エコノメトリックス
(2016年度 後期 講義ノート)

平成 29 年 1 月 20 日 (金) 版

教科書『計量経済学』
(山本拓著，新世社，1995年)

谷﨑　久志
大阪大学・経済学部

目 次

教科書

『計量経済学』(山本拓著，1995，新世社)

『基本統計学 (第 3 版)』(豊田他著，東洋経済新報社，2010年)

1 計量経済学について

- 経済理論 (ミクロ，マクロ，財政，金融，国際経済，$\cdots$)
- データ (GNP，消費，投資，金利，為替レート，$\cdots$)

計量経済学 $\Longrightarrow$ 経済理論が現実に成り立つものかどうかを，データを用いて，統計的に検証する。

1.1 例 1： マクロの消費関数

$$C = f(Y)$$

ただし，C は消費，Y は所得。

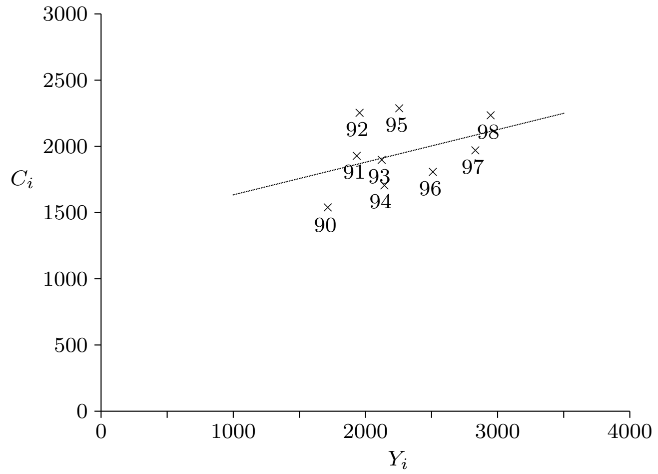
1. $Y \nearrow \Longrightarrow C \nearrow$
2. $\dfrac{\mathrm{d}C}{\mathrm{d}Y}$ = 限界消費性向 = 所得 1 円増加で消費が何円増加するか
3. すなわち，$\dfrac{\mathrm{d}C}{\mathrm{d}Y} > 0$

モデルの定式化

1. $C = a + bY$
2. $b = \dfrac{\mathrm{d}C}{\mathrm{d}Y}$ = 限界消費性向
3. a = 基礎消費 ($Y = 0$ のときに必要な消費)
4. 符号条件： $a > 0$，$b > 0$ (しかも，$1 > b$)

- この講義ノートは，
`http://www2.econ.osaka-u.ac.jp/~tanizaki/class/2016`
からダウンロード可。

図 1： 消費 (C_i) と所得 (Y_i)



1. $\times \longrightarrow$ 実際のデータ
2. $(Y_i, C_i) \Longrightarrow t$ 期のデータ, i.e., $i = 1, 2, \cdots, 9$
3. $i = 1 \Longrightarrow 1990$ 年,

   $i = 2 \Longrightarrow 1991$ 年,

   $\cdots,$

   $i = 9 \Longrightarrow 1998$ 年,

1. 実際のデータを用いて，a, b を求める。
2. a, b を求める $\equiv$ 現実の経済構造を求める
3. その結果，もし $a > 0$，$1 > b > 0$ なら，経済理論は現実経済を説明していると言える。

1.2　例 2： 日本酒の需要関数

$$Q = f(Y, P_1, P_2)$$

ただし，Q は日本酒の需要量，Y は所得，P_1 は日本酒の価格，P_2 は洋酒の価格。

1. $Y \nearrow \Longrightarrow Q \nearrow$,

   $P_1 \nearrow \Longrightarrow Q \searrow$,

   $P_2 \nearrow \Longrightarrow Q \nearrow$
2. $\dfrac{\partial Q}{\partial Y} > 0$, $\dfrac{\partial Q}{\partial P_1} < 0$, $\dfrac{\partial Q}{\partial P_2} > 0$
3. 日本酒と洋酒は代替財
4. モデルの定式化 (A)

$$Q = a + b_1 Y + b_2 P_1 + b_3 P_2$$

5. Q, Y, P_1, P_2 を用いて，a, b_1, b_2, b_3 を求める (日本酒の需要構造を求める)。
6. 符号条件： $b_1 > 0$, $b_2 < 0$, $b_3 > 0$, a ?
7. t 期のデータ $(Q_i, Y_i, P_{1i}, P_{2i})$
8. n 組のデータ, i.e., $i = 1, 2, \cdots, n$
9. モデルの定式化 (B)

$$Q = a + b_1 Y + b_2 \frac{P_1}{P_2}$$

   符号条件： $b_1 > 0$, $b_2 < 0$

10. モデルの定式化 (C)

$$\log(Q) = a + b_1 \log(Y) + b_2 \log(\frac{P_1}{P_2})$$

   符号条件： $b_1 > 0$, $b_2 < 0$

11. モデル (A), (B), (C) のどれが最も現実的かを得られた結果から判断する。

2　行列について

A を 2×2 行列とすると，

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

と表される。
$a_{ij} = A$ の第 i 行，第 j 列の要素

a を 2×1 行列 (縦ベクトル) とすると，

$$a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

と表される。
$a_i = a$ の第 i 要素

a を 1×2 行列 (横ベクトル) とすると，

$$a = \begin{pmatrix} a_1 & a_2 \end{pmatrix}$$

と表される。
$a_i = a$ の第 i 要素

A を $n \times k$ 行列とすると，

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nk} \end{pmatrix}$$

と表される。
$a_{ij} = A$ の第 i 行，第 j 列の要素 (ij 要素)

a を $n \times 1$ 行列 (縦ベクトル) とすると，

$$a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

と表される。
$a_i = a$ の第 i 要素

a を $1 \times k$ 行列 (横ベクトル) とすると，

$$a = \begin{pmatrix} a_1 & \cdots & a_k \end{pmatrix}$$

と表される。
$a_i = a$ の第 i 要素

行列の等号： A，B を $n \times k$ 行列とする。$A = B$ は，すべての $i = 1, \cdots, n$, $j = 1, \cdots, k$ について，$a_{ij} = b_{ij}$ を意味する。ただし，a_{ij}, b_{ij} は，それぞれ，A, B の ij 要素とする。
$x = 3$, $y = 2$ の 2 つの等式を行列で表す。

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \qquad \text{または} \qquad \begin{pmatrix} x & y \end{pmatrix} = \begin{pmatrix} 3 & 2 \end{pmatrix}$$

行列の和と差： A, B を $n \times k$ 行列とする。

$$\begin{aligned} A + B &= \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nk} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1k} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nk} \end{pmatrix} \\ &= \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1k} + b_{1k} \\ \vdots & \ddots & \vdots \\ a_{n1} + b_{n1} & \cdots & a_{nk} + b_{nk} \end{pmatrix} \end{aligned}$$

すなわち，$A + B$ の ij 要素は，$a_{ij} + b_{ij}$ となる。

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \qquad B = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$$

$$A + B = \begin{pmatrix} 1+5 & 2+6 \\ 3+7 & 4+8 \end{pmatrix} = \begin{pmatrix} 6 & 8 \\ 10 & 12 \end{pmatrix}$$

$$A - B = \begin{pmatrix} 1-5 & 2-6 \\ 3-7 & 4-8 \end{pmatrix} = \begin{pmatrix} -4 & -4 \\ -4 & -4 \end{pmatrix}$$

要素と行列の積： A を $n \times k$ 行列とする。c を スカラー (1×1 行列のこと) とする。

$$cA = c \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nk} \end{pmatrix} = \begin{pmatrix} ca_{11} & \cdots & ca_{1k} \\ \vdots & \ddots & \vdots \\ ca_{n1} & \cdots & ca_{nk} \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \qquad c = 5 \qquad \text{のとき}$$

$$cA = 5 \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 5 \times 1 & 5 \times 2 \\ 5 \times 3 & 5 \times 4 \end{pmatrix} = \begin{pmatrix} 5 & 10 \\ 15 & 20 \end{pmatrix}$$

行列と行列の積： A, B を $n \times k$，$k \times n$ 行列とする。

$$\begin{aligned} AB &= \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nk} \end{pmatrix} \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{k1} & \cdots & b_{kn} \end{pmatrix} \\ &= \begin{pmatrix} \sum_{m=1}^{k} a_{1m} b_{m1} & \cdots & \sum_{m=1}^{k} a_{1m} b_{mn} \\ \vdots & \ddots & \vdots \\ \sum_{m=1}^{k} a_{nm} b_{m1} & \cdots & \sum_{m=1}^{k} a_{1m} b_{mn} \end{pmatrix} \end{aligned}$$

すなわち，AB は $n \times n$ 行列で，AB の ij 要素は，$a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{ik} b_{kj} = \sum_{m=1}^{k} a_{ik} b_{kj}$ となる。

$$\begin{aligned} BA &= \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{k1} & \cdots & b_{kn} \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nk} \end{pmatrix} \\ &= \begin{pmatrix} \sum_{m=1}^{n} b_{1m} a_{m1} & \cdots & \sum_{m=1}^{n} b_{1m} a_{mk} \\ \vdots & \ddots & \vdots \\ \sum_{m=1}^{n} b_{km} a_{m1} & \cdots & \sum_{m=1}^{n} b_{1m} a_{mk} \end{pmatrix} \end{aligned}$$

すなわち，BA は $k \times k$ 行列で，BA の ij 要素は，$b_{i1} a_{1j} + b_{i2} a_{2j} + \cdots + b_{ik} a_{kj} = \sum_{m=1}^{k} a_{ik} b_{kj}$ となる。
このように，AB と BA の次元は異なる。

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \qquad B = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$$

$$
\begin{aligned}
AB &= \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \\
&= \begin{pmatrix} 1\times 5 + 2\times 7 & 1\times 6 + 2\times 8 \\ 3\times 5 + 4\times 7 & 3\times 6 + 4\times 8 \end{pmatrix} \\
&= \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
BA &= \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \\
&= \begin{pmatrix} 5\times 1 + 6\times 3 & 5\times 2 + 6\times 4 \\ 7\times 1 + 8\times 3 & 7\times 2 + 8\times 4 \end{pmatrix} \\
&= \begin{pmatrix} 23 & 34 \\ 31 & 46 \end{pmatrix}
\end{aligned}
$$

一般的に，$AB \neq BA$ となる。

c をスカラーとする。

$$cAB = AcB = (Ac)B = A(cB) = ABc$$

c をどこで掛けても値は変わらない。

連立方程式：

$$
\begin{cases} x + 2y = 3 \\ 4x + 5y = 6 \end{cases}
$$

行列表示すると，

$$
\begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}
$$

となる。

また，

$$
\begin{cases} x + 2y + 3z = 4 \\ 5x + 6y + 7z = 8 \\ 9x + 10y + 11z = 12 \end{cases}
$$

行列表示すると，

$$
\begin{pmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \\ 9 & 10 & 11 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 4 \\ 8 \\ 12 \end{pmatrix}
$$

