付録 A

重要な関係式などの導出

A.1　四分位数を求める

10 ページでは 3 つの四分位数（ひとつはメジアン）を決定する手順を示したが，2 個のデータ点を不均等に内分するところの説明は飛ばして天下りに記述してある．ここではデータの数が $n = 4m$ $(m = 1, 2, 3, \ldots)$ の場合についてもう少し詳しく説明する．

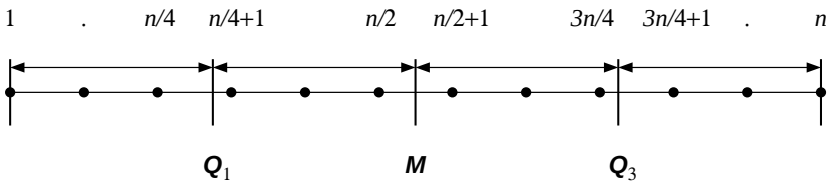

図 A.1　データ数が 4 の倍数の場合に四分位数 $\boldsymbol{Q_1}, \boldsymbol{M}, \boldsymbol{Q_3}$ を求めるための図．データ点 $x_1, x_{n/2}$ 等は添字だけで示してある．

図にはデータ点 $x_1, x_2, \ldots, x_n$ を等間隔で描いてある．左端の座標を 1 としてデータ間の区間の長さを 1 とすると，全体の長さは $n-1$ になる．$n-1$ は 4 では割り切れないので，4 等分した区切りは区間の途中に入る．

$\boldsymbol{Q_1}$ の位置は左端から $\frac{n-1}{4}$ のところにあるので，その座標は $1 + \frac{n-1}{4} = \frac{n}{4} + \frac{3}{4}$ となる．つまり $x_{n/4}$ と $x_{n/4+1}$ のデータを 3 : 1 に内分する点が Q_1 となる．同様にして，$\boldsymbol{Q_3}$ の位置は左端から $\frac{3(n-1)}{4}$ のところにあることから，$x_{3n/4}$ と $x_{3n/4+1}$ のデータを 1 : 3 に内分する点が Q_3 となる．

またメジアン $\boldsymbol{M}$ については，$x_{n/2}$ と $x_{n/2+1}$ を 1 : 1 に内分する点であることが図からすぐに分かる．

A.2　ベイズの定理

事象 A, B, C が互いに排反で，かつ標本空間を尽くしているとする．このとき，A, B, C それぞれの下に，ある事象 E が起きる条件付き確率，

$$P(E|A), P(E|B), P(E|C)$$

が知られているとすると，次の式が成立する．

$$P(A|E) = \frac{P(A)P(E|A)}{P(A)P(E|A) + P(B)P(E|B) + P(C)P(E|C)} \tag{A.1}$$

まず，条件付き確率 $P(A|E)$ は次の式に従う (p.33, 式 (2.18) 参照)．

$$P(A|E) = \frac{P(A \cap E)}{P(E)} \tag{A.2}$$

$P(B|E)$, $P(C|E)$ についても同様の式が成立する．逆に次の形も成立することにも注意しよう．

$$P(E|A) = \frac{P(A \cap E)}{P(A)} \tag{A.3}$$

さらに，事象 A, B, C が排反であり，かつ標本空間を尽くしているから，$P(E)$ は次のようになる．

$$P(E) = P(E \cap A) + P(E \cap B) + P(E \cap C) \tag{A.4}$$

そこで，式 (A.2) に式 (A.3) と式 (A.4) を代入して整理すれば，求める式が得られる．

A.3　二項分布の平均と分散

A.3.1　二項分布の平均

二項分布の平均（期待値）は 51 ページの式 (3.10) で与えられている．

$$\mu = \sum_{x=0}^{n} x \ {}_nC_x p^x q^{n-x} = np, \ \ (q = 1 - p) \tag{A.5}$$

この式を導くには次のようなテクニックを使うのが面白い．まず，二項定理を使って $(p+q)^n$ を展開すると次のようになる．

$$\begin{aligned}(p+q)^n &= \sum_{x=0}^{n} {}_nC_x p^x q^{n-x} \\ &= p^n + np^{n-1}q + \frac{n(n-1)}{2}p^{n-2}q^2 + \cdots \end{aligned} \tag{A.6}$$

