応用数学の教材作成（確率と統計）

石　　信　一*

Teaching materials of Probability and Statistics

Shin-ichi ISHI

要　旨

応用数学の授業改善の一方策として、シラバス用の授業教材の作成を試みた。

Abstract

We have made and showed a lecture-note about Probability and Statistics according to syllabus to imptrove upon the scheme of teaching in Applied Mathematics.

第１週　確率

確率とは

確率論は、「賭け」の問題に一つの起源をもつ。言い換えれば、それは、偶然に関する考察から生まれた。ここにいう“偶然”は漠然とした偶然（思い付き、予期せぬこと等）による出来事ではなく、次のような制限された意味である：コイン投げにおいて、「表」あるいは「裏」がでる、という事柄は偶然に依る。しかし、この事柄を問題にするときには、一定条件で何回でも繰り返しが出来るし、または、少なくとも実験可能であると考えることができる、ということを暗に含んでいる（無限回投げ続けることはできないが、そう考えることはできる)。従って、偶然の事柄が全て考察の対象になるということではない。コイン投げでの「表」あるいは「裏」がでるという事柄のように、偶然のもとで生まれる個々の結果を**事象**(event)と呼ぶ。こうして、確率とは、ある事象が起こる確からしさの度合（偶然性）を数量化したもの、であるといえる。事象Aの起こる確率をP(A)と書く。事象を生み出す一回々の実験・観測・調査等のことを**試行**(trial)と呼ぶ。

1つの集合Ω(universal set)を決めて、その要素（元）や部分集合について考察するとき、Ωを空間(space)という。集合と空間は同じ内容を表す。集合の‘要素(元)’は、空間では‘点’という。一回のコイン投げの試行において、表あるいは裏のでる事象の集まりを{表, 裏}=Ωで表し、Ωを**全事象**あるいは**基礎空間**と呼ぶ。特に、基礎空間Ωの点（要素）を**素事象**（**根元事象**；elementary event）と呼ぶ（標本空間、見本空間Ωの場合には、その点（要素）を標本点、見本点と呼ぶことになる)。Ωの部分集合(subset)は、

$\{\phi\}$=(空事象), { 表 }, { 裏}, {表, 裏} = Ω

の4つである。ϕとΩを考慮するのは、後述する集合演算に対して閉じることを要請するからである。一般に、Ωの部分集合の全体からなる集合を2^{Ω}で表し、Ωのベキ集合(power set)という。上記コイン投げのベキ集合は、2^{Ω} = { ϕ,(表),(裏),(表,裏)} である。集合を要素とする集合（単に、「集合の集合」という表現になる）を集合族(family of set)という。事象はΩの部分集合として表現されるから、確率$P(A)$は、集合を定義域とする関数とみなせる。この関数を、特に集合関数(set function)という。表あるいは裏がでる事象の起こる確率をそれぞれ数値p,q（但し、$p+q=1$「全確率は1である」）とすれば、確率$P(A)$を、

$$P(\{表\})=p, \quad P(\{裏\})=q$$

と書く。次に、表・裏のでる確率が**同程度に確からしく起こることを仮定すれば、**

＊　助教授　一般教科

$$P(\{表\}) = P(\{裏\}) = \frac{1}{2}$$

と書ける。上式をより機能的な表式に書き替えたい。そこで、表がでたら1、裏がでたら0、即ち、X(表)=1, X(裏)=0、なる関数Xを導入する。{表,裏}=Ω の集合が定義域で、{0,1}の集合が値域とする関数である。これによって、上式を、

$$P(X=1) = P(X=0) = \frac{1}{2} \qquad (1)$$

書く。上式ではXは、関数 $P(X)$ の変数の働きである。そこで、関数Xを特に**確率変数**（言い方を代えて、Xの値が偶然性によるという意味で確率変数）、$P(X)$ を**確率関数**と呼ぶ。関数 $P(X)$ は、Xの取りうる値に対して全確率1がどのように分布しているかを示すものである、即ち、(1) 式は X に対する確率分布を与えている。

(Ω,2^Ω,P) の組によって一回のコイン投げで表・裏のでる確率を示すモデルができた。この (Ω,2^Ω,P) の組を**確率空間**と呼ぶ。その理解のために幾つかの具体例を挙げよう：

[例1] サイコロを一回投げる試行とする。その試行に対する確率空間 (Ω ,2^Ω,P) を作れ。

[解]

(a) 基礎空間 $\Omega = \{1,2,3,4,5,6\}$

(b) 部分集合 2^Ω; ϕ, {1},{2},{3},{4},{5},{6}, {1,2},{1,3},···,{5,6}, {1,2,3},···,{4,5,6}, {1,2,3,4}, ···, {3,4,5,6}, {1,2,3,4,5},···,{2,3,4,5,6}, {1,2,3,4,5,6}=Ω； 部分集合の数 2^6=64個。

(c) 確率 $P(X) = \frac{1}{6}$, $(X = 1,2,3,4,5,6)$

但し、1から6の目がでるのは同程度に確からしく起こると仮定した。

事象が部分集合で表されることを[例1]の部分集合を用いて再度説明する。サイコロを振って（試行の結果）2の目がでたとする。この結果は、事象に関して幾つかの解釈が可能となる。例えば、2以下の目のでる事象、2以上の目のでる事象、偶数の目のでる事象等々である。対応する部分集合は、$\{1,2\}, \{2,3,4,5,6\}, \{2,4,6\}$ となる。2の目のみに注目する事象は $\{2\}$ で、これを特に素事象という。

[例2] コイン投げを二回続けて行なったとき、表・裏に注目する試行とする。（一回々の試行が、互いに何の関係もないとする。これを**独立試行**という。）この試行に対する確率空間 (Ω,2^Ω,P) を作れ。

[解]

(a) 基礎空間；Ω={ (表,表),(表,裏),(裏,表),(裏,裏) }

一回の試行による基礎空間を Ω′={表,裏} とすれば、Ω は、Ω′ × Ω′={表,裏} × {表,裏} と書ける。

Ω′ \ Ω′	表	裏
表	(表,表)	(表,裏)
裏	(裏,表)	(裏,裏)

これを Ω′ と Ω′ の直積(direct product)という。

(b) 部分集合 2^Ω; ϕ, {(表,表)} ,{(表,裏)},{(裏,表)} ,{(裏,裏)} , {(表,表),(表,裏)} ,{(表,表),(裏,表)} ,{(表,表),(裏,裏)} ,{(表,裏),(裏,表)},{(表,裏),(裏,裏) } ,{(裏,表),(裏,裏)} ,{(表,表),(表,裏),(裏,表)} ,{(表,表),(表,裏),(裏,裏)} , {(表,表),(裏,表),(裏,裏) } ,{(表,裏),(裏,表),(裏,裏)} , {(表,表),(表,裏),(裏,表),(裏,裏)} =Ω ;

部分集合の総数 2^4=16

(c) 試行の独立を仮定しているので、

確率 $P(\{X,Y\}) = \frac{1}{4}$, (X,Y=表 あるいは 裏)。

一般に、n 個の要素からなる集合の全ての部分集合の数は 2^n個 である。

[例3] コイン投げを二回繰り返し、表がでる回数に注目する。この試行に対する確率空間 (Ω,2^Ω,P) を作れ。

[解]

(a) 基礎空間；$\Omega = \{2,1,0\}$

(b) 部分集合 2^Ω; ϕ, {2},{1},{0},{2,1},{2,0},{1,0}, {2,1,0}=Ω

部分集合の数 2^3=8 個。

(c) 試行の独立を仮定しているので、

$P(\{2\})=P(\{表\})P(\{表\})=\frac{1}{4}$

$P(\{1\})=P(\{表\})P(\{裏\})+P(\{裏\})P(\{表\})=\frac{1}{2}$

$P(\{0\})=P(裏)P(裏)=\frac{1}{4}$

上記コイン投げのように、結果が二通りの事象で独立な試行を**ベルヌイ試行**と呼ぶ。ベルヌイ試行を n 回行なった結果を**長さ n のベルヌイ列**という.

ベルヌイ試行を n 回行なった結果（1回の試行で、事象 A の起こる確率を p、起こらな確率を $q = 1 - p$ とする。但し、$p + q = 1$）、事象 A が x 回起こる確率は、

$$P(x) = {}_nC_x p^x q^{n-x} \qquad (2)$$

と表せる。x を確率変数として、$x = 1, 2, \cdots$のときの確率が、上式で与えられる離散的確率関数を**二項分布**[1]という（(2) は二項展開式の各項になっている)。

集合・事象の記号の説明

このように、確率は部分集合に対して定義することになる。事象の和、積等の演算を、集合の演算として取り扱うことになる。従って、ここに集合の記号と幾つかの演算規則を要約しよう（大文字 A,B 等は集合、小文字 a,b 等は集合の要素(元)を表す)：

a)a$\in$ A(a$\notin$A)； a は A に含まれる（含まれない)。

b)A$\subset$B(A$\not\subset$B)； A は B に含まれる（含まれない)。

c)A ＝B；A$\subseteq$BかつA$\subseteq$B。

d)A$\cup B = \{x \mid a \in A$ and $b \in B\}$； A と B の和集合（合併集合)。

e)A$\cap B = \{x \mid a \in A$ or $b \in B\}$； A と B の積集合（共通集合)。

f)$A^c (= \overline{A}) = \{x \mid x \notin A\}$； A の補集合（余集合)。

g)A$\cap$A$^c = \phi$ ； ϕ は空集合。

h)A$\cup$A^c=Ω； Ω は全集合。

i)A$\cap$B=ϕ; A と B は互いに素（排反）である。

上記 d) - i) の集合の記号の説明を事象の言葉に翻訳しておこう；

d)' $A \cup B$； A と B のいずれかが起こるという事象（和事象)。

e)' $A \cap B$； A と B のどちらも起こるという事象（積事象)。

f)' $A^c(= \overline{A}) = \{x|x \notin A\}$ ；A が起こらないという事象(余事象)。

g)' A $\cap A^c = \phi$;　ϕ は空事象。

h)' A$\cup$A^c = Ω； Ω は全事象。

i)' $A \cap B = \phi$；A と B は同時には起こらない事象（排反事象)。

次に、集合演算の公式を上げておこう。

1)　(A$\cup$B)$\cup$C=A$\cup$(B$\cup$C)（結合則）
　　(A$\cap$B)$\cap$C=A$\cap$(B$\cap$C)