となる。

単位行列：　単位行列とは，対角要素 1，その他 0 となる行列であり，I で表す。

$$
I = \begin{pmatrix} 1 & 0 & \cdots & & 0 \\ 0 & 1 & & & \\ \vdots & & \ddots & & \vdots \\ & & & 1 & 0 \\ 0 & & \cdots & 0 & 1 \end{pmatrix}
$$

I が $n \times n$ 行列のとき，I_n と書くことも多い。

A を $n \times n$ 行列，x を $n \times 1$ 行列 (ベクトル) とする。

$$I_n A = A I_n = A \qquad I_n x = x$$

$$
\begin{aligned}
&\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \\
&= \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \\
&= \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}
\end{aligned}
$$

$$
\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}
$$

逆行列：　A を $n \times n$ とする。A の逆行列とは，$AB = I_n$ または $BA = I_n$ となる B を指す。A も B も次元は同じ。B を A^{-1} と表す。
すなわち，A の逆行列は A^{-1} であり，A^{-1} の逆行列は A である。

$$
A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}
$$

のとき，

$$
A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}
$$

となる。

$$
\begin{aligned}
A^{-1}A &= \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \\
&= \frac{1}{ad - bc} \begin{pmatrix} da - bc & db - bd \\ -ca + ac & -bc + ad \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2
\end{aligned}
$$

$$
\begin{aligned}
AA^{-1} &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \times \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \\
&= \frac{1}{ad-bc} \begin{pmatrix} ad-bc & -ab+ba \\ cd-dc & -cb+da \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2
\end{aligned}
$$

連立方程式の解：　A を $n \times n$ 行列，x と b を $n \times 1$ 行列 (ベクトル) とする。

$$Ax = b$$

両辺に A^{-1} を左から掛ける。

$$A^{-1}Ax = A^{-1}b$$

$A^{-1}A = I_n$ なので，

$$I_n x = A^{-1}b$$

となる。また，

$$I_n x = x$$

なので，x を A, b で表すと，

$$x = A^{-1}b$$

となる。

例

$$
\begin{cases} x + 2y = 3 \\ 4x + 5y = 6 \end{cases}
$$

の行列表示は，

$$
\begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}
$$

となる。
x, y の解は，

$$
\begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}^{-1} \begin{pmatrix} 3 \\ 6 \end{pmatrix}
$$

なので，

$$
\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}^{-1} \begin{pmatrix} 3 \\ 6 \end{pmatrix}
$$

すなわち，

$$
\begin{aligned}
\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}^{-1} \begin{pmatrix} 3 \\ 6 \end{pmatrix} \\
&= \frac{1}{1 \times 5 - 2 \times 4} \begin{pmatrix} 5 & -2 \\ -4 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 6 \end{pmatrix} \\
&= -\frac{1}{1 \times 3} \begin{pmatrix} 5 \times 3 - 2 \times 6 \\ -4 \times 3 + 1 \times 6 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}
\end{aligned}
$$

例

$$
\begin{cases} x + 2y + 3z = 4 \\ 5x + 6y + 7z = 8 \\ 9x + 10y + 11z = 12 \end{cases}
$$

の行列表示は，

$$
\begin{pmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \\ 9 & 10 & 11 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 4 \\ 8 \\ 12 \end{pmatrix}
$$

となる。x, y, z の解は，

$$
\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \\ 9 & 10 & 11 \end{pmatrix}^{-1} \begin{pmatrix} 4 \\ 8 \\ 12 \end{pmatrix}
$$

となる。

転置行列：　A を $n \times k$ 行列とする。
A の ij 要素を a_{ij} とする。
A の転置行列 (A' または ${}^tA$) の ij 要素は，a_{ji} となる。

$$
A = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nk} \end{pmatrix}
$$

$$
A' = \begin{pmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & \ddots & \vdots \\ a_{1k} & \cdots & a_{nk} \end{pmatrix}
$$

A' は $k \times n$ となる。

$$(A')' = A$$

$$
x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad x' = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix}
$$

3　最小二乗法について

経済理論に基づいた線型モデルの係数の値をデータから求める時に用いられる手法 $\Longrightarrow$ 最小二乗法

3.1　最小二乗法と回帰直線

(X_1, Y_1), (X_2, Y_2), $\cdots$, (X_n, Y_n) のように n 組のデータがあり，X_i と Y_i との間に以下の線型関係を想定する。

$$Y_i = \alpha + \beta X_i,$$

X_i は説明変数，Y_i は被説明変数，α, β はパラメータとそれぞれ呼ばれる。
上の式は回帰モデル (または，回帰式) と呼ばれる。目的は，切片 α と傾き β をデータ $\{(X_i, Y_i),\ i = 1, 2, \cdots, n\}$ から推定すること，

データについて：

1. タイム・シリーズ (時系列)・データ： i が時間を表す (第 i 期)。
2. クロス・セクション (横断面)・データ： i が個人や企業を表す (第 i 番目の家計，第 i 番目の企業)。

3.2　切片 α と傾き β の推定

次のような関数 $S(\alpha, \beta)$ を定義する。

$$S(\alpha, \beta) = \sum_{i=1}^{n} u_i^2 = \sum_{i=1}^{n} (Y_i - \alpha - \beta X_i)^2$$

このとき，

$$\min_{\alpha, \beta} S(\alpha, \beta)$$

となるような α, β を求める (最小自乗法)。このときの解を $\widehat{\alpha}$, $\widehat{\beta}$ とする。
最小化のためには，

$$\frac{\partial S(\alpha, \beta)}{\partial \alpha} = 0$$

$$\frac{\partial S(\alpha, \beta)}{\partial \beta} = 0$$

を満たす α, β が $\widehat{\alpha}$, $\widehat{\beta}$ となる。

すなわち，$\widehat{\alpha}$, $\widehat{\beta}$ は，

$$\sum_{i=1}^{n} (Y_i - \widehat{\alpha} - \widehat{\beta} X_i) = 0, \tag{1}$$

$$\sum_{i=1}^{n} X_i (Y_i - \widehat{\alpha} - \widehat{\beta} X_i) = 0, \tag{2}$$

を満たす。
さらに，

$$\sum_{i=1}^{n} Y_i = n\widehat{\alpha} + \widehat{\beta} \sum_{i=1}^{n} X_i, \tag{3}$$

$$\sum_{i=1}^{n} X_i Y_i = \widehat{\alpha} \sum_{i=1}^{n} X_i + \widehat{\beta} \sum_{i=1}^{n} X_i^2,$$

行列表示によって，

$$\begin{pmatrix} \sum_{i=1}^n Y_i \\ \sum_{i=1}^n X_i Y_i \end{pmatrix} = \begin{pmatrix} n & \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i^2 \end{pmatrix} \begin{pmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{pmatrix},$$

逆行列の公式：

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

$\widehat{\alpha}$, $\widehat{\beta}$ について，まとめて，

$$\begin{aligned}
&\begin{pmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{pmatrix} \\
&= \begin{pmatrix} n & \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{i=1}^n Y_i \\ \sum_{i=1}^n X_i Y_i \end{pmatrix} \\
&= \frac{1}{n \sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2} \\
&\quad \times \begin{pmatrix} \sum_{i=1}^n X_i^2 & -\sum_{i=1}^n X_i \\ -\sum_{i=1}^n X_i & n \end{pmatrix} \begin{pmatrix} \sum_{i=1}^n Y_i \\ \sum_{i=1}^n X_i Y_i \end{pmatrix}
\end{aligned}$$

さらに，$\widehat{\beta}$ について解くと，

$$\begin{aligned}
\widehat{\beta} &= \frac{n \sum_{i=1}^n X_i Y_i - (\sum_{i=1}^n X_i)(\sum_{i=1}^n Y_i)}{n \sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2} \\
&= \frac{\sum_{i=1}^n X_i Y_i - n \overline{X}\overline{Y}}{\sum_{i=1}^n X_i^2 - n \overline{X}^2} \\
&= \frac{\sum_{i=1}^n (X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^n (X_i - \overline{X})^2}
\end{aligned}$$

連立方程式の (3) 式から，

$$\widehat{\alpha} = \overline{Y} - \widehat{\beta}\overline{X}$$

となる。ただし，

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i, \quad \overline{Y} = \frac{1}{n} \sum_{i=1}^{n} Y_i,$$

とする。

数値例：　以下の数値例を使って，回帰式 $Y_i = \alpha + \beta X_i$ の α，β の推定値 $\widehat{\alpha}$，$\widehat{\beta}$ を求める。

i	Y_i	X_i
1	6	10
2	9	12
3	10	14
4	10	16

$\widehat{\alpha}$，$\widehat{\beta}$ を求めるための公式は

$$\widehat{\beta} = \frac{\sum_{i=1}^n X_i Y_i - n\overline{X}\overline{Y}}{\sum_{i=1}^n X_i^2 - n\overline{X}^2}$$
$$\widehat{\alpha} = \overline{Y} - \widehat{\beta}\overline{X}$$

なので，必要なものは $\overline{X}$，$\overline{Y}$，$\displaystyle\sum_{i=1}^n X_i^2$，$\displaystyle\sum_{i=1}^n X_i Y_i$ である。

i	Y_i	X_i	$X_i Y_i$	X_i^2
1	6	10	60	100
2	9	12	108	144
3	10	14	140	196
4	10	16	160	256
合計	$\sum Y_i$	$\sum X_i$	$\sum X_i Y_i$	$\sum X_i^2$
	35	52	468	696
平均	$\overline{Y}$	$\overline{X}$		
	8.75	13		

よって，

$$\widehat{\beta} = \frac{468 - 4 \times 13 \times 8.75}{696 - 4 \times 13^2} = \frac{13}{20} = 0.65$$
$$\widehat{\alpha} = 8.75 - 0.65 \times 13 = 0.3$$

となる。

注意事項：

1. α, β は真の値で未知
2. $\widehat{\alpha}, \widehat{\beta}$ は α, β の推定値でデータから計算される

回帰直線は

$$\widehat{Y}_i = \widehat{\alpha} + \widehat{\beta} X_i,$$

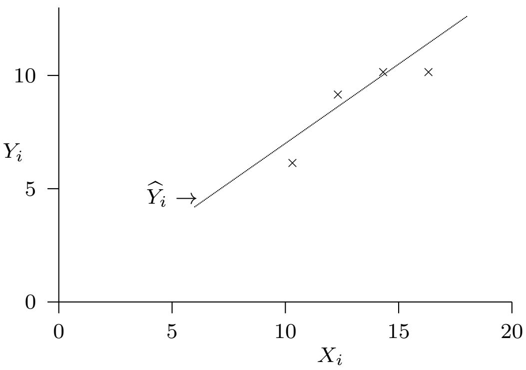
として与えられる。
上の数値例では，

$$\widehat{Y}_i = 0.3 + 0.65 X_i$$

となる。

i	Y_i	X_i	$X_i Y_i$	X_i^2	$\widehat{Y}_i$
1	6	10	60	100	6.8
2	9	12	108	144	8.1
3	10	14	140	196	9.4
4	10	16	160	256	10.7
合計	$\sum Y_i$	$\sum X_i$	$\sum X_i Y_i$	$\sum X_i^2$	$\sum \widehat{Y}_i$
	35	52	468	696	35.0
平均	$\overline{Y}$	$\overline{X}$			
	8.75	13			