ここで式 (A.6) を p で微分してみると，次のようになる．

$$n(p+q)^{n-1} = \sum_{x=0}^{n} x\ {}_nC_x p^{x-1} q^{n-x} \tag{A.7}$$

両辺に p を掛けて，$p+q=1$ を使えば，

$$np = \sum_{x=0}^{n} x\ {}_nC_x p^x q^{n-x} \tag{A.8}$$

となって，求める関係式が得られている．

A.3.2　二項分布の分散

次に分散 σ^2 を求めてみよう．最初に，

分散 ＝ 2 乗の平均 − 平均の 2 乗

であることを思い出しておこう．

式 (A.7) をもう一度 p で微分すると，

$$n(n-1)(p+q)^{n-2} = \sum_{x=0}^{n} x(x-1)_nC_x p^{x-2} q^{n-x} \tag{A.9}$$

が得られる．両辺に p^2 を掛けて，左辺と右辺を展開し，さらに $p+q=1$ と置くと，

$$n^2p^2 - np^2 = \sum_{x=0}^{n} x^2\ {}_nC_x p^x q^{n-x} - \sum_{x=0}^{n} x\ {}_nC_x p^x q^{n-x} \tag{A.10}$$

式 (A.5) より，$np=\mu$．従って左辺第 1 項は平均の 2 乗である．また，右辺第 1 項は 2 乗の平均を，右辺第 2 項は平均を意味している．これらに注意して整理すると，次のように変形して，最後の結果が得られる．

$$\begin{aligned}\sigma^2 &= \sum_{x=0}^{n} x^2\ {}_nC_x p^x q^{n-x} - n^2p^2 \\ &= \sum_{x=0}^{n} x\ {}_nC_x p^x q^{n-x} - np^2 = np - np^2 = npq \end{aligned} \tag{A.11}$$

A.4　ポアソン分布

ポアソン分布は，二項分布において n が非常に大きく，かつ p が非常に小さい極限で成立する確率分布である．つまり，$n \longrightarrow \infty,\ p \longrightarrow 0$ とした時に，

$$ {}_nC_x p^x (1-p)^{n-x} \longrightarrow \frac{\mu^x}{x!} \mathrm{e}^{-\mu} \tag{A.12} $$

となる．なおここで，$\mu = np$ であり，この値は平均値であって，有限の値をもつ．これを次のように段階に分けて証明する．

■x が小さい場合の ${}_nC_x$

$$ \begin{aligned} {}_nC_x &= \frac{n!}{x! \times (n-x)!} \\ &= \frac{n \cdot (n-1) \cdots 2 \cdot 1}{x \cdot (x-1) \cdots 2 \cdot 1 \times (n-x) \cdot (n-x-1) \cdots 2 \cdot 1} \end{aligned} \tag{A.13} $$

この式の分子の方をよく考えよう．n の階乗で，$n > x \geq 0$ だから $n \geq (n-x)$ となることから，次のように書いておく．

$$ n! = n \cdot (n-1) \cdots (n-x+1) \cdot (n-x) \cdot (n-x-1) \cdots 2 \cdot 1 $$

これから，式 (A.13) の分母と分子を通分して次の式を得る．

$$ {}_nC_x = \frac{n \cdot (n-1) \cdots (n-x+1)}{x \cdot (x-1) \cdots 2 \cdot 1} \tag{A.14} $$

ここで分子の $n \cdot (n-1) \cdots (n-x+1)$ は，x 個の因子の掛けあわせであることを押さえておこう[*1]．さらに $n \to \infty$ より $n \gg x$ であるから，$n-1$ や $n-x+1$ などはすべて n と近似的に等しいとみなせて，式 (A.14) は $n \to \infty$ の極限で次のように近似できる．

$$ {}_nC_x \longrightarrow \frac{n^x}{x!} \tag{A.15} $$

■$p^x(1-p)^{n-x}$ の極限値

まず，自然対数の底 (てい) である e の定義は次の式で与えられることを心に留めておこう．この定義の式の意味については適当な数学の参考書を見ていただきたい．

$$ \mathrm{e} = \lim_{q \to \infty} \left(1 + \frac{1}{q} \right)^q \tag{A.16} $$

[*1] $n-0$ から $n-(x-1)$ までの積であるから，$0, 1, \ldots, (x-1)$ と数え挙げると x 個ある．

定義 (A.16) から次の式が導けることを証明なしにあげておく．

$$\lim_{q\to\infty}\left(1-\frac{1}{q}\right)^{q}=\frac{1}{\mathrm{e}} \tag{A.17}$$