2)　$(\bigcup_{i=1}^{n} A_i) \cap B = \bigcup_{i=1}^{n} (A_i \cap B)$（分配則）
　　$(\bigcap_{i=1}^{n} A_i) \cup B = \bigcap_{i=1}^{n} (A_i \cup B)$

3)　$(\bigcup_{i=1}^{n} A_i)^c = \bigcap_{i=1}^{n} A_i^c$（ド・モルガン則）
　　$(\bigcap_{i=1}^{n} A_i)^c = \bigcup_{i=1}^{n} A_i^c$

上記の確率空間では、集合の要素は全て有限であった。しかし、コインやサイコロを投げ続けような試行を考察するときは、要素の数が有限とは限らない。従って、要素の数が無限の場合にも確率が定義できるように集合族の概念を拡張しておく必要がある。

Ω の部分集合からなる集合族$\mathbf{B}$で、次の三つの条件

i)　$\Omega \in \mathbf{B}$
ii)　$A \in \mathbf{B} \implies A^c \in \mathbf{B}$
iii)　$A_i \in \mathbf{B} \implies \bigcup_{i=1}^{\infty} A_i \in \mathbf{B}$

を満たすとき、$\mathbf{B}$ は Ω における σ-加法的集合族あるいは完全加法族、単に σ-加法族 という[2]。記号 'σ-' は、可算無限（自然数と 1 ： 1 の対応がつくこと）の意味である。添字 $_i$ が有限であれば、有限集合族という。(無論、有限は可算無限に含まれる。) $\mathbf{B}$ の集合の意味は、集合演算（$\cup, \cap, ^c$等）を有限または可算無限回行なった結果はまた $\mathbf{B}$ の集合に含まれる、ということである。このことは、事象の組み合わせによって新たな複雑な事象を取り扱うための集合論での言葉を用意している。

上記 i) - iii) から導かれる、次の二つの事柄を付け加えておこう。i) と ii) から、

$$A = \Omega \longrightarrow A^c = \Omega^c = \phi \in \mathbf{B}$$

となる。更に、iii) から、

$$A_i \in \mathbf{B} \longrightarrow \bigcap_{i=1}^{\infty} A_i \in \mathbf{B}.$$

確率の定義 (公理論的立場)

全事象 Ω とその集合族 $\mathbf{B}$ (上記i)-iii) を満たす) が与えられ, $\mathbf{B}$ では集合の演算が閉じている。このとき, $\mathbf{B}$ に属する各事象（部分集合）A に対して、実数値 $P(A)$ が対応し、

I)　$0 \leq P(A)$

[1] 二項展開式:$(a+b)^n = a^n + {}_nC_1 a^{n-1} b + \cdots + {}_nC_x a^x b^{n-x} + \cdots + b^n$

[2] ボレル集合族：Ω の部分集合のうち、どれが開集合であるかが決まっているとき、Ω の開集合を全て含む最小のσ-加法的集合族をボレル集合族、それに属する集合をボレル集合という。

II)　$P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i) \quad (A_i \cap A_j \neq \phi)$

III)　$P(\Omega) = 1$

が満たされるとき、$P(A)$ を事象 A の起こる確率、単に、"事象 A の確率" という。

I) は、確率値が非負で、その上限値は III) から 1 である ($0 \leq P(A) \leq 1$)。II) の性質を「完全加法性」という。これによって、色々な組合せの確率が求められる。一般的に、$(\Omega, \mathbf{B}, \mathrm{P})$ の組を確率空間と呼ぶ。I)-III) で確率を定義する立場を公理論的立場と呼ぶ。歴史的には、これに先行する確率の定義として次の二つがある：

[A] ある試行の結果、n 個の場合があって、それらは同等に確からしく起こるとする。事象 A の場合が r 個あるとき、

$$(\text{事象A の起こる確率}) = \frac{(\text{事象} A \text{の場合の数})}{(\text{全体の場合の数})} = \frac{r}{n}$$

という。これを、「先験的確率（ラプラスの確率）」という。また、試行の結果が場合の数として表現できないとき、線分や空間の領域で表される事象を「幾何学的確率」と呼ぶ。

[B] ある試行を n を繰り返し行い、事象Aが起こった回数が r 回であるとき、$\frac{r}{n}$ を相対度数という。

$$n \longrightarrow \infty \text{　として、} \frac{r}{n} \longrightarrow p \ (\text{実数値})$$

は、事象Aの起こる確率という。言い換えれば、確率は試行回数が無限に多くなったときの相対度数の極限値といえる。これを「経験的確率」という。公理論的立場は、これらの定義の一般化・拡張になっている。

第2週　確率の計算公式

前節の公理 I) − III) から導ける幾つかの確率の計算公式 (1) - 7)) を以下に説明しよう。

1) $P(A^c) = 1 - P(A)$　（余事象の定理）

証明.

$$A + A^c = \Omega, \quad A \cap A^c = \phi$$

として，II), III) から，

$$1 = P(\Omega) = P(A) + P(A^c) \quad \therefore P(A^c) = 1 - P(A)$$

2) $P(\phi) = 0$

証明.

$$\Omega^c = \phi$$

だから，1) から

$$P(\Omega^c) = P(\phi) = 1 - P(\Omega) = 1 - 1 = 0$$

3) $A \supseteq B \longrightarrow P(A) \geq P(B)$

証明.

$$A = B + (A - B)$$

$$\longrightarrow P(A) = P(B) + P(A - B)$$

$$\longrightarrow P(A) \geq P(B) \quad (1 \geq P(A - B) \geq 0)$$

4) 和公式（加法定理）

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

証明.

$$A = (A \cap B) \cup (A \cap B^c), \ B = (A \cap B) \cup (A^c \cap B)$$

と表せるので，

$$A \cup B = (A \cap B) \cup (A \cap B^c) \cup (A^c \cap B)$$

と書ける。II) により，

$$P(A \cup B) = P(A \cap B) + P(A \cap B^c) + P(A^c \cap B)$$

同様に，

$$P(A \cap B^c) = P(A) - P(A \cap B)$$

$$P(A^c \cap B) = P(B) - P(A \cap B)$$

よって，

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

を得る。

上式の特別の場合、$P(A \cap B) = P(\phi) = 0$ (「AとBは排反である」) のとき、

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

と書ける。この性質を「有限加法性」という。

この一般化した表式は，

$$P(\bigcup_{i=1}^{n} A_i) = P(\sum_{i=1}^{\infty} A_i) - \sum_{i<j}\sum P(A_i \cap A_j) +$$

$$\sum_{i<j<k}\sum\sum P(A_i \cap A_j \cap A_k) -$$

$$\cdots+(-1)^{n-1}\sum_{i<j<\cdots<n}\cdots\sum P(\bigcap_{i=1}^{n} A_i)$$

と書ける。

5) 積公式（乗法定理）

事象Aが起こったという条件の下で事象Bの起こる確率（「**条件付き確率**」）をP(B | A) あるいは P_A (B) で表す。

［例］サイコロを振る試行とする。Aは偶数の目のでる事象 A={2,4,6}，Bは2の目のでる事象 B={2} とする。偶数の目がでることがわかったとして，2の目がでる確率を求めよ。

［解］P({2,4,6})=$\frac{1}{2}$, P({2})=$\frac{1}{6}$だから,

$$\frac{(\text{2の目のでる確率})}{(\text{偶数の目のでる確率})}=\frac{\frac{1}{6}}{\frac{1}{2}}=\frac{1}{3}=P_A(B)$$

「条件付き確率」の式を用いると次式（乗法定理）が成り立つ：

$$P(A\cap B)=P(A)P_A(B)=P(B)P_B(A) \quad (3)$$

証明.

全事象Ωを、A∩B、A∩B^c、A^c∩B、A^c∩B^c の4つの排反事象に分ける。各事象の起こる場合の数を a, b, c, d とすれば、全事象Ω、事象 A のそれらは、a+b+c+d=N, a+c である。事象 A∩B と A の確率は、

$$P(A\cap B)=\frac{b}{N},\quad P(A)=\frac{a+b}{N}$$

で、Aが既に起こったものとしてBの確率は、

$$P_A(B)=\frac{b}{a+b}$$

であるから、

$$P(A)P_A(B)=\frac{b}{N}=P(A\cap B)$$

を得る。右辺の等式は同様にして、

$$P(A\cap B)=\frac{b}{N},\quad P(B)=\frac{b+c}{N},\quad P_B(A)=\frac{b}{b+c}$$

であるから、

$$P(A\cap B)=P(B)P_B(A)=\frac{b}{N}$$

を得る。

こうして、$P(A\cap B)=P(A)P_A(B)$ から

$$P_A(B)=P(B)\longrightarrow P(A\cap B)=P(A)P(B)$$

と書けるとき，事象**A と B は独立である**、という。他方、独立でないときは**従属である**、という。一般に、$A_i(i=1,2,\cdots,n)$ について、

$$P(A_i\cap B_j)=P(A_i)P(B_j)$$

$$P(A_i\cap A_j\cap A_k)=P(A_i)P(A_j)P(A_k)$$

$$\cdots$$

$$P(\bigcap_{i=1}^{n} A_i)=P(A_1)P(A_2)P(A_3)\cdots P(A_n)$$

が成り立つとき、$A_i(i=1,2,\cdots,n)$ は独立であるという。

5)' 一般的な積公式（乗法定理）

$$P(\bigcap_{i=1}^{n} A_i)$$
$$=P(A_1)P_{A_1}(A_2)P_{A_1\cap A_2}(A_3)$$
$$\cdots P_{A_1\cap A_2\cap\cdots\cap A_{n-1}}(A_n)$$

6) $\Omega=\bigcup_{i=1}^{n} A_i$，$A_i\cap A_j\neq\phi$　のとき、

$$P(B)=P(A_1)P_{A_1}(B)+P(A_2)P_{A_2}(B)+\cdots$$

$$\cdots+P(A_n)P_{A_n}(B)=\sum_{i=1}^{n}P(A_i)P(B|A_i) \quad (4)$$

証明.

$$B=(B\cap A_1)\cup(B\cap A_2)\cup(B\cap A_3)\cup\cdots(B\cap A_n)$$

だから、

$$P(B\cap A_1)=P(B)P_B(A_1),\cdots$$

よって、II) により上式を得る。これを「**全確率の公式**」と呼ぶ。

7) 上式 6) の仮定の下で，即ち、n 個の互いに排反する事象 $A_1,A_2,\cdots,A_n$ があって、かつ、その中の1つが必ず起こるものとする。B を別の確率事象とすると、一般に、

$$P_B(A_i)=\frac{P(A_i)P(B|A_i)}{\sum P(A_i)P(B|A_i)}$$

この式を「**ベイズの公式**」(Bayers' formula) という。

証明.