図 2： Y_i，X_i，$\widehat{Y}_i$



$\widehat{Y}_i$ を実績値 Y_i の予測値または理論値と呼ぶ。

$$\widehat{u}_i = Y_i - \widehat{Y}_i,$$

$\widehat{u}_i$ を残差と呼ぶ。

$$Y_i = \widehat{Y}_i + \widehat{u}_i = \widehat{\alpha} + \widehat{\beta} X_i + \widehat{u}_i,$$

さらに，$\overline{Y}$ を両辺から引いて，

$$(Y_i - \overline{Y}) = (\widehat{Y}_i - \overline{Y}) + \widehat{u}_i,$$

3.3　残差 $\widehat{u}_i$ の性質について

$\widehat{u}_i = Y_i - \widehat{\alpha} - \widehat{\beta} X_i$ に注意して，(1) 式から，

$$\sum_{i=1}^n \widehat{u}_i = 0,$$

を得る。
(2) 式から，

$$\sum_{i=1}^n X_i \widehat{u}_i = 0,$$

を得る。
$\widehat{Y}_i = \widehat{\alpha} + \widehat{\beta} X_i$ から，

$$\sum_{i=1}^{n} \widehat{Y}_i \widehat{u}_i = 0,$$

を得る。なぜなら，

$$\begin{aligned}\sum_{i=1}^{n} \widehat{Y}_i \widehat{u}_i &= \sum_{i=1}^{n} (\widehat{\alpha} + \widehat{\beta} X_i) \widehat{u}_i \\ &= \widehat{\alpha} \sum_{i=1}^{n} \widehat{u}_i + \widehat{\beta} \sum_{i=1}^{n} X_i \widehat{u}_i \\ &= 0\end{aligned}$$

である。

i	Y_i	X_i	$\widehat{Y}_i$	$\widehat{u}_i$	$X_i \widehat{u}_i$	$\widehat{Y}_i \widehat{u}_i$
1	6	10	6.8	−0.8	−8.0	−5.44
2	9	12	8.1	0.9	10.8	7.29
3	10	14	9.4	0.6	8.4	5.64
4	10	16	10.7	−0.7	−11.2	−7.49
合計	$\sum Y_i$	$\sum X_i$	$\sum \widehat{Y}_i$	$\sum \widehat{u}_i$	$\sum X_i \widehat{u}_i$	$\sum \widehat{Y}_i \widehat{u}_i$
	35	52	35.0	0.0	0.0	0.00

3.4 決定係数 R^2 について

次の式

$$(Y_i - \overline{Y}) = (\widehat{Y}_i - \overline{Y}) + \widehat{u}_i,$$

の両辺を二乗して，総和すると，

$$\begin{aligned}&\sum_{i=1}^{n} (Y_i - \overline{Y})^2 \\ &= \sum_{i=1}^{n} \Big((\widehat{Y}_i - \overline{Y}) + \widehat{u}_i \Big)^2 \\ &= \sum_{i=1}^{n} (\widehat{Y}_i - \overline{Y})^2 + 2 \sum_{i=1}^{n} (\widehat{Y}_i - \overline{Y}) \widehat{u}_i + \sum_{i=1}^{n} \widehat{u}_i^2 \\ &= \sum_{i=1}^{n} (\widehat{Y}_i - \overline{Y})^2 + \sum_{i=1}^{n} \widehat{u}_i^2\end{aligned}$$

となる。まとめると，

$$\sum_{i=1}^{n} (Y_i - \overline{Y})^2 = \sum_{i=1}^{n} (\widehat{Y}_i - \overline{Y})^2 + \sum_{i=1}^{n} \widehat{u}_i^2$$

を得る。さらに，

$$1 = \frac{\sum_{i=1}^{n} (\widehat{Y}_i - \overline{Y})^2}{\sum_{i=1}^{n} (Y_i - \overline{Y})^2} + \frac{\sum_{i=1}^{n} \widehat{u}_i^2}{\sum_{i=1}^{n} (Y_i - \overline{Y})^2}$$

それぞれの項は，

1. $\sum_{i=1}^{n} (Y_i - \overline{Y})^2 \Longrightarrow y$ の全変動
2. $\sum_{i=1}^{n} (\widehat{Y}_i - \overline{Y})^2 \Longrightarrow \widehat{Y}_i$ (回帰直線) で説明される部分
3. $\sum_{i=1}^{n} \widehat{u}_i^2 \Longrightarrow \widehat{Y}_i$ (回帰直線) で説明されない部分

となる。
回帰式の当てはまりの良さを示す指標として，決定係数 R^2 を以下の通りに定義する。

$$R^2 = \frac{\sum_{i=1}^{n} (\widehat{Y}_i - \overline{Y})^2}{\sum_{i=1}^{n} (Y_i - \overline{Y})^2}$$

または，

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \widehat{u}_i^2}{\sum_{i=1}^{n} (Y_i - \overline{Y})^2},$$

として書き換えられる。
または，$Y_i = \widehat{Y}_i + \widehat{u}_i$ と

$$\begin{aligned}&\sum_{i=1}^{n} (\widehat{Y}_i - \overline{Y})^2 \\ &= \sum_{i=1}^{n} (\widehat{Y}_i - \overline{Y})(Y_i - \overline{Y} - \widehat{u}_i) \\ &= \sum_{i=1}^{n} (\widehat{Y}_i - \overline{Y})(Y_i - \overline{Y}) - \sum_{i=1}^{n} (\widehat{Y}_i - \overline{Y}) \widehat{u}_i \\ &= \sum_{i=1}^{n} (\widehat{Y}_i - \overline{Y})(Y_i - \overline{Y})\end{aligned}$$

を用いて，

$$\begin{aligned}R^2 &= \frac{\sum_{i=1}^{n} (\widehat{Y}_i - \overline{Y})^2}{\sum_{i=1}^{n} (Y_i - \overline{Y})^2} \\ &= \frac{\Big(\sum_{i=1}^{n} (\widehat{Y}_i - \overline{Y})^2\Big)^2}{\sum_{i=1}^{n} (Y_i - \overline{Y})^2 \sum_{i=1}^{n} (\widehat{Y}_i - \overline{Y})^2} \\ &= \left(\frac{\sum_{i=1}^{n} (\widehat{Y}_i - \overline{Y})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n} (Y_i - \overline{Y})^2 \sum_{i=1}^{n} (\widehat{Y}_i - \overline{Y})^2}} \right)^2\end{aligned}$$

と書き換えられる。すなわち，R^2 は Y_i と $\widehat{Y}_i$ の相関係数の二乗と解釈される。
$\sum_{i=1}^{n} (Y_i - \overline{Y})^2 = \sum_{i=1}^{n} (\widehat{Y}_i - \overline{Y})^2 + \sum_{i=1}^{n} \widehat{u}_i^2$ から，明らかに，

$$0 \leq R^2 \leq 1,$$

となる。R^2 が 1 に近づけば回帰式の当てはまりは良いと言える。しかし，t 分布のような数表は存在しない。したがって,「どの値よりも大きくなるべき」というような基準はない。
慣習的には，メドとして 0.9 以上を判断基準にする。

数値例： 決定係数の計算には以下の公式を用いる。

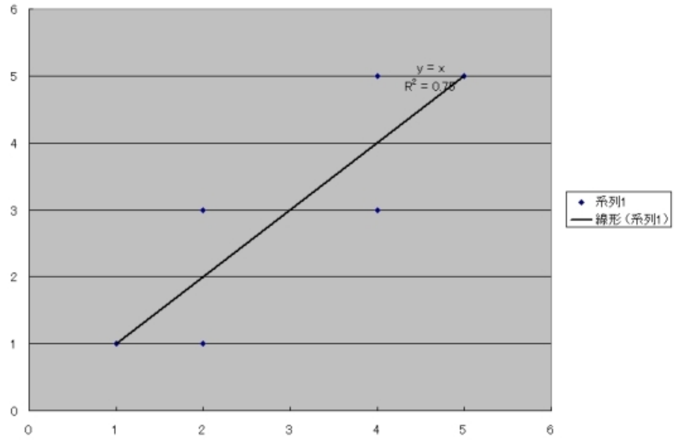
$$R^2 = 1 - \frac{\sum_{i=1}^n \widehat{u}_i^2}{\sum_{i=1}^n (Y_i - \overline{Y})^2} = 1 - \frac{\sum_{i=1}^n \widehat{u}_i^2}{\sum_{i=1}^n Y_i^2 - n\overline{Y}^2}$$

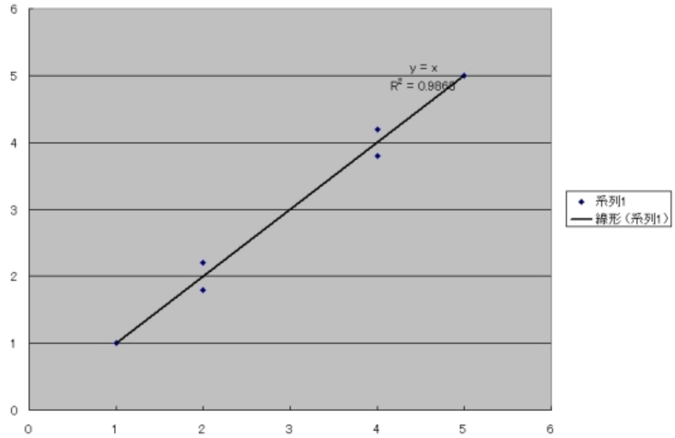
計算に必要なものは，$\widehat{u}_i = Y_i - (\widehat{\alpha} + \widehat{\beta}X_i)$，$\overline{Y}$，$\displaystyle\sum_{i=1}^n Y_i^2$ である。

i	Y_i	X_i	$\widehat{Y}_i$	$\widehat{u}_i$	$\widehat{u}_i^2$	Y_i^2
1	6	10	6.8	-0.8	0.64	36
2	9	12	8.1	0.9	0.81	81
3	10	14	9.4	0.6	0.36	100
4	10	16	10.7	-0.7	0.49	100
合計	$\sum Y_i$	$\sum X_i$	$\sum \widehat{Y}_i$	$\sum \widehat{u}_i$	$\sum \widehat{u}_i^2$	$\sum Y_i^2$
	35	52	35.0	0.0	2.30	317

$\sum \widehat{u}_i^2 = 2.30$，$\overline{X} = 13$，$\overline{Y} = 8.75$，$\displaystyle\sum_{i=1}^n Y_i^2 = 317$ なので，

$$R^2 = 1 - \frac{2.30}{317 - 4 \times 8.75^2} = 1 - \frac{2.30}{10.75} = 0.786$$





3.5 まとめ

$\widehat{\alpha}$，$\widehat{\beta}$ を求めるための公式は

$$\widehat{\beta} = \frac{\sum_{i=1}^n X_iY_i - n\overline{XY}}{\sum_{i=1}^n X_i^2 - n\overline{X}^2}$$
$$\widehat{\alpha} = \overline{Y} - \widehat{\beta}\overline{X}$$