それでは二項分布の式 (A.12) 左辺の後半に現れる因子[*2]を変形していこう．

$$p^{x}(1-p)^{n-x}=\left(\frac{p}{1-p}\right)^{x}\times(1-p)^{n} \tag{A.18}$$

式 (A.18) の右辺の前半の因子については，$p\to 0$ で $(1-p)\to 1$ となるから，

$$\left(\frac{p}{1-p}\right)^{x}\longrightarrow p^{x} \tag{A.19}$$

となる[*3]．さらに $(1-p)^{n}$ について考える．この式で $q=1/p$ と置くと次のように変形でき，これから $p\longrightarrow 0$ における極限値として次の式が得られる．

$$\begin{aligned}(1-p)^{n} &= (1-p)^{\frac{1}{p}\times np}\\ &= \left(1-\frac{1}{q}\right)^{q\times\mu}\\ &\longrightarrow \mathrm{e}^{-\mu}\end{aligned} \tag{A.20}$$

3 番目の式を得るのには，式 (A.17) が使われている．

さて，ここで式 (A.15) (A.19) (A.20) をまとめると，

$$\begin{aligned}{}_{n}C_{x}p^{x}(1-p)^{n-x} &\longrightarrow \frac{n^{x}}{x!}\times p^{x}\times\mathrm{e}^{-\mu}\\ &= \frac{(np)^{x}}{x!}\mathrm{e}^{-\mu}\\ &= \frac{\mu^{x}}{x!}\mathrm{e}^{-\mu}\end{aligned} \tag{A.21}$$

というポアソン分布の式を得る．

[*2] 掛け算の形の式があるときに，掛け合わされるものを因子 (英語では factor) という．たとえば $ax(x-1)$ とあるときには，a, x, $x-1$ が因子である．因数ともいう．

[*3] $p\longrightarrow 0$ なら $1-p$ が 1 に近づくときに分子の p も同時にゼロにするべきではないかと思う人もいるだろう．しかし，そうすると全体がゼロになってしまって式としては意味を失う．$p\longrightarrow 0$ ということは，あくまで p は正の実数であって，何がしかの意味をもった値であることには違いはないのである．それでも気になる人は 0.99999 を 1 に近似しても大した違いはないが，0.00001 をゼロにしたらまずいということを考えればわかるはずだ．

A.5　標本平均の平均と分散の関係

$$\begin{aligned} E[\overline{X}] &= E[\frac{1}{n}(X_1 + X_2 + ... + X_n)] \\ &= \frac{1}{n}\left(E[X_1] + E[X_2] + ... + E[X_n]\right) \\ &= \frac{1}{n}(\mu + \mu + ... + \mu) = \mu \end{aligned}$$

2 つ目の式を得るために，式 (3.13) が使われている．また， $E[X_1]$ などを μ と置けるのは，母集団から要素を 1 個だけ取り出した時の期待値は母集団の中での平均値，すなわち母平均そのものであるからである．この結果を使うと，式 (6.4) も導かれる．

$$\begin{aligned} V[\overline{X}] =& E[(\overline{X} - E[\overline{X}])^2] \\ =& E\left[\left(\frac{1}{n}(X_1 + X_2 + ... + X_n) - \mu\right)^2\right] \\ =& \frac{1}{n^2}E\left[(X_1 + X_2 + ... + X_n - n\mu)^2\right] \\ =& \frac{1}{n^2}E\left[((X_1 - \mu) + (X_2 - \mu) + ... + (X_n - \mu))^2\right] \\ =& \frac{1}{n^2}E\left[(X_1 - \mu)^2 + (X_2 - \mu)^2 + ... + (X_n - \mu)^2\right] \\ & -\frac{2}{n^2}E\left[((X_1 - \mu)(X_2 - \mu) + (X_1 - \mu)(X_3 - \mu) + ...)\right] \\ =& \frac{1}{n^2}\left(E[(X_1 - \mu)^2] + E[(X_2 - \mu)^2] + \ldots + E[(X_n - \mu)^2]\right) - 0 \\ =& \frac{1}{n^2}(\sigma^2 + \sigma^2 + \ldots + \sigma^2) = \frac{\sigma^2}{n} \end{aligned} \tag{A.22}$$

ここでは，$X_1, X_2, ...$ が独立であるということを前提に，式 (3.18) が使われている．すなわち，標本抽出が無作為かつ復元的に[*4]行われていることが必要である．

A.6　標本分散の平均と母分散の関係

式 (6.5) は，次のようにやや技巧的な導き方で得られる．まず s^2 を変形しておく．

[*4] 有限集団からの非復元抽出では，前の結果が後の結果に影響を及ぼすので，独立性が失われる．

$$
\begin{aligned}
s^2 &= \frac{1}{n}\left((X_1-\overline{X})^2+(X_2-\overline{X})^2+\ldots+(X_n-\overline{X})^2\right) \\
&= \frac{1}{n}\left(X_1^2+X_2^2+\cdots+X_n^2\right)-\overline{X}^2 \ (2\text{ 乗の平均 }-\text{ 平均の 2 乗})
\end{aligned}
$$