条件付き確率の定義式から、積公式 5) と全確率の公式 6) から、

$$P_B(A_i)\overset{(3)}{=}\frac{P(A_i\cap B)}{P(B)}\overset{(4)}{=}\frac{P(A_i)P(B|A_i)}{\sum P(A_i)P(B|A_i)}$$

を得る。

上式の解釈として，A_i (i=1,2, $\cdots$) を事象 B の起こる原因の事象、P(A_i) を A_i の事前確率、P(A_i | B) を A_i の事後確率と呼ぶ。結果Bが起こったときに、その原因 A_i がどの程度であるかという推論の可能性を与えることになる。これが「ベイズの公式」の意義（統計的推論の先駆）である。

第３週　確率変数と確率分布

一般的に、確率分布が与えられたとき、分布の特性量（平均 (Mean value; μ) と分散 (Varience;σ^2)) の定義を与えよう。「平均」と「分散」を表す記号は、次のように幾通りかが用いられる：

$$\mu \ = \ E(x) \ = < x >$$

$$\sigma^2 = \ E((x-\mu)^2) = \ V(x) \ = < x^2 > - < x >^2$$

ここに、$E(\cdot)$ あるいは < ・ >は、確率変数・の期待値である（第５週参照)。それらの量の意味するところは、平均は分布の釣り合いの尺度（分布の安定性についての尺度)、分散は分布の平均値からのバラツキを示す尺度である。分散と同様な量、標準偏差は $\sqrt{分散} = \sqrt{\sigma^2} = \sigma$ で定義する。

確率変数 x として、その確率分布 $P(x)$ の平均と分散は、

$$\mu \ = \ \sum_{x=0}^{n} xP(x) \longrightarrow \int_{-\infty}^{\infty} xP(x)dx$$

$$\sigma^2 \ = \ \sum_{x=0}^{n} (x-\mu)^2 P(x) \longrightarrow \int_{-\infty}^{\infty} (x-\mu)^2 P(x)dx$$

と書ける。$\longrightarrow$ は、離散変数から連続変数に代わったときの表式である。特に、連続変数のときには、$P(x)$ を"確率密度関数"という（以下では $f(x)$ で表す)。実際の分散の計算では、

$$\sigma^2 = \sum_{x=0}^{n} x^2 P(x) - \mu^2 \longrightarrow \int_{-\infty}^{\infty} x^2 f(x)dx - \mu^2$$

を用いる。

二項分布

二項分布は、既に、「長さ n のベルヌイ列」の所で触れた。繰り返すと、n 回の独立な試行を行うとする。１回の試行で、結果は二通りしかなく（コイン投げのように表か裏）、かつ、必ずどちらかが起こるとする。n 回の試行のうち、表が x 回でる確率（表のでる確率を p、裏のでる確率は $q = 1-p$ として）は、

$$\overbrace{p \cdot p \cdots p}^{x}\overbrace{q \cdot q \cdots q}^{n-x} \longrightarrow p^x q^{n-x}$$

並べ方に依らない数（組み合わせ数）を考慮して、

$$_nC_x p^x q^{n-x} \quad (p+q=1,\ x=0,1,2,\cdots,n) \tag{5}$$

となる。表のでる回数 x に従う確率分布、つまり、離散的確率変数 x による確率分布を二項分布 (Binomial Distribution) といい、記号 $\mathbf{Bin(n,p)}$ で表す。表・裏の代わり、成功・失敗、当たり・外れ、on・off 等に置き換えると、いろいろな二項分布ができる。次に、二項分布の例を示そう。

［例１］　サイコロを n 回投げるとき、６の目のでる回数を x とする確率分布は、二項分布 $Bin(n, \frac{1}{6})$ で書ける。

［例２］　箱の中にリンゴ m 個、ナシ n 個が入っている。この中から１個ずつ10個取り出すとき、リンゴの取り出す回数を x とすると、x の確率分布は、二項分布 $Bin(10, \frac{m}{m+n})$ で書ける。

［例３］　製品の不良品の割合が m %である製品の中から、n 個取り出したときの不良品の個数を x とする確率分布は、二項分布 $Bin(n, \frac{m}{100})$ で書ける。

二項分布での特性量ー平均と分散の定義式ーは、

$$\mu = \sum_{x=0}^{n} xBin(n,p) = \sum_{x=0}^{n} x\ _nC_x p^x q^{n-x} \tag{6}$$

$$\sigma^2 = \sum_{x=0}^{n} (x-\mu)^2 Bin(n,p)$$

$$= \sum_{x=0}^{n} (x-\mu)^2\, _nC_x p^x q^{n-x} = \sum_{x=0}^{n} x^2\, _nC_x p^x q^{n-x} - \mu^2 \tag{7}$$

と書ける。そこで、**二項分布の平均** ($\mu = np$) **と分散** ($\sigma^2 = npq$) の表式を確かめよう：

まず、二項展開式

$$\sum_{x=0}^{n} {}_nC_x p^x q^{n-x} = (p+q)^n$$

両辺を p について微分すれば、

$$\sum_{x=0}^{n} x\ _nC_x p^{x-1} q^{n-x} = n(p+q)^{n-1} \tag{8}$$

両辺に p をかけると、

$$\sum_{x=0}^{n} x\ {}_nC_x p^x q^{n-x} = np(p+q)^{n-1} \overset{p+q=1}{=} np$$

左辺は、平均の定義式(6)そのものになっている。こうして、$\mu = np$ を得る。

分散を求めるには、$< x^2 >$ の表式が必要です。そこで、上式 (8) を更に p について微分して、

$$\sum_{x=0}^{n} x(x-1)_n C_x p^{x-2} q^{n-x} = n(n-1)(p+q)^{n-2}$$

(両辺)$\times p^2$ によって、

$$\sum_{x=0}^{n} x^2{}_nC_x p^x q^{n-x} - \sum_{x=0}^{n} x_n C_x p^x q^{n-x}$$

$$=< x^2 > -\mu = n(n-1)p^2(p+q)^{n-1} - np$$

$$\overset{p+q=1}{=} n(n-1)p^2 - np$$

で、分散の定義式に代入すれば、

$$\sigma^2 =< x^2 > -\mu^2 = n(n-1)p^2 - np - (np)^2$$

$$= np(1-p) = npq$$

を得る。

標準偏差は、$\sqrt{\sigma^2} = \sigma = \sqrt{npq}$ となる。

[問題]　二項分布の平均と分散を定義式((6) と (7)) の変形から導け。

ポアソン分布

二項分布 $Bin(n,p)$ の計算は，n が大きくなる ($n \to \infty$) と大変厄介になる。そこで，$n \to \infty$ に対して二項分布に代る近似的な分布を考える。

今，二項分布の極限 $\lim_{n\to\infty} Bin(n,p)$ を $np = \lambda(= const.)$ として、

$$Bin(n,p) = {}_nC_x p^x q^{n-x}$$

$$= \frac{n(n-1)\cdots(n-x+1)}{x!}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{\lambda^x}{x!}\frac{n}{n}\frac{n-1}{n}\cdots\frac{n-x+1}{n}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-x}$$

$$\overset{n\to\infty}{\Longrightarrow} \frac{\lambda^x}{x!} 1\cdots 1\ e^{-\lambda} 1 \ \left(\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n = e^{-\lambda}\right)$$

$$= \frac{\lambda^x}{x!}e^{-\lambda} = Po(\lambda) \qquad (9)$$

ここに、離散確率変数 x の確率関数 $\frac{\lambda^x}{x!}e^{-\lambda}$ の分布をポアソン分布 (Poisson Distribution; $Po(\lambda)$) という。上式の導出からわかるように、ポアソン分布の特徴は、$np = \lambda(= constant)$ で $n \to \infty$ としているので、事象は稀にしか起こらない ($p \to 0$) ということになる。ポアソン分布は、次の例のように、放射性物質の崩壊過程のように、一定時間内に、稀に起こる事象を記述する。

[例 1]　Rutherford と Geiger の実験によると、放射性物質の崩壊を7.5秒の時間間隔で2608回観察したところ、崩壊粒子の総数10094個で、次のような度数分布を得た：

崩壊粒子数	0	1	2	3	4	5
度数(回数)	57	203	385	525	532	408
崩壊粒子数	6	7	8	9	10以上	合計
度数(回数)	273	139	45	9	16	2608