なので，必要なものは $\overline{X}$，$\overline{Y}$，$\displaystyle\sum_{i=1}^n X_i^2$，$\displaystyle\sum_{i=1}^n X_iY_i$ である。
決定係数の計算には以下の公式を用いる。

$$R^2 = 1 - \frac{\sum_{i=1}^n \widehat{u}_i^2}{\sum_{i=1}^n (Y_i - \overline{Y})^2} = 1 - \frac{\sum_{i=1}^n \widehat{u}_i^2}{\sum_{i=1}^n Y_i^2 - n\overline{Y}^2}$$

計算に必要なものは，$\sum \widehat{u}_i^2$，$\overline{Y}$，$\displaystyle\sum_{i=1}^n Y_i^2$ である。

4 統計学の回帰分析への応用

$(X_1, Y_1), (X_2, Y_2), \cdots, (X_n, Y_n)$ のように n 組のデータがあり，X_i と Y_i との間に線型関係を想定する。

$$Y_i = \alpha + \beta X_i$$

最小二乗法を用いて，データに直線のあてはめを行った。
$\widehat{\alpha}$，$\widehat{\beta}$，$\widehat{Y}_i$ を求めるための公式は

$$\widehat{\beta} = \frac{\sum_{i=1}^n (X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^n (X_i - \overline{X})^2} = \frac{\sum_{i=1}^n (X_i - \overline{X})Y_i}{\sum_{i=1}^n (X_i - \overline{X})^2}$$
$$\widehat{\alpha} = \overline{Y} - \widehat{\beta}\overline{X},$$
$$\widehat{Y}_i = \widehat{\alpha} + \widehat{\beta}X_i,$$

である。

Y_i，$\widehat{Y}_i$，$\widehat{u}_i$，$\widehat{\alpha}$，$\widehat{\beta}$ の関係は以下の通りである。

$$\begin{aligned} Y_i &= \widehat{Y}_i + \widehat{u}_i \\ &= \widehat{\alpha} + \widehat{\beta} X_i + \widehat{u}_i \end{aligned}$$

残差 $\widehat{u}_i$ が必ず含まれることから，

$$Y_i = \alpha + \beta X_i + u_i,$$

として誤差項 (または，撹乱項) u_i を含め，それを確率変数として考える。

$\Longrightarrow$ 確率的モデル

Y_i： 被説明変数，従属変数

X_i： 説明変数，独立変数

α, β： 未知母数 (未知パラメータ)

$\widehat{\alpha}, \widehat{\beta}$： 推定量 (特に，最小二乗推定量)

1. 残差 $\widehat{u}_i$ は u_i の実現値としてみなすことができる。
2. $\widehat{\alpha}$，$\widehat{\beta}$ の性質を統計学的に考察可能となる。

統計学の復習 (統計量，推定量，推定値について)

1. 理論標本，理論観測値

   $\Longrightarrow X_1, X_2, \cdots, X_n$

   $\Longrightarrow$ 確率変数

2. 実現された標本，実現された観測値，実現値

   $\Longrightarrow x_1, x_2, \cdots, x_n$

   $\Longrightarrow$ 数値

1. 理論観測値 $X_1, X_2, \cdots, X_n$ の関数 $\Longrightarrow$ 統計量
2. すべての i について，$\mu = \mathrm{E}(X_i)$ と仮定する。
3. 母平均 μ の推定に使われる統計量 $\Longrightarrow$ μ の推定量
   (a) $\overline{X} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$ は μ の推定量
   (b) $S^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}(X_i - \overline{X})^2$ は σ^2 の推定量
4. 実現された標本を用いて実際に計算された推定量の値 $\Longrightarrow$ 推定値
   (a) $\overline{x} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} x_i$ は μ の推定値
   (b) $s^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}(x_i - \overline{x})^2$ は σ^2 の推定値
5. μ や σ^2 の推定量の候補は無数に考えられる。
6. α, β は母数。
7. $\widehat{\alpha}, \widehat{\beta}$ は α, β の推定量である。

4.1 回帰モデルの仮定

回帰モデル

$$Y_i = \alpha + \beta X_i + u_i,$$

の仮定：

1. X_i は確率変数でないと仮定する (固定された値)。
2. すべての i について，$\mathrm{E}(u_i) = 0$ とする。
3. すべての i について，$\mathrm{V}(u_i) = \sigma^2$ とする。($\mathrm{V}(u_i) = \mathrm{E}(u_i^2) = \sigma^2$ に注意)
4. すべての $i \neq j$ について，$\mathrm{Cov}(u_i, u_j) = 0$ とする。($\mathrm{Cov}(u_i, u_j) = \mathrm{E}(u_i u_j) = 0$ に注意)
5. すべての i について，$u_i \sim N(0, \sigma^2)$ とする。
6. $n \longrightarrow \infty$ のとき，$\sum_{i=1}^{n}(X_i - \overline{X})^2 \longrightarrow \infty$ とする。

撹乱項 $u_1, u_2, \cdots, u_n$ はそれぞれ独立に平均ゼロ，分散 σ^2 の正規分布する。

再度，まとめて，回帰モデル：

$$Y_i = \alpha + \beta X_i + u_i, \quad u_i \sim N(0, \sigma^2),$$

ただし，

Y_i： 被説明変数，従属変数

X_i： 説明変数，独立変数

α, β, σ^2： 未知母数 (未知パラメータ)

$\widehat{\alpha}, \widehat{\beta}$： 推定量 (特に，最小二乗推定量)

特に，回帰直線は，

$$\mathrm{E}(Y_i) = \alpha + \beta X_i$$

として解釈される。

4.2 誤差項 (攪乱項) の経済学的意味

1. 経済理論自身が不完全： X 以外にも他の説明変数が必要であるにもかかわらず，それを誤って除いている可能性がある。
2. モデルの定式化が不完全： Y と X との間の線形関係が誤りかもしれない。
3. 理論モデルとデータとの対応： 理論モデルで考えられる変数と実際に用いたデータが適当でないかもしれない。例： 所得のデータについては国民総生産，国民所得，可処分所得，労働所得 $\cdots$，金利では公定歩合，国債利回り，定期預金金利，全国銀行平均約定金利 $\cdots$
4. 測定上の誤差： 経済データは一般的に推計されているため完全ではない。誤差を含む。

4.3 $\widehat{\alpha}$，$\widehat{\beta}$ の統計的性質

準備：

$$Y_i = \alpha + \beta X_i + u_i,$$

$$\overline{Y} = \alpha + \beta\overline{X} + \overline{u},$$

ただし，

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^n Y_i,$$
$$\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i,$$
$$\overline{u} = \frac{1}{n}\sum_{i=1}^n u_i,$$

とする。辺々を引いて，

$$Y_i - \overline{Y} = \beta(X_i - \overline{X}) + (u_i - \overline{u}),$$

を得る。

4.3.1 $\widehat{\beta}$ について

β の最小二乗推定量 $\widehat{\beta}$ に代入すると，

$$\begin{aligned}
\widehat{\beta} &= \frac{\sum_{i=1}^n (X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^n (X_i - \overline{X})^2} \\
&= \frac{\sum_{i=1}^n (X_i - \overline{X})\Big(\beta(X_i - \overline{X}) + (u_i - \overline{u})\Big)}{\sum_{i=1}^n (X_i - \overline{X})^2} \\
&= \frac{\sum_{i=1}^n (X_i - \overline{X})\Big(\beta(X_i - \overline{X})\Big)}{\sum_{i=1}^n (X_i - \overline{X})^2} \\
&\quad + \frac{\sum_{i=1}^n (X_i - \overline{X})(u_i - \overline{u})}{\sum_{i=1}^n (X_i - \overline{X})^2} \\
&= \beta + \frac{\sum_{i=1}^n (X_i - \overline{X})u_i}{\sum_{i=1}^n (X_i - \overline{X})^2}
\end{aligned}$$

である。途中の計算で，$\sum_{i=1}^n (X_i - \overline{X})\overline{u} = 0$ に注意せよ。よって，まとめると，

$$\begin{aligned}
\widehat{\beta} &= \frac{\sum_{i=1}^n (X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^n (X_i - \overline{X})^2} \\
&= \beta + \frac{\sum_{i=1}^n (X_i - \overline{X})u_i}{\sum_{i=1}^n (X_i - \overline{X})^2} \\
&= \beta + \sum_{i=1}^n \omega_i u_i,
\end{aligned}$$

となる。ただし，$\omega_i = \dfrac{(X_i - \overline{X})}{\sum_{i=1}^n (X_i - \overline{X})^2}$ とする。

4.3.2 $\widehat{\alpha}$ について

α の最小二乗推定量 $\widehat{\alpha}$ については，

$$\begin{aligned}
\widehat{\alpha} &= \overline{Y} - \widehat{\beta}\overline{X} \\
&= \alpha - (\widehat{\beta} - \beta)\overline{X} + \overline{u}
\end{aligned}$$

ただし，$\overline{u} = \dfrac{1}{n}\displaystyle\sum_{i=1}^n u_i$ である。$\overline{Y} = \alpha + \beta\overline{X} + \overline{u}$ を途中で使う。

4.3.3 $\widehat{\alpha}$，$\widehat{\beta}$ の平均

統計学の復習 (期待値の公式)：

1. X を確率変数とする。

$$\mathrm{E}(a + bX) = a + b\mathrm{E}(X),$$

となる。ただし，a, b は定数とする。

2. X_1, X_2, $\cdots$, X_n の n 個の確率変数を考える。このとき，

$$\mathrm{E}(\sum_{i=1}^{n} c_i X_i) = \sum_{i=1}^{n} \mathrm{E}(c_i X_i), = \sum_{i=1}^{n} c_i \mathrm{E}(X_i),$$

となる。ただし，$c_1, c_2, \cdots, c_n$ は定数とする。

$\widehat{\alpha}$，$\widehat{\beta}$ の平均： $\widehat{\beta}$ は次のように書き換えられた。

$$\widehat{\beta} = \beta + \sum_{i=1}^{n} \omega_i u_i,$$

の両辺に期待値をとると，

$$\begin{aligned}
\mathrm{E}(\widehat{\beta}) &= \mathrm{E}(\beta + \sum_{i=1}^{n} \omega_i u_i) \\
&= \beta + \mathrm{E}(\sum_{i=1}^{n} \omega_i u_i) \\
&= \beta + \sum_{i=1}^{n} \mathrm{E}(\omega_i u_i) \\
&= \beta + \sum_{i=1}^{n} \omega_i \mathrm{E}(u_i) \\
&= \beta,
\end{aligned}$$

となり，$\widehat{\beta}$ は β の不偏推定量であると言える。
$\widehat{\alpha}$ については，

$$\widehat{\alpha} = \alpha - (\widehat{\beta} - \beta)\overline{X} + \overline{u}$$

を利用して，辺々に期待値をとると，

$$\begin{aligned}
\mathrm{E}(\widehat{\alpha}) &= \alpha - \mathrm{E}(\widehat{\beta} - \beta)\overline{X} + \mathrm{E}(\overline{u}) \\
&= \alpha
\end{aligned}$$

