立つ．

(i) $\displaystyle\sum_{k=1}^{\infty} P(A_k) < \infty$ ならば $P(A) = P(A_n \, \text{i.o.}) = 0$

(ii) $\displaystyle\sum_{k=1}^{\infty} P(A_k) = \infty$ で，$A_1, A_2, \cdots, A_n, \cdots$ が互いに独立であれば

$$P(A) = P(A_n \, \text{i.o.}) = 1$$

（証明）
(i) A の定義から，任意の n に対して,

$$A \subset \bigcup_{k=n}^{\infty} A_k$$

であるから,

$$P(A) \le \sum_{k=n}^{\infty} P(A_k) \quad \to \quad 0 \qquad (n \to \infty)$$

を得る．
(ii) de Morgan の法則と確率の連続性を使って，

$$P(A^c) = P\left(\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} A_k^c\right) = \lim_{n\to\infty} P\left(\bigcap_{k=n}^{\infty} A_k^c\right)$$

を得る．ここで，独立性の仮定から

$$P\left(\bigcap_{k=n}^{\infty} A_k^c\right) = \lim_{m\to\infty} P\left(\bigcap_{k=n}^{m} A_k^c\right) = \lim_{m\to\infty}\prod_{k=n}^{m} P(A_k^c) = \lim_{m\to\infty}\prod_{k=n}^{m}(1 - P(A_k)) \qquad (*)$$

となる．さらに，不等式 $1 - x \leq e^{-x}$ $(x = P(A_k))$ が成り立つことを使って，

$$(*) \leq \lim_{m\to\infty}\prod_{k=n}^{m} e^{-P(A_k)} = \exp\left(-\sum_{k=n}^{\infty} P(A_k)\right) = e^{-\infty} = 0$$

を得る．したがって，

$$P\left(\bigcap_{k=n}^{\infty} A_k^c\right) = 0, \qquad P(A^c) = \lim_{n\to\infty} P\left(\bigcap_{k=n}^{\infty} A_k^c\right) = 0, \qquad P(A) = 1 - P(A^c) = 1$$

を得る．

・応用　ゆがみのないサイコロを無限回投げたとき，１から６の目がそれぞれ無限回出る確率は１であることを示す．
(解)　例えば，n 回目に１が出る事象を A_n とすれば，$A_1, A_2, \cdots$ は独立な事象の系列と考えることができる．そして，$P(A_n) = 1/6$ $(n = 1, 2, \cdots)$ であるから，

$$\sum_{k=1}^{\infty} P(A_k) = \infty$$

となる．したがって，Borel-Cantelli の補題 (ii) により，$P(A_n\ \text{i.o.}) = 1$，すなわち，無限に多く１の目が出る確率は１である．

６．確率変数の導入

事象を数値に変換することにより，数学的操作可能性を高めることを考える．そのために，確率変数を導入する．

確率変数 $X = X(\omega)$ **の定義：**全事象 Ω で定義され，実数の集合 R^1 に値をとる**可測関数 (measurable function)**，すなわち，任意の実数 a に対して

$$\{\omega:\ X(\omega) \leq a\} \in \mathcal{F}$$

可測性の意味合い：確率変数は事象を数値に変換したものであるが，どのような変換でも確率変数になるわけではなく，確率が定義されるような変換（可測変換）に限定する．

確率変数は詳しく書けば $X(\omega)$ となるが，ω を省略して単に X と書くことが多い．

確率変数の利点:事象から数値に変換された確率変数は，数学的な操作性に優れている．すなわち，確率分布，分布関数，密度関数などの概念や，それらに基づく期待値，分散などを定義することができ，確率的な現象を多面的に考察することが可能となる．

7．分布関数

確率変数 X に対して，X が x 以下の値をとる確率を x の関数とみなしたもの，すなわち，

$$F(x) = P(X \le x)$$

を確率変数 X の分布関数という．分布関数は X の性質を特徴付けるものであり，分布関数を推測するのが統計学の主要な目的の１つである．

分布関数の性質：

(D1) 任意の x に対して $0 \le F(x) \le 1$ であり，かつ，

$$F(-\infty) = 0, \quad F(+\infty) = 1$$

(D2) 単調非減少，すなわち，$x < y$ ならば $F(x) \le F(y)$
(D3) 右連続，すなわち，任意の x に対して $F(x) = F(x+)$ である．しかし，必ずしも左連続ではない．

(D3) の証明：

$$A_n = \left\{\omega : X(\omega) \le x + \frac{1}{n}\right\}$$

とおくと，$\{A_n\}$ は単調減少な集合列となり，かつ，

$$P(A_n) = F\left(x + \frac{1}{n}\right), \quad \lim_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} A_n = \{\omega : X(\omega) \le x\}$$

となる．したがって，確率測度の連続性より，

$$\begin{aligned} F(x) &= P\left(\lim_{n\to\infty} A_n\right) = \lim_{n\to\infty} P(A_n) \\ &= \lim_{n\to\infty} F\left(x + \frac{1}{n}\right) = F(x+) \end{aligned}$$

（問）分布関数の左連続性は必ずしも成り立たないことを示せ．