崩壊粒子数 x を確率変数として、崩壊粒子の平均値は$\lambda = \frac{10094}{2608} = 3.87$ であるから、ポアソン分布は

$$Po(\lambda) = \frac{\lambda^x}{x!}e^{-\lambda} = \frac{3.87^x}{x!}e^{-3.87}$$

崩壊粒子数	0	1	2	3
度数	54.40	210.52	407.36	525.50
崩壊粒子数	4	5	6	7
度数	508.42	393.52	253.82	140.33
崩壊粒子数	8	9	10以上	
度数	67.88	29.19	17.08	

[例 2]　体積 V の箱の中に N 個の気体分子が入っている。体積 V の中の微少体積を v として、その中の n 個の気体分子が入っているとする。v の中で 1 個の気体分子が見い出される確率は $\frac{v}{V}$ であるとする（系（箱の中）は熱平衡状態であると仮定する)。この時、分子数 n の確率分布 $f(n)$ は、二項分布 $f(n) = Bin(N, p = \frac{v}{V})$ で表せる。次のような極限をとれば、

$$f(n) = Bin(N, p = \frac{v}{V})$$

$$\overset{p=\frac{v}{V}\to 0,\ \frac{N}{V}=const.\ \ (V\to\infty)}{\Longrightarrow} \frac{(\overline{n})^n}{n!}e^{-\overline{n}}, \quad (\overline{n} = np)$$

ポアソン分布になる。

ポアソン分布の平均 ($\mu = np = \lambda$) **と分散** ($\sigma^2 = \lambda$) を確かめよう：

ポアソン分布の平均 ($\mu = np = \lambda$) と分散 ($\sigma^2 = \lambda$) は、二項分布の平均 ($\mu = np$) と分散 ($\sigma^2 = npq$) の極限から容易に求められる。

平均は ($\mu = np = \lambda$) で、分散は、

$$\sigma^2 = np(1-p) = np(1-\frac{\lambda}{n}) \overset{n\to\infty}{\longrightarrow} np = \lambda$$

である。これらの結果を前述の二項分布での導出で用いた微分法で確かめておこう。

まず、指数関数を展開して、

$$e^{\lambda}=\sum_{x=0}^{\infty}\frac{\lambda^x}{x!}$$

両辺をλについて微分すれば、

$$e^{\lambda}=\sum_{x=0}^{\infty}x\frac{\lambda^{x-1}}{x!} \tag{10}$$

(両辺)$\times\lambda$ とすると、

$$\lambda e^{\lambda}=\sum_{x=0}^{\infty}x\frac{\lambda^x}{x!}$$

(両辺)$\times e^{-\lambda}$ とすれば、

$$\lambda=\sum_{x=0}^{\infty}x\frac{\lambda^x}{x!}e^{-\lambda}=\sum_{x=0}^{\infty}xPo(\lambda)$$

右辺は、平均の定義式そのものになっている。こうして、$\mu=\lambda$ を得る。

分散を求めるには、上式 (10) を再び λ について微分して、

$$e^{\lambda}+\lambda e^{\lambda}=\sum_{x=0}^{\infty}x^2\frac{\lambda^{x-1}}{x!}$$

(両辺) $\times\lambda e^{-\lambda}$ によって、

$$\lambda+\lambda^2=\sum_{x=0}^{\infty}x^2\frac{\lambda^x}{x!}e^{-\lambda}=\sum_{x=0}^{\infty}x^2Po(\lambda)$$

で、分散の定義式に代入すれば、

$$\sigma^2=<x^2>-\mu^2=\lambda+\lambda^2-\lambda^2=\lambda$$

を得る。更に、標準偏差は、$\sqrt{\sigma^2}=\sigma=\sqrt{\lambda}$ となる。

[問題] ポアソン分布の平均と分散を定義式の変形から導け。

上記ではポアソン分布を二項分布の極限として示したが、次のような1つの確率過程（確率変数が時間 t にも依存する）として、それとは独立に導くことができる。

時間 $(0,t+\Delta t)$ の間に、x 個の粒子が観測される確率を $P_x(t+\Delta t)$ で表す。Δt の間に粒子が観測される確率を $\lambda\Delta t$ とする。その際、$\lambda\Delta t$ の確率は、Δt の時間内だけで決まり、その前後の影響を受けない。更に、Δt の間に 2 個以上の粒子が観測される確率は無視できる。これらの仮定の下で、確率 $P_x(t+\Delta t)$ は、次の２つの排反な事象の和で書ける：

（1）$(0,t)$ の間に $(x-1)$ 個の粒子が観測され、$(t,t+\Delta t)$ の間に 1 個観測される。この確率は $P_{x-1}(t)\times\lambda\Delta t$ と書ける。

（2）$(0,t)$ の間に x 個の粒子が観測され、$(t,t+\Delta t)$ の間に観測されない。この確率は $P_x(t)\times(1-\lambda\Delta t)$ と書ける。

即ち、

$$P_x(t+\Delta t)=P_{x-1}(t)\times\lambda\Delta t+P_x(t)\times(1-\lambda\Delta t).$$

整理して、

$$\frac{P_x(t+\Delta t)-P_{x-1}(t)}{\Delta t}=-\lambda(P_x(t)-P_{x-1}(t))$$

$\Delta t\longrightarrow 0$ のときには、

$$\frac{\partial P_x(t)}{\partial t}=-\lambda(P_x(t)-P_{x-1}(t))$$

となる。$x=0,1,\ldots,n$ に対して、

$$\frac{\partial P_0(t)}{\partial t} = -\lambda P_0(t) \tag{11}$$

$$\frac{\partial P_1(t)}{\partial t} = -\lambda(P_1(t)-P_0(t)) \tag{12}$$

$$\vdots = \vdots$$

$$\frac{\partial P_n(t)}{\partial t} = -\lambda(P_n(t)-P_{n-1}(t)). \tag{13}$$

これらの方程式を、初期条件 $t=0,\ P_0(t)=1$ のもとで解けば、

$$P_x(t)=\frac{(\lambda t)^x}{x!}e^{-\lambda t}$$

を得る。

第4週　正規分布

離散的確率変数 x とする二項確率分布 ${}_nC_xp^xq^{n-x}$ は、$n\to\infty$ のとき、連続的確率変数 x とする確率密度関数

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\left(\mu=np,\ \sigma^2=npq\right) \tag{14}$$

で近似できる（付録参照）。この分布をもつとき、連続的確率変数 x は正規分布 (14) 式に従うという。正規分布は二つのパタメーター μ と σ で特徴つけられるので $f(x)$ の代わりに、正規分布を $N(x;\mu,\sigma^2)$、あるいは、単に $N(\mu,\sigma^2)\equiv N(\mu,\sigma)$ と記す。

[問題] $g(x)=e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ に関して、

(1)　$g(x)$ の極値と変曲点を求めよ（グラフを描け）。

(2)　$g(x)$ の規格化因子を求めよ。

解.

(1) $g'(x)=-\dfrac{2(x-\mu)}{2\sigma^2}e^{-\frac{(x-\mu)^2}{2\sigma^2}}=0$

これより、$x=\mu$（極値）。

$$g''(x) = -\frac{1}{\sigma^2}\left(1 - \frac{(x-\mu)^2}{\sigma^2}\right)e^{-\frac{(x-\mu)^2}{2\sigma^2}} = 0$$

これより、$x = \mu \pm \sigma$ (変曲点)。

x		$\mu-\sigma$		μ	
$g'(x)$ の情報	↗		↗	極大	↘
$g''(x)$ の情報	下に凸	変曲点	上に凸		上に凸
$g(x)$		$e^{-\frac{1}{2}}$		1	

x	$\mu+\sigma$	
$g'(x)$ の情報		↗
$g''(x)$ の情報	変曲点	下に凸
$g(x)$	$e^{-\frac{1}{2}}$	

(2) $\displaystyle\int_{-\infty}^{\infty} Ne^{-\frac{(x-\mu)^2}{2\sigma^2}}dx = 1$

変数変換 $\dfrac{x-\mu}{\sqrt{2}\sigma} = t$ とおくと、$dx = \sqrt{2}\sigma dt$、

$$N\int_{-\infty}^{\infty} e^{-t^2}\sqrt{2}\sigma dt = N\sqrt{2}\sigma\int_{-\infty}^{\infty} e^{-t^2}\sqrt{2}\sigma dt$$

$$= N\sqrt{2}\sigma\sqrt{\pi} = 1 \quad\longrightarrow\quad N = \frac{1}{\sqrt{2\pi}\sigma}$$

確率密度関数 ($f(x) = N(\mu,\sigma^2)$) の意味は、確率変数 x の大きさが $x = a, b(a < b)$ の間にある確率 $P(a \le x \le b) = P(a < x < b)$ [3]が、

$$\frac{1}{\sqrt{2\pi}\sigma}\int_a^b e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx = \int_a^b N(\mu,\sigma^2)dx \quad (15)$$

で表され、$a = -\infty$, $b = \infty$ のとき、

$$\int_{-\infty}^{\infty} N(\mu,\sigma^2)dx = 1 \quad (16)$$

である。

$\dfrac{x-\mu}{\sigma} = t$ と変数変換すると、(14) 式の確率密度関数は、

$$f(t) = \frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}} \quad (17)$$

ここに、$f(t) = N(0,1)$ と記す。$N(0,1)$ を標準(あるいは、基準) 正規分布 (Standard Normal Distribution) という。通常，$f(t)$ の定積分

$$\int_{-\infty}^{x} f(t)dt \quad あるいは、\quad \int_{x}^{\infty} f(t)dt$$

の値が表（正規分布表）として与えられている。特に次の値は記憶しておくと便利である：

$$\int_{\mu+\sigma}^{\mu-\sigma} N(\mu,\sigma^2)dx \;=\; 0.685 \quad (18)$$

$$\int_{\mu+2\sigma}^{\mu-2\sigma} N(\mu,\sigma^2)dx \;=\; 0.954 \quad (19)$$

$$\int_{\mu+3\sigma}^{\mu-3\sigma} N(\mu,\sigma^2)dx \;=\; 0.997 \quad (20)$$

これらの積分の意味は、確率変数 x について、

$\mu-\sigma < x < \mu+\sigma$ の確率が 0.683

$\mu-2\sigma < x < \mu+2\sigma$ の確率が 0.954

$\mu-3\sigma < x < \mu+3\sigma$ の確率が 0.997

ということである。

$t^2 = y$ と変数変換すれば、(14) 式の確率密度関数は、

$$f(y) = \frac{1}{\sqrt{2\pi}}y^{-\frac{1}{2}}e^{-\frac{y}{2}} \quad (21)$$

となる．(これを，自由度1の χ^2-分布（カイ自乗分布）という。)