となる。$\mathrm{E}(\widehat{\beta} - \beta) = 0$ に注意。また，$\mathrm{E}(\overline{u})$ の計算は以下のとおり。

$$\begin{aligned}
\mathrm{E}(\overline{u}) &= \mathrm{E}(\frac{1}{n}\sum_{i=1}^{n} u_i) \\
&= \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}(u_i) \\
&= 0
\end{aligned}$$

$\widehat{\alpha}$ は α の不偏推定量であると言える。

4.3.4 $\widehat{\alpha}$，$\widehat{\beta}$ の分散

統計学の復習 (分散の公式)：

1. X を確率変数とする。

$$\mathrm{V}(X) = \mathrm{E}(X - \mu)^2,$$

となる。ただし，$\mu = \mathrm{E}(X)$ とする。

2. X を確率変数とする。

$$\mathrm{V}(a + bX) = \mathrm{V}(bX) = b^2 \mathrm{V}(X),$$

となる。ただし，a, b は定数とする。

3. X_1, X_2, $\cdots$, X_n の n 個の確率変数は互いに独立とする。このとき，

$$\mathrm{V}(\sum_{i=1}^{n} c_i X_i) = \sum_{i=1}^{n} \mathrm{V}(c_i X_i) = \sum_{i=1}^{n} c_i^2 \mathrm{V}(X_i),$$

となる。ただし，$c_1, c_2, \cdots, c_n$ は定数とする。

$\widehat{\alpha}$，$\widehat{\beta}$ の分散： $\widehat{\beta}$ の分散について，$\widehat{\beta} = \beta + \sum_{i=1}^{n} \omega_i u_i$ を用いると，

$$\begin{aligned}
\mathrm{V}(\widehat{\beta}) &= \mathrm{V}(\beta + \sum_{i=1}^{n} \omega_i u_i) \\
&= \mathrm{V}(\sum_{i=1}^{n} \omega_i u_i) \\
&= \sum_{i=1}^{n} \omega_i^2 \mathrm{V}(u_i^2) \\
&= \sigma^2 \sum_{i=1}^{n} \omega_i^2 \\
&= \frac{\sigma^2}{\sum_{i=1}^{n} (X_i - \overline{X})^2}
\end{aligned}$$

誤差項 (または，攪乱項) の仮定より，

$$\mathrm{V}(u_i) = \sigma^2,$$

を用いる。
最後の行は，$\omega_i = \dfrac{X_i - \overline{X}}{\sum_{i=1}^{n} (X_i - \overline{X})^2}$ に注意して，

$$\sum_{i=1}^{n} \omega_i^2 = \sum_{i=1}^{n} \left(\frac{X_i - \overline{X}}{\sum_{i=1}^{n} (X_i - \overline{X})^2} \right)^2$$

$$
\begin{aligned}
&= \sum_{i=1}^{n} \frac{(X_i - \overline{X})^2}{\left(\sum_{i=1}^{n}(X_i - \overline{X})^2\right)^2} \\
&= \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\left(\sum_{i=1}^{n}(X_i - \overline{X})^2\right)^2} \\
&= \frac{1}{\sum_{i=1}^{n}(X_i - \overline{X})^2}
\end{aligned}
$$

を用いる。

よって，$\widehat{\beta}$ の平均は β，分散は $\dfrac{\sigma^2}{\sum(X_i - \overline{X})^2}$ となることが示された。

$\widehat{\alpha}$ の分散について，$\widehat{\alpha} = \alpha - (\widehat{\beta} - \beta)\overline{X} + \overline{u}$ を利用すると，

$$
\begin{aligned}
\mathrm{V}(\widehat{\alpha}) &= \mathrm{E}(\widehat{\alpha} - \alpha)^2 \\
&= \mathrm{E}\big(-(\widehat{\beta} - \beta)\overline{X} + \overline{u}\big)^2 \\
&= \overline{X}^2 \mathrm{E}(\widehat{\beta} - \beta)^2 - 2\overline{X}\mathrm{E}\big((\widehat{\beta} - \beta)\overline{u}\big) + \mathrm{E}(\overline{u}^2) \\
&= \sigma^2 \left(\frac{\overline{X}^2}{\sum_{i=1}^{n}(X_i - \overline{X})^2} + \frac{1}{n} \right) \\
&= \frac{\sigma^2 \sum_{i=1}^{n} X_i^2}{n \sum_{i=1}^{n}(X_i - \overline{X})^2}
\end{aligned}
$$

途中で，以下の計算が使われる。

$$
\begin{aligned}
&\mathrm{E}\big((\widehat{\beta} - \beta)\overline{u}\big) \\
&= \mathrm{E}\left(\frac{\sum_{i=1}^{n}(X_i - \overline{X})u_i}{\sum_{i=1}^{n}(X_i - \overline{X})^2} \frac{1}{n} \sum_{i=1}^{n} u_i \right) \\
&= \frac{1}{n \sum_{i=1}^{n}(X_i - \overline{X})^2} \mathrm{E}\left(\sum_{i=1}^{n}(X_i - \overline{X})u_i \sum_{j=1}^{n} u_j \right) \\
&= \frac{1}{n \sum_{i=1}^{n}(X_i - \overline{X})^2} \sum_{i=1}^{n}\sum_{j=1}^{n}(X_i - \overline{X})\mathrm{E}(u_i u_j) \\
&= \frac{1}{n \sum_{i=1}^{n}(X_i - \overline{X})^2} \sigma^2 \sum_{i=1}^{n}(X_i - \overline{X}) \\
&= 0
\end{aligned}
$$

$\sum_{i=1}^{n}(X_i - \overline{X}) = 0$ であることに注意。

$$
\begin{aligned}
\mathrm{E}(\overline{u}^2) &= \mathrm{E}(\frac{1}{n}\sum_{i=1}^{n} u_i)^2 \\
&= \frac{1}{n^2}\mathrm{E}(\sum_{i=1}^{n}\sum_{j=1}^{n} u_i u_j) \\
&= \frac{1}{n^2}\mathrm{E}(\sum_{i=1}^{n} u_i^2) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{E}(u_i^2) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 \\
&= \frac{\sigma^2}{n}
\end{aligned}
$$

よって，$\widehat{\alpha}$ の平均は α，分散は $\dfrac{\sigma^2 \sum_{i=1}^{n} X_i^2}{n \sum_{i=1}^{n}(X_i - \overline{X})^2}$ となることが示された。

$\widehat{\alpha}$ と $\widehat{\beta}$ の共分散について，$\widehat{\alpha} = \alpha - (\widehat{\beta} - \beta)\overline{X} + \overline{u}$ を利用すると，

$$
\begin{aligned}
\mathrm{Cov}(\widehat{\alpha}, \widehat{\beta}) &= \mathrm{E}\big((\widehat{\alpha} - \alpha)(\widehat{\beta} - \beta)\big) \\
&= \mathrm{E}\big((-(\widehat{\beta} - \beta)\overline{X} + \overline{u})(\widehat{\beta} - \beta)\big) \\
&= -\mathrm{E}(\widehat{\beta} - \beta)^2\overline{X} + \mathrm{E}\big(\overline{u}(\widehat{\beta} - \beta)\big) \\
&= -\mathrm{E}(\widehat{\beta} - \beta)^2\overline{X} \\
&= -\frac{\sigma^2\overline{X}}{\sum_{i=1}^{n}(X_i - \overline{X})^2}
\end{aligned}
$$

となる。

数値例：

i	Y_i	X_i	X_iY_i	X_i^2
1	6	10	60	100
2	9	12	108	144
3	10	14	140	196
4	10	16	160	256
合計	$\sum Y_i$	$\sum X_i$	$\sum X_iY_i$	$\sum X_i^2$
	35	52	468	696
平均	$\overline{Y}$	$\overline{X}$		
	8.75	13		

$$
\begin{aligned}
\mathrm{V}(\widehat{\beta}) &= \frac{\sigma^2}{\sum_{i=1}^{n}(X_i - \overline{X})^2} \\
&= \frac{\sigma^2}{\sum_{i=1}^{n} X_i^2 - n\overline{X}^2} \\
&= \frac{\sigma^2}{696 - 4 \times 13^2} \\
&= \frac{\sigma^2}{20} \\
&= 0.05\sigma^2
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{V}(\widehat{\alpha}) &= \frac{\sigma^2 \sum_{i=1}^n X_i^2}{n\sum_{i=1}^n (X_i - \overline{X})^2} \\
&= \frac{\sigma^2 \sum_{i=1}^n X_i^2}{n(\sum_{i=1}^n X_i^2 - n\overline{X}^2)} \\
&= \frac{\sigma^2 696}{4(696 - 4\times 13^2)} \\
&= \frac{696\sigma^2}{80} \\
&= 8.7\sigma^2
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Cov}(\widehat{\alpha}, \widehat{\beta}) &= -\frac{\sigma^2\overline{X}}{\sum_{i=1}^n (X_i - \overline{X})^2} \\
&= -\frac{\sigma^2\overline{X}}{\sum_{i=1}^n X_i^2 - n\overline{X}^2} \\
&= -\frac{13\sigma^2}{696 - 4\times 13^2} \\
&= -0.65\sigma^2
\end{aligned}
$$

注意：　最小二乗法を復習すると，まず，次のような関数 $S(\alpha, \beta)$ を定義する。

$$
S(\alpha, \beta) = \sum_{i=1}^n u_i^2 = \sum_{i=1}^n (Y_i - \alpha - \beta X_i)^2
$$

$S(\alpha, \beta)$ の最小化によって，

$$
\frac{\partial S(\alpha, \beta)}{\partial \alpha} = 0
$$

$$
\frac{\partial S(\alpha, \beta)}{\partial \beta} = 0
$$

を満たす α, β が $\widehat{\alpha}, \widehat{\beta}$ となる。
すなわち，$\widehat{\alpha}, \widehat{\beta}$ は，

$$
\begin{aligned}
&\sum_{i=1}^n (Y_i - \widehat{\alpha} - \widehat{\beta}X_i) = 0, \\
&\sum_{i=1}^n X_i(Y_i - \widehat{\alpha} - \widehat{\beta}X_i) = 0
\end{aligned}
$$

を満たす。
さらに，

$$
\begin{aligned}
&\sum_{i=1}^n Y_i = n\widehat{\alpha} + \widehat{\beta}\sum_{i=1}^n X_i, \\
&\sum_{i=1}^n X_iY_i = \widehat{\alpha}\sum_{i=1}^n X_i + \widehat{\beta}\sum_{i=1}^n X_i^2,
\end{aligned}
$$

行列表示によって，

$$
\begin{pmatrix} \sum_{i=1}^n Y_i \\ \sum_{i=1}^n X_iY_i \end{pmatrix} = \begin{pmatrix} n & \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i^2 \end{pmatrix} \begin{pmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{pmatrix},
$$

$\widehat{\alpha}, \widehat{\beta}$ について，まとめて，

$$
\begin{aligned}
&\begin{pmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{pmatrix} \\
&= \begin{pmatrix} n & \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{i=1}^n Y_i \\ \sum_{i=1}^n X_iY_i \end{pmatrix} \\
&= \frac{1}{n\sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2} \\
&\quad \times \begin{pmatrix} \sum_{i=1}^n X_i^2 & -\sum_{i=1}^n X_i \\ -\sum_{i=1}^n X_i & n \end{pmatrix} \begin{pmatrix} \sum_{i=1}^n Y_i \\ \sum_{i=1}^n X_iY_i \end{pmatrix}
\end{aligned}
$$