ここで，

$$
E[X_1^2] = E[X_2^2]=\cdots=\sigma^2 \tag{A.23}
$$

また，

$$
E[\overline{X}^2] = V[\overline{X}]=\frac{\sigma^2}{n} \tag{A.24}
$$

これらを使って，

$$
\begin{aligned}
E[s^2] &= \frac{1}{n}(\sigma^2+\sigma^2+\cdots)-\frac{\sigma^2}{n} \\
&= \frac{n-1}{n}\sigma^2
\end{aligned}
$$

この導出の途中では式 (6.3), (6.4) の導出過程をも用いた．

A.7　最小二乗法

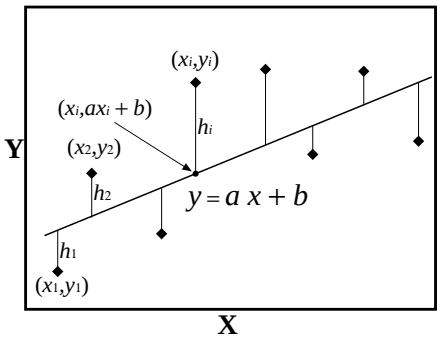

図 A.2　最小二乗法の原理: $h_1^2+h_2^2+\ldots$ が最小になるように a, b の値を決めてやる．

図 A.2 のようにデータ点 $(x_1,y_1),(x_2,y_2),\ldots,(x_n,y_n)$ が与えられていたとき，$y=ax+b$ で表される直線を引いたとしよう．このとき，i 番目のデータ点と直線の縦のずれ

を h_i とすると，

$$h_i = y_i - (ax_i + b) \tag{A.25}$$

となる．直線 $y = ax + b$ は，a と b の値を変化させることで，傾きを変えたり平行にずらしたりできる．それでは a, b がどのような値をとったときに，直線はデータ点をもっともよく近似できるだろうか．

そのためにまず次の S を定義しておこう．近似がもっともよい時には，S は極小になるはずである．

$$S = \frac{1}{n}(h_1^2 + h_2^2 + \ldots + h_n^2) \tag{A.26}$$

ここで右辺に $\frac{1}{n}$ を掛けているのは，後の計算をうまく処理するためである．

式 (A.26) に式 (A.25) を代入して整理すると，次の式が得られる．

$$S = \overline{y^2} + a^2\overline{x^2} - 2b\overline{y} - 2a\overline{xy} + 2ab\overline{x} + b^2 \tag{A.27}$$

a, b を変化させて S が最小になる条件を求めるには，次の 2 つの偏微分がゼロになればよい．

$$\begin{aligned} \frac{\partial S}{\partial a} &= 2a\overline{x^2} - 2\overline{xy} + 2b\overline{x} = 0 \\ \frac{\partial S}{\partial b} &= -2\overline{y} + 2a\overline{x} + 2b = 0 \end{aligned} \tag{A.28}$$

これを整理して，次のような連立方程式が得られる．ただしここでは a, b が未知数であることに注意！

$$\overline{x^2}a + \overline{x}b = \overline{xy} \tag{A.29}$$

$$\overline{x}a + b = \overline{y} \tag{A.30}$$

これを解くと次の結果が得られる．

$$a = \frac{\overline{xy} - \overline{x}\,\overline{y}}{\overline{x^2} - \overline{x}^2} = \frac{\sigma_{xy}}{\sigma_x^2} \tag{A.31}$$

$$b = \overline{y} - a\overline{x} \tag{A.32}$$

♡ **偏微分って何？** ♡

上の計算に出てきた**偏微分**を初めて見る人がいるかも知れない．これは複数の変数をもつ関数を微分するのに 1 つの変数だけで微分し，他の変数は定数として扱うものだ．だからある関数の偏微分は，変数の数だけある．例として，x, y を変数とする関数

$$f(x,y) = ax^2 + bxy + cy^2 + dy$$

を考える．このとき次のように x, y それぞれに関する偏微分が存在する．偏微分では，微分記号に通常の d ではなく，∂ を使う．

$$\begin{aligned} \frac{\partial f(x,y)}{\partial x} &= 2ax + by \\ \frac{\partial f(x,y)}{\partial y} &= bx + 2cy + d \end{aligned}$$

x による偏微分では y は定数とみなされるために，$f(x,y)$ の cy^2と dy はゼロになることに留意してほしい．

★