正規分布の特性値、平均及び分散は、

$$\int_{-\infty}^{\infty} xN(\mu,\sigma^2)dx = \mu \quad (22)$$

$$\int_{-\infty}^{\infty} (x-\mu)^2N(\mu,\sigma^2)dx = \sigma^2 \quad (23)$$

である（積分計算の演習問題)。

正規分布の著しい特徴は、変数 $(x_1, x_2, \cdots, x_n)$ が増えても、$\displaystyle\sum_{i=1}^{n} a_i x_i$ の確率分布は、

$$N\left(\sum_{i=1}^{n} a_i x_i, \sum_{i=1}^{n} a_i^2 x_i^2\right)$$

の分布になる。この性質を正規分布の「再生性」という（第5週参照)。

x がある特定の値 x_0 より小さいか、同じ大きさである確率 $F(x_0;\mu,\sigma^2)$ は、

$$F(x_0;\mu,\sigma^2) = \int_{-\infty}^{x_0} N(\mu,\sigma^2)dx \quad (24)$$

で与えられ、x_0 が無限大 ($x_0 \to \infty$) になるとき、

$$F(+\infty;\mu,\sigma^2) = \int_{-\infty}^{\infty} N(\mu,\sigma^2)dx = 1 \quad (25)$$

である。関数 $F(x;\mu,\sigma^2)$ のグラフを図に示す ($y = 0$ と $y = 1$ が漸近線となる)。$F(x)$ は単調増加関数である。この関数を**累積分布関数**、単に分布関数という。分布関数の微分が確率密度関数である：

$$\frac{dF(x)}{dx} = f(x) \quad (26)$$

ここに、$F(x)$ の一般的性質をまとめておこう：

(1) 単調非減少 $x_i \le x_j \to F(x_i) \le F(x_j)$

(2) 右連続 $F(a+0) = \lim_{x\downarrow a} F(x) = F(a)$

(3) $F(-\infty) = \lim_{x\to-\infty} F(x) = 0$

$F(+\infty) = \lim_{x\to+\infty} F(x) = 1$

[3] 1点の積分値は0であるから、上限と下限に等号を含めても含めなくてもどちらでも良い。

確率変数 x の分布が正規分布であるかどうかの判定には、正規確率紙を用いるのが便利である。正規分布に従う集団において、任意にとった値 x より大きくない数値をもつものの百分率を y とすると、

$$y = 100 \times \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x} e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx \stackrel{\frac{x-\mu}{\sigma}=t}{=} 100\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt$$

となる。正規確率紙の縦軸は t なる距離に y を目盛り、横軸には x を等間隔にとる。こうして、各点 x_i に対する y_i をプロットする。正規分布であれば、プロットした点は一直線上にのらなければならない。

第5週 モーメントと母関数・特性関数

モーメント

ある確率分布が与えられたとき、その分布の特性量を統一的に取り扱う方法を述べよう。今、確率密度関数 $f(x)$ に関して、関数 $g(x)$ の平均（あるいは期待値）の定義式を、

$$E(g(x)) =< g(x) >= \int_{-\infty}^{\infty} g(x)f(x)dx \quad (27)$$

と書こう。$g(x)$ の平均（期待値）をとる、あるいは、$g(x)$ に（平均をとる）演算 E を作用させる、という。$g(x)$ が、x, x^2 のときは、

$$E(x) = < x > = \int_{-\infty}^{\infty} xf(x)dx = \mu$$
$$E(x^2) = < x^2 > = \int_{-\infty}^{\infty} x^2 f(x)dx$$

で、x, x^2 の平均として既に用いている。ここに、

$$E(x^n) =< x^n >= \int_{-\infty}^{\infty} x^n f(x)dx \quad (28)$$

を n 次のモーメントという。更に、一般化して、

$$E((x-\mu)^n) =< (x-\mu)^n > = \int_{-\infty}^{\infty} (x-\mu)^n f(x)dx \quad (29)$$

を平均値 μ の回りの n 次のモーメントという。$\mu = 0$ のときは、原点の回りの n 次のモーメントという。特に、$n = 2$ のとき、

$$E((x-\mu)^2) =< (x-\mu)^2 >= \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx =< x^2 > - < x >^2 = \sigma^2 \quad (30)$$

は分散で、$\sqrt{\sigma^2} = \sigma$ は標準偏差であった。

確率変数が二つ、例えば、x と y の二次元確率密度関数を、

$$f(x,y) \geq 0, \qquad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = 1 \quad (31)$$

と書く。$f(x,y)$ を同時確率密度関数ともいう。この関数 $f(x,y)$ において、変数の一つ、例えば、変数 y が x に無関係であれば、その y について積分した

$$f_1(x) = \int_{-\infty}^{\infty} f(x,y)dy \quad (32)$$

を x の周辺分布関数という。同様に、

$$f_2(y) = \int_{-\infty}^{\infty} f(x,y)dx \quad (33)$$

を y の周辺分布関数という。

y の値が決まっているときの x の条件付き確率を、

$$f(x|y) = \frac{f(x,y)}{f_2(y)} \quad (34)$$

で表す。同様に、x の値が決まっているときの y の条件付き確率を、

$$f(y|x) = \frac{f(x,y)}{f_1(x)} \quad (35)$$

で表す。これらの式から、$f(x,y)$ が

$$f(x,y) = f_1(x)f_2(y) \quad (36)$$

と書けるとき、確率変数 x, y は独立である、という。

二変数の確率密度関数 $f(x,y)$ に関して、関数 $g(x,y)$ の平均（あるいは期待値）の定義式を、上記一変数の場合のように、

$$E(g(x,y)) =< g(x,y) > = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)dxdy \quad (37)$$

と書こう。

$g(x,y) = x + y$ （二変数の和）のとき、和の平均は定義式に従って、

$$E(x+y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y)f(x,y)dxdy$$

と書ける。更に、x, y が独立であれば、

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f(x)f(y)dxdy$$

$$= \int_{-\infty}^{\infty}xf(x)dx + \int_{-\infty}^{\infty}yf(y)dy = E(x)+E(y)$$

となる（平均値の線形性）。また、和 $x+y=z$ とおけば、z の確率分布 $(h(z))$ は、

$$h(z) = \int_{-\infty}^{\infty}f(x,z-x)dx = \int_{-\infty}^{\infty}f(z-y,y)dy$$

$$= \int_{-\infty}^{\infty}f(x)f(z-x)dx = \int_{-\infty}^{\infty}f(z-y)f(y)dy$$

$$= (f*f)(z)$$

と書ける（$*$ はたたみ込み積である）。和 z の平均は、

$$\begin{aligned}
E(z) &= \int_{-\infty}^{\infty}zf(z)dz \\
&= \int_{-\infty}^{\infty}z\left\{\int_{-\infty}^{\infty}f(x)f(z-x)dx\right\}dz \\
&= \int_{-\infty}^{\infty}(x+y)\left\{\int_{-\infty}^{\infty}f(x)f(y)dx\right\}dy \\
&= \int_{-\infty}^{\infty}xf(x)dx + \int_{-\infty}^{\infty}yf(y)dy \\
&= E(x)+E(y)
\end{aligned}$$

となる。

前週の課題であった正規分布の再生性が、たたみ込みを用いて容易に示せる。それを二変数の場合で示そう。x,y が $N(\mu,\sigma^2)$ 正規分布に従うときの和 $z=x+y$ の確率分布は、

$$(f*f)(z) = \int_{-\infty}^{\infty}f(z-y)g(y)dy$$

$$= \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_x}e^{-\frac{(z-\mu_x-y)^2}{2\sigma_x^2}}\cdot\frac{1}{\sqrt{2\pi}\sigma_y}e^{-\frac{(y^2-\mu_y)^2}{2\sigma_y^2}}dy$$

$$= \frac{1}{2\pi\sigma_x\sigma_y}\int_{-\infty}^{\infty}e^{-\frac{\sigma_x^2+\sigma_y^2}{2\sigma_x^2\sigma_y^2}\left(y-\frac{\sigma_y^2 z-\left(\sigma_y^2\mu_x+\sigma_x^2\mu_y\right)}{\sigma_x^2+\sigma_y^2}z\right)^2}$$

$$\times e^{-\frac{(z-(\mu_x+\mu_x))^2}{2\left(\sigma_x^2+\sigma_y^2\right)}}dy$$

$$= \frac{1}{\sqrt{2\pi(\sigma_x^2+\sigma_y^2)}}e^{-\frac{(z-(\mu_x+\mu_x))^2}{2(\sigma_x^2+\sigma_y^2)}}$$

$$= N(\mu_x+\mu_x, \sigma_x^2+\sigma_y^2)$$

となる。n 変数の場合についても同様である。

確率変数 $x_1, x_2, \cdots, x_n$ が互いに独立で同一の $N(\mu,\sigma^2)$ 確率分布に従うとき、新しい確率変数 $\overline{x} = \frac{1}{n}\sum_{i=1}^{n}x_i$ は、$N\left(\mu, \frac{\sigma^2}{n}\right)$ 確率分布に従う。つまり、独立な確率分布が多くなるほど、平均値 $(\overline{x})$ のバラツキは小さくなる。更に、新たな確率変数 $\frac{\overline{x}-\mu}{\sigma} = t$ は、$N(0,1)$ 確率分布に従う。これらの事柄が統計的推測の基礎になる。