逆行列の部分と分散，共分散とは以下のような関係がある。

$$
\begin{aligned}
&\begin{pmatrix} \mathrm{V}(\widehat{\alpha}) & \mathrm{Cov}(\widehat{\alpha}, \widehat{\beta}) \\ \mathrm{Cov}(\widehat{\alpha}, \widehat{\beta}) & \mathrm{V}(\widehat{\beta}) \end{pmatrix} \\
&= \sigma^2 \begin{pmatrix} n & \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i^2 \end{pmatrix}^{-1} \\
&= \frac{\sigma^2}{n\sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2} \\
&\quad \times \begin{pmatrix} \sum_{i=1}^n X_i^2 & -\sum_{i=1}^n X_i \\ -\sum_{i=1}^n X_i & n \end{pmatrix} \\
&= \begin{pmatrix} \dfrac{\sigma^2 \sum_{i=1}^n X_i^2}{n\sum(X_i - \overline{X})^2} & -\dfrac{\sigma^2\overline{X}}{\sum(X_i - \overline{X})^2} \\ -\dfrac{\sigma^2\overline{X}}{\sum(X_i - \overline{X})^2} & \dfrac{\sigma^2}{\sum(X_i - \overline{X})^2} \end{pmatrix}
\end{aligned}
$$

4.3.5　$\widehat{\alpha}$，$\widehat{\beta}$ の分布 (σ^2 が既知の場合)

統計学の復習 (正規分布について):

1. n 個の独立な確率変数 $X_1, X_2, \cdots, X_n$ が同一の分布に従うものとする。このとき，

   $$
   \begin{aligned}
   \mathrm{E}(\sum_{i=1}^n c_iX_i) &= \sum_{i=1}^n c_i\mathrm{E}(X_i), \\
   \mathrm{V}(\sum_{i=1}^n c_iX_i) &= \sum_{i=1}^n c_i^2\mathrm{V}(X_i),
   \end{aligned}
   $$

   となる。

2. n 個の独立な確率変数 $X_1, X_2, \cdots, X_n$ が同一の正規分布に従うものとする。このとき，

   $$
   \sum_{i=1}^n c_iX_i \sim N\Big(\mathrm{E}(\sum_{i=1}^n c_iX_i), \mathrm{V}(\sum_{i=1}^n c_iX_i)\Big)
   $$

となる。すなわち，

$$\sum_{i=1}^{n} c_i X_i \sim N\Big(\sum_{i=1}^{n} c_i \mathrm{E}(X_i), \sum_{i=1}^{n} c_i^2 \mathrm{V}(X_i)\Big)$$

3. 特に，$X_i \sim N(\mu, \sigma^2)$，標本平均 $\overline{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$ を考えると，

$$\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$$

となる。(すべての i について，$c_i = \dfrac{1}{n}$ の場合を考えればよい。)

$\widehat{\alpha}$，$\widehat{\beta}$ の分布:

1. $\widehat{\beta} = \beta + \sum_{i=1}^{n} \omega_i u_i$
2. $\mathrm{E}(\widehat{\beta}) = \beta$
3. $\mathrm{V}(\widehat{\beta}) = \dfrac{\sigma^2}{\sum_{i=1}^{n}(X_i - \overline{X})^2}$

よって，

$$\widehat{\beta} \sim N\left(\beta, \frac{\sigma^2}{\sum_{i=1}^{n}(X_i - \overline{X})^2}\right),$$

となる。

1. $\widehat{\alpha} = \alpha - (\widehat{\beta} - \beta)\overline{X} + \overline{u}$
2. $\mathrm{E}(\widehat{\alpha}) = \alpha$
3. $\mathrm{V}(\widehat{\alpha}) = \sigma^2\left(\dfrac{\overline{X}^2}{\sum_{i=1}^{n}(X_i - \overline{X})^2} + \dfrac{1}{n}\right)$

よって，

$$\widehat{\alpha} \sim N\left(\alpha, \sigma^2\left(\frac{\overline{X}^2}{\sum_{i=1}^{n}(X_i - \overline{X})^2} + \frac{1}{n}\right)\right),$$

となる。

4.3.6 $\widehat{\alpha}$，$\widehat{\beta}$ の性質：最良線型不偏性と一致性

統計学の復習 (推定量の望ましい性質)： $\widehat{\alpha}$，$\widehat{\beta}$ の性質を求めるために

1. **不偏性：**

   ある母集団のある母数 θ に対して，θ の推定量として $\widehat{\theta}$ を考える。

   このとき，

   $$\mathrm{E}(\widehat{\theta}) = \theta$$

   となるとき，$\widehat{\theta}$ は θ の不偏推定量であると言う。

   $\widehat{\theta}$ は不偏性を持つと言う。

   $\mathrm{E}(\widehat{\theta}) - \theta$ は偏りと定義される。

   (a) 標本平均 $\overline{X}$ は μ の不偏推定量である。

   証明：

   $$\begin{aligned}\mathrm{E}(\overline{X}) &= \mathrm{E}(\frac{1}{n}\sum_{i=1}^{n} X_i) \\ &= \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}(X_i) \\ &= \frac{1}{n}\sum_{i=1}^{n}\mu \\ &= \mu\end{aligned}$$

   このように，$\mathrm{E}(\overline{X}) = \mu$ なので，標本平均 $\overline{X}$ は μ の不偏推定量となる。

2. **有効性 (最小分散性)：**

   ある母数 θ に対して，$\widehat{\theta}_1$ と $\widehat{\theta}_2$ の 2 つの不偏推定量を考える。

   このとき，$\mathrm{V}(\widehat{\theta}_1) \le \mathrm{V}(\widehat{\theta}_2)$ が成り立つとき，$\widehat{\theta}_1$ は $\widehat{\theta}_2$ より有効であると言う。

   ある母数 θ に対して，可能なすべての不偏推定量を考え，$\widehat{\theta}$ が最も小さな分散を持つ不偏推定量であるとする。

   このとき，$\widehat{\theta}$ を最小分散不偏推定量，または，最良不偏推定量と言う。

   (a) 推定量 $\sum_{i=1}^{n} c_i X_i$ の中で，$\overline{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$ が最も小さな分散を持つ推定量となる。

   $\Longrightarrow$ 最良線型不偏推定量

3. **一致性：**

   ある母数 θ について推定量 $\widehat{\theta}$ を考える。n 個の標本から構成された推定量を $\widehat{\theta}^{(n)}$ と定義する。

数列 $\widehat{\theta}^{(1)}, \widehat{\theta}^{(2)}, \cdots, \widehat{\theta}^{(n)}, \cdots$ を考える。

十分大きな n について，$\widehat{\theta}^{(n)}$ が θ に確率的に収束するとき，$\widehat{\theta}$ は θ の一致推定量であると言う。

$$\operatorname*{plim}_{n\to\infty} \widehat{\theta} = \theta$$

と表現する。

(a) $\mathrm{E}(\widehat{\theta}) = \theta$ とする。$n \to \infty$ のとき，$\mathrm{V}(\widehat{\theta}) \to 0$ が成り立てば，$\widehat{\theta}$ は θ の一致推定量である。

(b) μ の推定量 $\overline{X}$ を調べる。

$$\mathrm{E}(\overline{X}) = \mu$$

である。

$$\mathrm{V}(\overline{X}) = \frac{\sigma^2}{n}$$

となる。$n \to \infty$ のとき，

$$\mathrm{V}(\overline{X}) = \frac{\sigma^2}{n} \longrightarrow 0$$

となるので，$\overline{X}$ は μ の一致推定量であると言える。

$\widehat{\boldsymbol{\alpha}}$，$\widehat{\boldsymbol{\beta}}$ の最良線型不偏性と一致性

不偏性： 既に証明したとおり，$\mathrm{E}(\widehat{\beta}) = \beta$，$\mathrm{E}(\widehat{\alpha}) = \alpha$ なので，$\widehat{\alpha}$，$\widehat{\beta}$ は β，α の不偏推定量である。

最良線型不偏性： $\widehat{\beta}$ を変形すると以下の通りとなる。

$$\begin{aligned}
\widehat{\beta} &= \frac{\sum_{i=1}^n (X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^n (X_i - \overline{X})^2} \\
&= \frac{\sum_{i=1}^n (X_i - \overline{X})Y_i}{\sum_{i=1}^n (X_i - \overline{X})^2} \\
&= \sum_{i=1}^n \omega_i Y_i
\end{aligned}$$

ただし，$\omega_i = \dfrac{(X_i - \overline{X})}{\sum_{i=1}^n (X_i - \overline{X})^2}$ とする。このように，$\widehat{\beta}$ は線型不偏推定量であると言える。

別の線型不偏推定量を次のように考える。

$$\widetilde{\beta} = \sum_{i=1}^n c_i Y_i$$

ただし，$c_i = \omega_i + d_i$ とする。$\widetilde{\beta}$ もまた β の不偏推定量と仮定したので，

$$\begin{aligned}
\widetilde{\beta} &= \sum_{i=1}^n c_i Y_i \\
&= \sum_{i=1}^n (\omega_i + d_i)(\alpha + \beta X_i + u_i) \\
&= \alpha \sum_{i=1}^n \omega_i + \beta \sum_{i=1}^n \omega_i X_i + \sum_{i=1}^n \omega_i u_i \\
&\quad + \alpha \sum_{i=1}^n d_i + \beta \sum_{i=1}^n d_i X_i + \sum_{i=1}^n d_i u_i \\
&= \beta + \alpha \sum_{i=1}^n d_i + \beta \sum_{i=1}^n d_i X_i \\
&\quad + \sum_{i=1}^n \omega_i u_i + \sum_{i=1}^n d_i u_i
\end{aligned}$$

と変形される。$\sum_{i=1}^n \omega_i = 0$，$\sum_{i=1}^n \omega_i X_i = 1$ に注意。

よって，期待値をとると，

$$\begin{aligned}
\mathrm{E}(\widetilde{\beta}) &= \beta + \alpha \sum_{i=1}^n d_i + \beta \sum_{i=1}^n d_i X_i \\
&\quad + \sum_{i=1}^n \omega_i \mathrm{E}(u_i) + \sum_{i=1}^n d_i \mathrm{E}(u_i) \\
&= \beta + \alpha \sum_{i=1}^n d_i + \beta \sum_{i=1}^n d_i X_i
\end{aligned}$$