$g(x,y) = x^i y^j$ のとき（二変数の積）、

$$E(x^i y^j) = < x^i y^j >$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x^i y^j f(x,y)dx \qquad (38)$$

を $(i+j)$ 次のモーメントという。更に、一般化して、

$$E((x-\mu_x)^i(y-\mu_y)^j) = < (x-\mu_x)^i(y-\mu_y)^j >$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_x)^i(y-\mu_y)^j f(x,y)dxdy$$

を平均 μ_x, μ_y の回りの $(i+j)$ 次の結合モーメントという。従って、$\mu_i = 0$ $(i = x,y)$ のときは、原点の回りの $(i+j)$ 次の結合モーメントという。特に、$i = j = 1$ のときは **共分散** という。ここに、平均 $\mu_x(\mu_y)$ の定義式は、

$$\begin{aligned}
\mu_x &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf(x,y)dxdy \\
&= \int_{-\infty}^{\infty}xf_1(x)dx \qquad (39)
\end{aligned}$$

$$\begin{aligned}
\mu_y &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf(x,y)dxdy \\
&= \int_{-\infty}^{\infty}yf_1(y)dy \qquad (40)
\end{aligned}$$

となる。

モーメント母関数・特性関数

形式的に、指数関数 e^{sx}（s は x に無関係なパラメーター）の期待値を考える：

$$E(e^{sx}) \overset{(27)}{=} \int_{-\infty}^{\infty}e^{sx}f(x)dx = F(s) \qquad (41)$$

$F(s)$ をモーメント母関数という。離散変数の場合

は、$\int \longrightarrow \sum$ とすればよい[4]。$f(x)$ と $F(s)$ の対応が一意であることは仮定されている。

モーメント母関数の名前の由来は、モーメントが系統的に導けることによる。以下、それを見てみよう。まず、指数関数を形式的に展開して、

$$\begin{aligned} F(s) &= \int_{-\infty}^{\infty}\sum_{n=0}^{\infty}\frac{(sx)^n}{n!}f(x)dx \\ &= \sum_{n=0}^{\infty}\frac{s^n}{n!}\int_{-\infty}^{\infty}x^n f(x)dx \\ &= \sum_{n=0}^{\infty}\frac{s^n}{n!} < x^n > \end{aligned}$$

これより、モーメントは、

$$< x^n > = \left.\frac{d^n F(s)}{ds^n}\right|_{s=0} = \frac{d^n F(0)}{ds^n} \qquad (42)$$

で得られる。

[例1] 二項分布のモーメント母関数

$$F(s) = \sum_{x=0}^{\infty} {}_nC_x p^x q^{n-x} e^{sx} = \sum_{x=0}^{\infty} {}_nC_x (pe^s)^x q^{n-x}$$

$$= (pe^s + q)^n$$

[例2] ポアソン分布のモーメント母関数

$$F(s) = \sum_{x=0}^{\infty}\frac{\lambda^x}{x!}e^{-\lambda}e^{sx} = \sum_{x=0}^{\infty}\frac{(\lambda e^s)^x}{x!}e^{-\lambda}$$

$$= e^{\lambda e^s}e^{-\lambda} = e^{\lambda(e^s-1)}$$

あるいは、[例1]の結果を用いて、

$$(pe^s+q)^n = (pe^s+1-p)^n$$

$$= \{1-p(1-e^s)^n\} = \{1-\frac{\lambda}{n}(1-e^s)^n\}$$

$$\xrightarrow{n\to\infty} e^{\lambda(e^s-1)}.$$

[例3] 正規分布のモーメント母関数

$$F(s) = \int_{-\infty}^{\infty} e^{sx}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$$

$$= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}e^{-\frac{(x-\mu-s\sigma^2)^2}{2\sigma^2}+\mu s+\frac{1}{2}\sigma^2 s^2}dx$$

$$\overset{x-\mu-s\sigma^2=X}{=} e^{\mu s+\frac{1}{2}\sigma^2 s^2}\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}e^{-\frac{1}{2\sigma^2}X^2}dX$$

$$= e^{\mu s+\frac{1}{2}\sigma^2 s^2}\frac{1}{\sqrt{2\pi}\sigma}\sqrt{2\pi\sigma^2} = e^{\mu s+\frac{1}{2}\sigma^2 s^2}$$

[課題] モーメント母関数を用いて、二項分布、ポアソン分布及び正規分布の平均・分散を求めよ。

解. 結果を表にまとめておく：

確率分布	モーメント母関数	平均	分散
${}_nC_x p^x q^{n-x}$	$(pe^s+q)^n$	np	npq
$\frac{\lambda^x}{x!}e^{-\lambda}$	$e^{\lambda(e^s-1)}$	λ	λ
$\frac{1}{\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$e^{\mu s+\frac{1}{2}\sigma^2 s^2}$	μ	σ^2

次に、指数関数 e^{itx}（t は x に無関係なパラメーター）の期待値（確率密度関数 $f(x)$ のフーリエ変換）を考える:

$$E(e^{itx}) \overset{(27)}{=} \int_{-\infty}^{\infty} e^{itx}f(x)dx = F(t) \qquad (43)$$

$F(t)$ を確率密度関数 $f(x)$ の特性関数という。この特性関数 $F(t)$ とモーメントの関係を調べよう。まず、指数関数を形式的に展開して、

$$\begin{aligned} F(t) &= \int_{-\infty}^{\infty}\sum_{n=0}^{\infty}\frac{(itx)^n}{n!}f(x)dx \\ &= \sum_{n=0}^{\infty}\frac{(it)^n}{n!}\int_{-\infty}^{\infty}x^n f(x)dx \\ &= \sum_{n=0}^{\infty}\frac{(it)^n}{n!} < x^n > \end{aligned}$$

これより、モーメントは特性関数がわかれば、

$$< x^n > = (-i)^n\left.\frac{d^n F(t)}{dt^n}\right|_{t=0} = (-i)^n\frac{d^n F(0)}{dt^n} \qquad (44)$$

で得られる。

[例1] 二項分布の特性関数

$$F(t) = \sum_{x=0}^{\infty} {}_nC_x p^x q^{n-x} e^{itx} = (pe^{it}+q)^n$$

[例2] ポアソン分布の特性関数

$$F(t) = \sum_{x=0}^{\infty}\frac{\lambda^x}{x!}e^{-\lambda}e^{itx} = \sum_{x=0}^{\infty}\frac{(\lambda e^{it})^x}{x!}e^{-\lambda}$$

$$= e^{\lambda e^{it}}e^{-\lambda} = e^{\lambda(e^{it}-1)}$$

[4] 次の定義式でも同様：

$$E(s^n) \overset{(27)}{=} \sum_{x=0}^{n} s^n P(x) = F(s)$$

$F(s)$ は単に母関数という（$P(x)$ は離散的確率関数）。モーメントは次式による：

$$< x^n > = \left.\frac{d^n F(t)}{ds^n}\right|_{s=1} = \frac{d^n F(1)}{ds^n}$$

[例３]　正規分布の特性関数

$$F(t) = \int_{-\infty}^{\infty} e^{itx}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$$
$$= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-\frac{(x-\mu-it\sigma^2)^2}{2\sigma^2}+i\mu t-\frac{1}{2}\sigma^2t^2}dx$$
$$\overset{x-\mu-it\sigma^2=X}{=} e^{i\mu t-\frac{1}{2}\sigma^2t^2}\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}X^2}dX$$
$$= e^{i\mu t-\frac{1}{2}\sigma^2t^2}\frac{1}{\sqrt{2\pi}\sigma}\sqrt{2\pi\sigma^2} = e^{i\mu t-\frac{1}{2}\sigma^2t^2}$$

[問1]　次の確率分布の特性関数を求めよ。

$$(1) f(x) = \begin{cases} 1 & (|x| < a) \\ 0 & (|x| \geq a) \end{cases} \quad (2) f(x) = \begin{cases} \lambda e^{-\lambda x} & (x \geq 0) \\ 0 & (x < 0) \end{cases}$$

解.

$$(1)\ F(t) = \int_{-\infty}^{\infty} e^{itx}1dx = \frac{1}{it}\left[e^{itx}\right]_a^a$$
$$= \frac{1}{it}\left(e^{ita} - e^{-ita}\right) = \frac{2\sin at}{t}$$
$$(2)\ F(t) = \int_{-\infty}^{\infty} e^{itx}\lambda e^{-\lambda x}dx$$
$$= \frac{\lambda}{it-\lambda}\left[e^{(it-\lambda)x}\right]_0^{\infty} = \frac{\lambda}{\lambda - it}$$

[課題]　特性関数を用いて、二項分布、ポアソン分布及び正規分布の平均・分散を求めよ。

第６週　統計的推測

母集団と標本

問題にしようとする対象の観測値（実験値）の全集合を母集団という。母集団の要素の数を母集団の大きさ、その要素の数が有限、無限に対して、有限母集団、無限母集団という。母集団の特性値（母平均 μ, 母分散 σ^2 等）を**母数**という。それらの値は、既知であろうと未知であろうと定数である。

母集団から一つの要素 X を取り出すことを標本抽出(sampling)といい、その数値を標本値 x（標本値は、確率変数のとる値（実現値）のこと）、取り出されたものを標本という。以上を換言すれば、大きさ１の標本である。

次に、母集団から n 個の要素 $X_1, X_2, \cdots, X_n$ をとり、それらの標本値 $x_1, x_2, \cdots, x_n$ とすれば、大きさ n の標本を得る。X_i は独立で、X と同じ分布とする。言い換えれば、母集団の分布と標本のそれは、一定の関係で結ばれている。これによって、標本の分析・考察から母集団に関するの知見を得ることができる。

＜母集団と標本の区別＞

[例1] １枚のコインを投げて表のでる確率を調べる。表がでれば1、裏がでれば0とする。コイン投げを繰り返し行なうと、次のような 1,0 の無限個の数値の集まりが得られる：