となる。$\widetilde{\beta}$ が不偏であるためには，

$$\sum_{i=1}^n d_i = 0, \qquad \sum_{i=1}^n d_i X_i = 0,$$

の条件が必要となる。

この 2 つの条件が成り立っていると仮定すると，

$$\widetilde{\beta} = \beta + \sum_{i=1}^n (\omega_i + d_i) u_i$$

を利用して，

$$\begin{aligned}
\mathrm{V}(\widetilde{\beta}) &= \mathrm{E}(\widetilde{\beta} - \beta)^2 \\
&= \mathrm{E}\Big(\sum_{i=1}^n (\omega_i + d_i) u_i\Big)^2 \\
&= \sum_{i=1}^n (\omega_i + d_i)^2 \mathrm{E}(u_i^2) \\
&= \sigma^2 \left(\sum_{i=1}^n \omega_i^2 + 2 \sum_{i=1}^n \omega_i d_i + \sum_{i=1}^n d_i^2 \right)
\end{aligned}$$

$$= \sigma^2 \left(\sum_{i=1}^{n} \omega_i^2 + \sum_{i=1}^{n} d_i^2 \right)$$

$\widetilde{\beta}$ の不偏性の条件 $\sum_{i=1}^n d_i = 0$，$\sum_{i=1}^n d_i X_i = 0$ を利用すると，

$$\sum_{i=1}^{n} \omega_i d_i = \frac{\sum_{i=1}^n (X_i - \overline{X}) d_i}{\sum_{i=1}^n (X_i - \overline{X})^2}$$
$$= \frac{\sum_{i=1}^n X_i d_i - \overline{X} \sum_{i=1}^n d_i}{\sum_{i=1}^n (X_i - \overline{X})^2} = 0$$

を得る。
まとめると，$\widetilde{\beta}$ の分散は，

$$\mathrm{V}(\widetilde{\beta}) = \sigma^2 \left(\sum_{i=1}^{n} \omega_i^2 + \sum_{i=1}^{n} d_i^2 \right)$$

となる。$\widehat{\beta}$ の分散は，

$$\mathrm{V}(\widehat{\beta}) = \sigma^2 \sum_{i=1}^{n} \omega_i^2$$

なので，

$$\mathrm{V}(\widetilde{\beta}) \geq \mathrm{V}(\widehat{\beta})$$

となる。等号が成り立つときは，$\sum_{i=1}^n d_i^2 = 0$，すなわち，$d_1 = d_2 = \cdots = d_n = 0$ のときとなり，これは $\widehat{\beta}$ に一致する。
よって，$\widehat{\beta}$ は最小分散線型不偏推定量，または，最良線型不偏推定量であると言える。
$\Longrightarrow$ ガウス=マルコフの定理
$\widehat{\alpha}$ についても，同様で，α の最小分散線型不偏推定量となる。
証明は，

$$\widehat{\alpha} - \alpha = -(\widehat{\beta} - \beta)\overline{X} + \overline{u}$$

を利用すればよい。

推定量の関係 $\Longrightarrow$
最小分散 (最良) 線型不偏推定量 $\subset$ 線型不偏推定量 $\subset$ 線型推定量 $\subset$ 全推定量

一致性： $\mathrm{E}(\widehat{\beta}) = \beta$ となることが分かった。
n が大きくなると，$\widehat{\beta}$ は β に近づくかどうかを調べる。

$$\mathrm{V}(\widehat{\beta}) = \frac{\sigma^2}{\sum_{i=1}^n (X_i - \overline{X})^2}$$

$n \longrightarrow \infty$ のとき，$\mathrm{V}(\widehat{\beta}) \longrightarrow 0$ となれば，$\widehat{\beta}$ は β の一致推定量となる。
最小二乗法の仮定の一つに,「$n \longrightarrow \infty$ のとき，$\sum_{i=1}^n (X_i - \overline{X})^2 \longrightarrow \infty$」というものがあった。この仮定は,「$n \longrightarrow \infty$ のとき，$\mathrm{V}(\widehat{\beta}) \longrightarrow 0$」を保証する。よって，$\widehat{\beta}$ は β の一致推定量である。
$\widehat{\alpha}$ についても，同様に，$\mathrm{E}(\widehat{\alpha}) = \alpha$ であることは分かっている。

$$\mathrm{V}(\widehat{\alpha}) = \sigma^2 \left(\frac{1}{n} + \frac{\overline{X}^2}{\sum_{i=1}^n (X_i - \overline{X})^2} \right)$$

となり,「$n \longrightarrow \infty$ のとき，$\mathrm{V}(\widehat{\alpha}) \longrightarrow 0$」となるので，$\widehat{\alpha}$ も α の一致推定量であると言える。

4.4 誤差項 (または，攪乱項) u_i の分散 σ^2 について

$$Y_i = \alpha + \beta X_i + u_i,$$

誤差項 (または，攪乱項) の仮定：$u_i \sim N(0, \sigma^2)$

$$Y_i = \widehat{\alpha} + \widehat{\beta} X_i + \widehat{u}_i,$$

u_i の分散 σ^2 の不偏推定量：

$$\frac{\sum_{i=1}^n \widehat{u}_i^2}{\text{自由度}}$$

$$\begin{aligned} \text{自由度} &= \text{標本数}\ (n) - \text{推定すべき係数値の数}\ (2) \\ &= n - 2 \end{aligned}$$

誤差項 (または，攪乱項) の母分散 σ^2 の不偏推定量 s^2 は，

$$s^2 = \frac{1}{n-2} \sum_{i=1}^{n} \widehat{u}_i^2, = \frac{1}{n-2} \sum_{i=1}^{n} (Y_i - \widehat{\alpha} - \widehat{\beta} X_i)^2,$$

によって与えられる。

s^2 の不偏性の証明： まず，次のように書き直す。

$$\begin{aligned} u_i &= Y_i - \alpha - \beta X_i \\ &= (\widehat{\alpha} + \widehat{\beta} X_i + \widehat{u}_i) - \alpha - \beta X_i \\ &= (\widehat{\alpha} - \alpha) + (\widehat{\beta} - \beta) X_i + \widehat{u}_i, \end{aligned}$$

両辺を二乗する。

$$\begin{aligned}u_i^2 &= (\widehat{\alpha}-\alpha)^2 + (\widehat{\beta}-\beta)^2 X_i^2 + \widehat{u}_i^2 \\ &+ 2(\widehat{\alpha}-\alpha)(\widehat{\beta}-\beta)X_i \\ &+ 2(\widehat{\alpha}-\alpha)\widehat{u}_i \\ &+ 2(\widehat{\beta}-\beta)X_i\widehat{u}_i\end{aligned}$$

総和をとる。

$$\begin{aligned}\sum_{i=1}^n u_i^2 &= n(\widehat{\alpha}-\alpha)^2 + (\widehat{\beta}-\beta)^2\sum_{i=1}^n X_i^2 + \sum_{i=1}^n \widehat{u}_i^2 \\ &\quad + 2(\widehat{\alpha}-\alpha)(\widehat{\beta}-\beta)\sum_{i=1}^n X_i \\ &\quad + 2(\widehat{\alpha}-\alpha)\sum_{i=1}^n \widehat{u}_i \\ &\quad + 2(\widehat{\beta}-\beta)\sum_{i=1}^n X_i\widehat{u}_i \\ &\quad = n(\widehat{\alpha}-\alpha)^2 + (\widehat{\beta}-\beta)^2\sum_{i=1}^n X_i^2 + \sum_{i=1}^n \widehat{u}_i^2 \\ &\quad + 2n(\widehat{\alpha}-\alpha)(\widehat{\beta}-\beta)\overline{X}\end{aligned}$$

期待値をとる。

$$\begin{aligned}\mathrm{E}(\sum_{i=1}^n u_i^2) &= n\mathrm{E}(\widehat{\alpha}-\alpha)^2 + \mathrm{E}(\widehat{\beta}-\beta)^2\sum_{i=1}^n X_i^2 \\ &\quad + \mathrm{E}(\sum_{i=1}^n \widehat{u}_i^2) + 2n\mathrm{E}\big((\widehat{\alpha}-\alpha)(\widehat{\beta}-\beta)\big)\overline{X}\end{aligned}$$

$$\begin{aligned}n\sigma^2 &= \frac{\sigma^2\sum_{i=1}^n X_i^2}{\sum_{i=1}^n (X_i-\overline{X})^2} + \frac{\sigma^2\sum_{i=1}^n X_i^2}{\sum_{i=1}^n (X_i-\overline{X})^2} \\ &+ \mathrm{E}(\sum_{i=1}^n \widehat{u}_i^2) - \frac{2n\sigma^2\overline{X}^2}{\sum_{i=1}^n (X_i-\overline{X})^2} \\ &= 2\sigma^2\left(\frac{\sum_{i=1}^n X_i^2 - n\overline{X}^2}{\sum_{i=1}^n (X_i-\overline{X})^2}\right) + \mathrm{E}(\sum_{i=1}^n \widehat{u}_i^2) \\ &= 2\sigma^2 + \mathrm{E}(\sum_{i=1}^n \widehat{u}_i^2)\end{aligned}$$

途中の計算には以下が使われる。

$$\begin{aligned}\mathrm{E}(\sum_{i=1}^n u_i^2) &= n\sigma^2 \\ \mathrm{E}(\widehat{\alpha}-\alpha)^2 &= \frac{\sigma^2\sum_{i=1}^n X_i^2}{n\sum_{i=1}^n (X_i-\overline{X})^2} \\ \mathrm{E}(\widehat{\beta}-\beta)^2 &= \frac{\sigma^2}{\sum_{i=1}^n (X_i-\overline{X})^2} \\ \mathrm{E}\big((\widehat{\alpha}-\alpha)(\widehat{\beta}-\beta)\big) &= -\frac{\sigma^2\overline{X}}{\sum_{i=1}^n (X_i-\overline{X})^2}\end{aligned}$$

よって，

$$\mathrm{E}(s^2) = \mathrm{E}\left(\frac{\sum_{i=1}^n \widehat{u}_i^2}{n-2}\right) = \sigma^2$$

を得る。すなわち，s^2 は σ^2 の不偏推定量である。

統計学の復習 (χ^2 分布)： m 個の確率変数 $Z_1, Z_2, \cdots, Z_m$ は，互いに独立な標準正規分布に従うものとする。このとき，$Y = \sum_{i=1}^m Z_i^2$ は，自由度 m の χ^2 分布に従う。
$Y \sim \chi^2(m)$，または，$Y \sim \chi_m^2$ と表記する。
χ^2 (カイ二乗) 分布表から確率を求める。
$Y \sim \chi^2(m)$ のとき，$\mathrm{E}(Y) = m$，$\mathrm{V}(Y) = 2m$ となる。(証明略)

1. 2 つの独立な χ^2 分布からの確率変数 X, Y を考える。$X \sim \chi^2(n)$，$Y \sim \chi^2(m)$ とする。このとき，$Z = X + Y \sim \chi^2(n+m)$ となる。(証明略)
2. n 個の独立な確率変数 $X_1, X_2, \cdots, X_n$ が同一の正規分布 $N(\mu, \sigma^2)$ に従うものとする。
3. $\dfrac{X_i-\mu}{\sigma} \sim N(0,1)$ なので，$\left(\dfrac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2(1)$ となる。
   $\dfrac{X_1-\mu}{\sigma}, \dfrac{X_2-\mu}{\sigma}, \cdots, \dfrac{X_n-\mu}{\sigma}$ はそれぞれ独立なので，
   $$\sum_{i=1}^n \left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2(n)$$
   となる。
4. μ を $\overline{X}$ に置き換えると，
   $$\sum_{i=1}^n \left(\frac{X_i-\overline{X}}{\sigma}\right)^2 \sim \chi^2(n-1)$$
   となる。(証明は後述)
   さらに，
   $$S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i-\overline{X})^2$$