1,0,0,1,0,1,1,0, ⋯　　(無限個)

この 1,0 の無限個の数値の数値の集まり（集合）が無限母集団である。この中から、実際に 100 個取り出した 1,0 の集まり

0,1,0,0,1,1,0,1,⋯　　(100個)

が大きさ 100 の標本である。

[例2] ある学年 200 人の平均身長を調べる。無作為抽出で 10人選んで身長を測定した。測定値（標本値）は次の通り：

169, 175, 178, 162, 165, 170, 181, 173, 171, 170

この学年 200 人の身長値の集まりが有限母集団で，選んだ 10 人の身長値の集まりが大きさ 10 の標本となる。

n 個の要素の選び方は任意であるから（無作為抽出）、この操作を繰り返せば、一回ごとに異なる大きさ n の標本を得る。従って、一回ごとの標本の特性値も変わる。これを、標本の特性値（標本平均、標本分散等）は**標本変動**(偶然変動）という。このように、標本と母集団の特性値の相違を理解しておくことは，以下の議論において本質的である。

統計量とその分布

ある分布をもつ母集団の無作為抽出の大きさ n の標本において、

$$\overline{X} = \frac{1}{n}(X_1 + X_2 + \cdots + X_n) = \frac{1}{n}\sum_{i=1}^{n}X_i \quad (45)$$

を標本平均といい、

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \frac{1}{n}\sum_{i=1}^{n}X_i^2 - \overline{X}^2 \quad (46)$$

を標本分散という。$\sqrt{S^2} = S$ を標本標準偏差という。上記 $\overline{X}$ や S^2 のように、$X_1, X_2, \cdots, X_n$ の関数 $T(X_1, X_2,$
$\cdots, X_n)$ を一般に統計量と呼ぶ。上記「例２」の

統計量の値は次の通り：

$$\overline{X}=171.4,\quad S^2=29.04,\quad S=5.39$$

標本平均の期待値$E[\overline{X}]$は、

$$E[\overline{X}]=\mu \tag{47}$$

即ち、標本平均 $\overline{X}$ の平均は母集団の平均値 μ に等しい。これを大きさ n の標本の標本平均は母平均の**不偏推定量**である、という。一般に、母数 θ に対して、統計量 T が存在して $E[T]=\theta$ であるとき、T を θ の不偏推定量と呼ぶ。他方、S^2 の期待値は、

$$E[S^2]=\frac{n-1}{n}\sigma^2 \tag{48}$$

で、σ^2 にならない。しかし，$\dfrac{n-1}{n}S^2$ の期待値は母平均 σ^2 になる。こうして，

$$U^2=\frac{n-1}{n}S^2 \tag{49}$$

は、母分散 σ^2 の不偏推定量で**不偏分散**と呼ぶ。上記「例２」の標本値からそれらの値は次の通り：

$$U^2=32.27,\quad U=5.69$$

(??) 式において、係数 $\dfrac{n-1}{n}$ が意味をもつのは n が小さいときである。統計量は、抽出の度に異なるので確率変数である（つまり、統計量は標本変動する)。以下、特徴的な統計量（確率変数）とその分布を示す：

(I) 標本平均 $\overline{X}$ は、標本の抽出のされ方で値は異なる。X_i がいずれも正規分布 $\mathrm{N}(\mu,\sigma^2)$ に従うときは、n 個の標本は $\mathrm{N}(\mu,\dfrac{\sigma^2}{n})$ なる正規分布に従う。

$\dfrac{X_i-\overline{X}}{\sigma}$ を新しく X_iにとれば、X_i は正規分布 $\mathrm{N}(0,1^2)$ に従う。

(II) $\dfrac{X_i}{\sigma}$ を新しく X_i にとれば、(1) と (2) から

$$\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2=\frac{S^2}{\sigma^2} \tag{50}$$

で、$\dfrac{nS^2}{\sigma^2}$ は自由度 $n-1$ の χ^2-分布 (カイ二乗分布) に従う。

χ^2-分布 (カイ二乗分布) の確率密度関数 $f(x)$ は、

$$f(x)=\begin{cases}\dfrac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}x^{\frac{n}{2}-1}e^{-\frac{x}{2}}, & (x>0)\\ 0 & (x\leq 0)\end{cases} \tag{51}$$

である。χ^2-分布は母分散の推定等の解析に用いる。

(III) 更に統計量

$$t=\frac{\overline{X}-\mu}{\dfrac{S}{\sqrt{n-1}}}=\frac{\sqrt{n-1}(\overline{X}-\mu)}{S} \tag{52}$$

は、自由度 $n-1$ なる $t-$分布になる。母平均の推定等の解析に用いる。

$t-$ 分布の確率密度関数は、

$$f(t)=\frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\sqrt{\pi n}\Gamma\left(\dfrac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}$$

$$(-\infty<t<\infty) \tag{53}$$

である。$t-$分布は $n\to\infty$ のとき、$\mathrm{N}(0,1^2)$ となる（$n\geq 30$ に対して、$\mathrm{N}(0,1^2)$ とみなしてよい)。

第７・８週　統計的推定と検定

母集団の母数を母集団からの標本を分析して推定することを統計的推定という[5]。推定法には、母数を直接推定する**点推定**と、これをある幅の区間で推定する**区間推定**とがある。点推定値を求める方法についてまず述べる：

最尤法 (method of maximum likelihood)

ここでは、正規分布に従う母集団（正規母集団）の母平均 μ, 母分散 σ^2 の点推定を行う：

母集団から抽出した標本を $(x_1,x_2,\cdots,x_n)$ で表す。x_i の確率密度関数は、$f(x_i)=\dfrac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$ であるから、x_i の値が区間 $(x_i,\ x_i+\Delta x_i)$ に落ちる確率は $f(x_i)\Delta x_i$ である。$x_i\ (i=1,2,\cdots,n)$ は、互いに独立に抽出されているので、それぞれの x_i が区間$(x_i,\ x_i+\Delta x_i)$ に落ちる確率（同時確率分布）は、

$$f(x_1)\Delta x_1 f(x_2)\Delta x_2\cdots f(x_n)\Delta x_n$$

$$=\left(\frac{1}{\sqrt{2\pi}\,\sigma}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}\Delta x_1\cdots\Delta x_n$$

[5]母数 θ の推定量を $T(X_1,X_2,\cdots,X_n)$ とするとき、Tが推定量として望ましい性質は次のようなものである；

- 不偏性　$E[T]=\theta$
- 有効性　母数 θ の不偏推定量 T_1,T_2 において、分散

  $$V[T_1]<V[T_2]$$

  が成り立つとき、T_1 は T_2 より有効であると言う。つまり、(θ を中心として) T_1 の分散が小さいこと。
- 一致性　$\displaystyle\lim_{n\to\infty}P(|T-\theta|<\varepsilon)=1\quad(\varepsilon>0)$

となる。ここで、上式の係数を

$$A(\mu) = \left(\frac{1}{\sqrt{2\pi}\,\sigma}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

と書く。この関数を尤度関数、その値を尤度（likelihood; 尤もらしい度合）という。以下、この $A(\mu)$ を最大にする μ（最尤解；$\hat{\mu}$）を見つけることである。これは、直接微分法で、$\dfrac{\partial A(\mu)}{\partial \mu} = 0$ から求めることができるが、対数微分法を用いた方がみやすい：

$$\log A(\mu) = -\frac{n}{2}\log 2\pi - n\log\sigma$$

$$-\frac{1}{2\,\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 \quad \frac{1}{A(\mu)}\frac{\partial \log A(\mu)}{\partial \mu}$$

$$= -\frac{1}{\sigma}\sum_{i=1}^{n}(x_i-\mu) = 0$$

これより、母平均の最尤解として、

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n}x_i^2 \tag{54}$$

を得る。同様にして、母分散の最尤解は、

$$\frac{1}{A(\sigma)}\frac{\partial \log A(\sigma)}{\partial \sigma} = -\frac{n}{\sigma} - \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i-\mu)^2 = 0$$

$$\hat{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 \tag{55}$$

となる。これらの結果（(??) 式と (??) 式）は、平均の推定は最尤推定量が不偏推定量であるが、分散のそれは最尤推定量が不偏推定量にならない、ことをを示している。

母平均と母分散の区間推定

最尤推定法によって推定される正規母集団の母平均 (μ) と母分散 (σ^2) は、

$$\mu = \frac{1}{n}\sum_{i=1}^{n}X_i$$

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2 = S^2$$

と与えられた。これらの推定された μ, σ^2 がどの程度信頼おけるのか、ということが以下の課題となる。

（1）母平均の区間推定

確率変数 X の標本分布の分散は、母集団の分散 σ^2 が既知ならば $\dfrac{\sigma^2}{n}$ になるので、分布の幅は $\sqrt{n}$ に逆比例して狭くなる。従って、推定値 μ の精度も高くなる。しかし、一般には、σ^2 は既知でない場合が多い。その場合には、前節の (II) を用いて、望ましい μ は、$1-\alpha$ の信頼度で区間は、

$$\overline{X} - \lambda\frac{S}{\sqrt{n-1}} \le \mu \le \overline{X} + \lambda\frac{S}{\sqrt{n-1}} \tag{56}$$

の間にあることがいえる。ここに、λ は次のように決める；

$$t = \frac{\overline{X}-\mu}{\dfrac{S}{\sqrt{n-1}}} = \frac{\sqrt{n-1}(\overline{X}-\mu)}{S} \quad (n \ge 2)$$

とおいて、$f(t)$ をλから無限大まで積分したものが、

$$\int_{\lambda'}^{\infty} f(t)dt = \frac{\alpha}{2}$$

であるようなもので、$t-$分布表に与えられている。

（2）母分散の区間推定

母分散 σ^2 の信頼区間は、$\dfrac{nS^2}{\sigma^2}$ が自由度 $n-1$ の χ^2-分布 に従うことから、

$$\frac{nS^2}{\sigma^2} \le a_\nu \quad と \quad \frac{nS^2}{\sigma^2} \le b_\nu$$

のとる確率 α が、

$$\int_0^{a_\nu} f(x)dx + \int_{b_\nu}^{\infty} f(x)dx = \alpha$$

として、a_ν, b_ν を χ^2- 分布表から求める。こうして、母分散は、信頼度 $1-\alpha$

$$\frac{nS^2}{b_\nu} \le \sigma^2 \le \frac{nS^2}{a_\nu} \tag{57}$$

で与えられる。

[例題] ある高専の４年生の応用数学の平均点を知るために、４年生の１クラス　４１名　の平均点 ($\overline{X}$)、標準偏差 (S) を調べた所、それぞれ６６、１０点であった。これより、４年生全体の平均点を信頼度９５％で区間推定せよ[6]。

１．正規母集団 $N(\mu, \sigma = S) \overset{(抽出)}{\Longrightarrow} N(\mu, \frac{\sigma}{\sqrt{n}})$
標本の大きさ $n = 41$

（母集団の分散が既知でないとき、$\sigma = S$ を仮定する。）

２．統計量 $\dfrac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} = T$ は、N(0,1) に従う。信頼度９５％の意味は、

[6]区間推定の「手順」を示す：

1. 母集団より標本を抽出する
2. 統計量、信頼度を決める
3. 統計量の信頼区間を求める

$$P\left(\left|\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right| \leq 1.96\right)=0.95$$

区間 $-1.96 \leq \dfrac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \leq 1.96$ の確率が 0.95 である。

3．母数（平均点）μ の９５％の信頼区間は、

$$\left[\overline{X}-1.96\times\frac{\sigma}{\sqrt{n}}, \overline{X}+1.96\times\frac{\sigma}{\sqrt{n}}\right]$$

$$=[62.94, 69.06]$$

[問題１]　信頼度９９％で区間推定せよ。

[解] 母数（平均点）μ の９９％の信頼区間は,

$$\left[\overline{X}-2.58\times\frac{\sigma}{\sqrt{n}}, \overline{X}+2.58\times\frac{\sigma}{\sqrt{n}}\right]$$

$$=[62.94, 69.06]$$

[問題２]　ｔ一分布を用いて、信頼度９５％、９９％で区間推定せよ。

[解]　母数（平均点）μ の信頼度９５％、９９％の区間推定は、それぞれ

$$\left[\overline{X}-2.021\times\frac{S}{\sqrt{n-1}}, \overline{X}+2.021\times\frac{S}{\sqrt{n-1}}\right]$$

$$=[62.80, 69.20]$$

$$\left[\overline{X}-2.704\times\frac{S}{\sqrt{n-1}}, \overline{X}+2.704\times\frac{S}{\sqrt{n-1}}\right]$$

$$=[61.72, 70.28]$$

仮説とその検定

推計学においては、母集団の母数（母平均や母分散）の推定を行なうと共に、ある仮説 (hypothesis) を設定して検定を行なう。即ち、実際に起こった事象 A が、その仮説のもとで起こる確率 P(A) を求め、その確率が “非常に小さい” 場合には、その仮説は棄て去り別の仮説を考える、というように検定をする。この仮説は、棄却されることを念頭にした仮説なので帰無仮説 (Null Hypothesis) と呼ばれる。棄却するかどうかの検定の基準を有意水準 (level of significance) あるいは危険率という。

仮説の検定の手続きを次に示す：まず、有意水準 (α) を設定する．確率変数 z として、

$$\int_{Z_\alpha}^{\infty} f(z)dz$$

を満たすような (Z_α) を求める。つまり、棄却域を決める。この Z_α と、標本からの検定統計量の実現値 z_0 とを比較する。z_0 が Z_α より小さければ、仮説は “棄却されず”、z_0 が Z_α より大きければ “棄却する” ということになる。この場合、仮説を棄却する危険率は α（100 α %）である[7]。

[例題]　ある高専の４年生の応用数学の平均点は例年７０点であった。今年度の４年生の１クラス　４１名　の平均点 $(\overline{X})$、標準偏差 (S) を調べた所、それぞれ６６、１０点であった。これより、このクラスは、例年の４年生に較べて能力が低いといえるか、危険率５％で吟味せよ。

1．$H_0 : \mu = 70$　（このクラスは例年と同じ能力である）

2．正規母集団 $N(\mu, \sigma=S) \overset{(\text{抽出})}{\Longrightarrow} N(\mu, \frac{\sigma}{\sqrt{n}})$

標本の大きさ $n=41$

統計量 $\dfrac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}=T$ は、$N(0,1)$ に従う。

3．$\alpha=0.05$, 両側検定として、

$$P\left(\left|\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right| \geq 1.96\right)=0.05$$

つまり、棄却域は、$Z \geq 1.96,\quad Z \leq -1.96$ である。

4．Tの実現値　$T=-2.56<-1.96$。仮説は棄却される。今後、何らかの対策が必要となる。

[課題１]　危険率１％で検定せよ。

[課題２]　ｔ一分布を用いて、危険率５％、１％で検定せよ。

＜ 付　録 ＞

二項分布 $Bin(n,p)$ に従う離散的確率変数 x が、

$$t=\frac{x-\mu}{\sigma}\quad (\mu=np,\ \sigma=\sqrt{npq}) \tag{58}$$

のとき、$a, b,\ (a<b)$ に対して、

$$\lim_{n\to\infty} P(a<t<b)=\frac{1}{2\pi}\int_a^b e^{-\frac{t^2}{2}}dt$$

である。つまり、二項分布は、$n\to\infty$ のとき、正規分布で近似できる：

[7]一般に、仮説の検定を行なって、ある有意水準 α でそれを棄却することができる場合は、その仮説の検定は有意水準 α で “有意である“ という。これに反して、有意水準 α では仮説は棄却できない場合には、有意水準 α でその仮説の検定は、“有意でない” という。換言すれば、仮説が正しいかどうかということには、その有意水準では結論がだせないということである。

$$Bin(n,p) = {}_nC_x p^x q^{n-x} \overset{n\to\infty}{\to} \frac{1}{2\pi}e^{-\frac{t^2}{2}}$$

これを示そう： $n \to \infty$ のとき、離散変数 $t = \dfrac{x-\mu}{\sigma}$ は連続変数 t になる。従って、その分布も離散分布から連続分布になる。

$$\lim_{n\to\infty} Bin(n,p)\Delta x = f(t)\Delta t \quad (\Delta x = \sigma\Delta t)$$

ここに、$f(t)$ は確率密度関数で、

$$\begin{aligned} f(t) &= \lim_{n\to\infty} Bin(n,p)\sigma \\ &= \lim_{n\to\infty} {}_nC_x p^x q^{n-x}\sqrt{npq}. \end{aligned}$$

$$f(t) = \lim_{n\to\infty} \frac{n!\ n^{\frac{1}{2}}}{x!\ (n-x)!} p^{x+\frac{1}{2}} q^{n-x+\frac{1}{2}}$$

$$\overset{\text{Stirling fomula}}{\Rightarrow} \frac{\sqrt{2\pi}n^{n+\frac{1}{2}}e^{-n}\ n^{\frac{1}{2}}}{\sqrt{2\pi}x^{x+\frac{1}{2}}e^{-x}\ \sqrt{2\pi}(n-x)^{n-x+\frac{1}{2}}e^{-(n-x)}}$$

$$\times\ p^{x+\frac{1}{2}}q^{n-x+\frac{1}{2}} \ = \ \frac{1}{\sqrt{2\pi}}\left(\frac{np}{x}\right)^{x+\frac{1}{2}}\left(\frac{nq}{n-x}\right)^{n-x+\frac{1}{2}}$$

$$= \frac{1}{\sqrt{2\pi}}\left(\frac{np}{np+t\sqrt{npq}}\right)^{np+t\sqrt{npq}+\frac{1}{2}}$$

$$\left(\frac{nq}{nq-t\sqrt{npq}}\right)^{nq-t\sqrt{n}}$$

$$= \frac{1}{\sqrt{2\pi}}\left(\frac{1}{1+t\sqrt{\dfrac{q}{np}}}\right)^{np+t\sqrt{npq}+\frac{1}{2}}$$

$$\left(\frac{1}{1-t\sqrt{\dfrac{p}{nq}}}\right)^{nq-t\sqrt{npq}+\frac{1}{2}}$$

$$= \frac{1}{\sqrt{2\pi}} \times A$$

この A を $\log A$[8] として以下のように計算する：

$$\begin{aligned} \log A &= -(np+t\sqrt{npq}+\frac{1}{2})\log\left(1+t\sqrt{\frac{q}{np}}\right) \\ &\quad -(nq+t\sqrt{npq}+\frac{1}{2})\log\left(1-t\sqrt{\frac{p}{nq}}\right) \\ &= -(np+t\sqrt{npq}+\frac{1}{2}) \\ &\quad \left(t\sqrt{\frac{q}{np}}-\frac{1}{2}\left(t\sqrt{\frac{q}{np}}\right)^2+\cdots\right) \\ &= -(nq-t\sqrt{npq}+\frac{1}{2}) \\ &\quad \left(-t\sqrt{\frac{p}{nq}}-\frac{1}{2}\left(t\sqrt{\frac{p}{nq}}\right)^2+\cdots\right) \\ &\Rightarrow\ -\frac{1}{2}(p+q)t^2 \ = \ -\frac{1}{2}t^2 \end{aligned}$$

$$\log A = -\frac{1}{2}t^2 \quad \longrightarrow \quad A = e^{-\frac{1}{2}t^2}$$

よって、求める式を得る：

$$f(t) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2}$$

（平成10年11月30日受理）

8

$$\log(1+x) = x - \frac{1}{2}x^2 + \ \cdots \ (x \longrightarrow 0)$$