を定義すると，

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

となる。S^2 は σ^2 の不偏推定量である (後述)。

5. すなわち，

$$\mathrm{E}\left(\frac{(n-1)S^2}{\sigma^2}\right) = n-1,$$
$$\mathrm{V}\left(\frac{(n-1)S^2}{\sigma^2}\right) = 2(n-1),$$

となる。

回帰分析に当てはめる。

$$\frac{\sum_{i=1}^n u_i^2}{\sigma^2} = \frac{\sum_{i=1}^n (Y_i-\alpha-\beta X_i)^2}{\sigma^2} \sim \chi^2(n),$$

α, β を推定値に置き換えると，

$$\frac{\sum_{i=1}^n \widehat{u}_i^2}{\sigma^2} = \frac{\sum_{i=1}^n (Y_i-\widehat{\alpha}-\widehat{\beta} X_i)^2}{\sigma^2} \sim \chi^2(n-2),$$

となる。さらに，

$$s^2 = \frac{1}{n-2}\sum_{i=1}^n (Y_i-\widehat{\alpha}-\widehat{\beta} X_i)^2,$$

なので，

$$\frac{\sum_{i=1}^n \widehat{u}_i^2}{\sigma^2} = \frac{(n-2)s^2}{\sigma^2} \sim \chi^2(n-2),$$

を得る。

s^2 の一致性の証明： s^2 は

$$s^2 = \frac{1}{n-2}\sum_{i=1}^n \widehat{u}_i^2 = \frac{1}{n-2}\sum_{i=1}^n (Y_i-\widehat{\alpha}-\widehat{\beta} X_i)^2$$

と定義される。
$\dfrac{(n-2)s^2}{\sigma^2} \sim \chi^2(n-2)$ なので (証明略)，

$$\mathrm{E}\left(\frac{(n-2)s^2}{\sigma^2}\right) = n-2,$$
$$\mathrm{V}\left(\frac{(n-2)s^2}{\sigma^2}\right) = 2(n-2),$$

となる。さらに，書き直すと，

$$\frac{(n-2)^2}{\sigma^4}\mathrm{V}(s^2) = 2(n-2),$$
$$\mathrm{V}(s^2) = \frac{2\sigma^4}{n-2},$$

を得る。「$\mathrm{E}(s^2) = \sigma^2$ で，しかも，$n \longrightarrow \infty$ のとき $\mathrm{V}(s^2) \longrightarrow 0$」が言えるので，$s^2$ は σ^2 の一致推定量である。

標準誤差について： 標準誤差 = 不偏分散の平方根
誤差項 (または，攪乱項) の標準誤差 s

$$s = \sqrt{\frac{\sum_{i=1}^n \widehat{u}_i^2}{n-2}}$$

数値例： $\widehat{\alpha} = 0.3$, $\widehat{\beta} = 0.65$ なので，$\widehat{Y}_i = 0.3 + 0.65X_i$, $\widehat{u}_i = Y_i - 0.3 - 0.65X_i$ により，$\widehat{Y}_i$, $\widehat{u}_i$ を計算する。

i	Y_i	X_i	X_iY_i	X_i^2	$\widehat{Y}_i$	$\widehat{u}_i$
1	6	10	60	100	6.8	−0.8
2	9	12	108	144	8.1	0.9
3	10	14	140	196	9.4	0.6
4	10	16	160	256	10.7	−0.7
合計	$\sum Y_i$	$\sum X_i$	$\sum X_iY_i$	$\sum X_i^2$	$\sum \widehat{Y}_i$	$\sum \widehat{u}_i$
	35	52	468	696	35	0
平均	$\overline{Y}$	$\overline{X}$				
	8.75	13				

誤差項 (または，攪乱項) の母分散 σ^2 の不偏推定量 s^2 は，

$$\begin{aligned} s^2 &= \frac{1}{n-2}\sum_{i=1}^n \widehat{u}_i^2 \\ &= \frac{1}{2}\big((-0.8)^2 + 0.9^2 + 0.6^2 + (-0.7)^2\big) \\ &= 1.15 \end{aligned}$$

によって与えられる。
s は「回帰の標準誤差 (Standard Error of Regression)」と呼ばれ，この例では，$s = \sqrt{1.15} = 1.07$ となる。

4.4.1 $\widehat{\alpha}$，$\widehat{\beta}$ の分散の不偏推定量

$\widehat{\alpha}$ と $\widehat{\beta}$ の分散は，

$$\mathrm{V}(\widehat{\alpha}) = \sigma_{\widehat{\alpha}}^2 = \frac{\sigma^2\sum_{i=1}^n X_i^2}{n\sum_{i=1}^n (X_i-\overline{X})^2}$$
$$\mathrm{V}(\widehat{\beta}) = \sigma_{\widehat{\beta}}^2 = \frac{\sigma^2}{\sum_{i=1}^n (X_i-\overline{X})^2}$$

によって，与えられる。

σ^2 をその不偏分散 s^2 に置き換えることによって，$\widehat{\alpha}$，$\widehat{\beta}$ の分散の不偏推定量を次のように得ることができる。

$$s^2_{\widehat{\alpha}} = \frac{s^2 \sum_{i=1}^n X_i^2}{n\sum_{i=1}^n (X_i - \overline{X})^2}$$

$$s^2_{\widehat{\beta}} = \frac{s^2}{\sum_{i=1}^n (X_i - \overline{X})^2}$$

さらに，平方根をとって，$\widehat{\alpha}$，$\widehat{\beta}$ の標準誤差はそれぞれ，

$$s_{\widehat{\alpha}} = s\sqrt{\frac{\sum_{i=1}^n X_i^2}{n\sum_{i=1}^n (X_i - \overline{X})^2}}$$

$$s_{\widehat{\beta}} = \frac{s}{\sqrt{\sum_{i=1}^n (X_i - \overline{X})^2}}$$

として与えられる。

数値例： $\widehat{\alpha} = 0.3$, $\widehat{\beta} = 0.65$ なので，$\widehat{Y}_i = 0.3 + 0.65X_i$, $\widehat{u}_i = Y_i - 0.3 - 0.65X_i$ により，$\widehat{Y}_i$, $\widehat{u}_i$ を計算する。

i	Y_i	X_i	X_iY_i	X_i^2	$\widehat{Y}_i$	$\widehat{u}_i$
1	6	10	60	100	6.8	−0.8
2	9	12	108	144	8.1	0.9
3	10	14	140	196	9.4	0.6
4	10	16	160	256	10.7	−0.7
合計	$\sum Y_i$	$\sum X_i$	$\sum X_iY_i$	$\sum X_i^2$	$\sum \widehat{Y}_i$	$\sum \widehat{u}_i$
	35	52	468	696	35	0
平均	$\overline{Y}$	$\overline{X}$				
	8.75	13				

$s^2 = 1.15$ なので，

$$\begin{aligned} s^2_{\widehat{\beta}} &= \frac{s^2}{\sum_{i=1}^n (X_i - \overline{X})^2} \\ &= \frac{s^2}{\sum_{i=1}^n X_i^2 - n\overline{X}^2} \\ &= 0.05 \times 1.15 \\ &= 0.0575 \end{aligned}$$

$$\begin{aligned} s^2_{\widehat{\alpha}} &= \frac{s^2 \sum_{i=1}^n X_i^2}{n\sum_{i=1}^n (X_i - \overline{X})^2} \\ &= \frac{s^2 \sum_{i=1}^n X_i^2}{n(\sum_{i=1}^n X_i^2 - n\overline{X}^2)} \\ &= 8.7 \times 1.15 \\ &= 10.005 \end{aligned}$$

$\widehat{\alpha}$，$\widehat{\beta}$ の標準誤差はそれぞれ，平方根をとって，

$$s_{\widehat{\beta}} = \sqrt{0.0575} = 0.240$$

$$s_{\widehat{\alpha}} = \sqrt{10.005} = 3.163$$

となる。

4.5 $\widehat{\alpha}$，$\widehat{\beta}$ の分布

4.5.1 統計学の復習 (t 分布)

正規分布の重要な定理： n 個の独立な確率変数 $X_1, X_2, \cdots, X_n$ が同一の正規分布 $N(\mu, \sigma^2)$ に従うものとする。このとき，

$$\sum_{i=1}^n c_i X_i \sim N(\mu \sum_{i=1}^n c_i, \sigma^2 \sum_{i=1}^n c_i^2)$$

となる。ただし，$c_1, c_2, \cdots, c_n$ は定数とする。

***t* 分布：** Z を標準正規分布，Y を自由度 m の χ^2 分布に従い，両者は独立な確率変数とする。このとき，$U = \dfrac{Z}{\sqrt{Y/m}}$ は，自由度 m の t 分布に従う。

$U \sim t(m)$，または，$U \sim t_m$ と表記する。

$U \sim t(m)$ のとき，$m > 1$ について $\mathrm{E}(U) = 0$，$m > 2$ について $\mathrm{V}(U) = \dfrac{m}{m-2}$ となる。(証明略)

t 分布表から確率を求める。(表 **??** を見よ)

1. ゼロを中心に左右対称。($\mathrm{E}(U) = 0$)
2. t 分布は，標準正規分布より裾野の広い分布 (なぜなら，$\mathrm{V}(U) = \dfrac{m}{m-2} > 1$)
3. $m \longrightarrow \infty$ のとき，$t(m) \longrightarrow N(0,1)$ となる。(期待値は $m > 1$ について $\mathrm{E}(U) = 0$，分散は $\mathrm{V}(U) = \dfrac{m}{m-2} \longrightarrow 1$)

標本平均 $\overline{X}$ の分布： $X_1, X_2, \cdots, X_n$ の n 個の確率変数は，互いに独立で，平均 μ，分散 σ^2 の正規分布に従うものとする。

1. $\overline{X} \sim N(\mu, \dfrac{\sigma^2}{n})$ なので，$\dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$ となる。
2. $\dfrac{(n-1)S^2}{\sigma^2} = \dfrac{\sum_{i=1}^n (X_i - \overline{X})^2}{\sigma^2} \sim \chi^2(n-1)$ である。(証明は略)
3. $\dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}}$ と $\dfrac{(n-1)S^2}{\sigma^2}$ は独立。(証明は略)
   すなわち，$\overline{X}$ と S^2 は独立。

4. したがって，

$$\frac{\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\dfrac{(n-1)S^2}{\sigma^2}\Big/ n-1}} = \frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$$

を得る。

重要な結果は，

$$\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$$

ただし，$\overline{X}=\dfrac{1}{n}\displaystyle\sum_{i=1}^n X_i$，$S^2=\dfrac{1}{n-1}\displaystyle\sum_{i=1}^n (X_i-\overline{X})^2$ である。

σ^2 を S^2 に置き換えると，正規分布から t 分布になる。

$$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1) \quad\Longrightarrow\quad \frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$$

4.5.2 $\widehat{\beta}$ について：

$$\begin{aligned}\widehat{\beta} &= \beta + \frac{\sum_{i=1}^n (X_i-\overline{X})u_i}{\sum_{i=1}^n (X_i-\overline{X})^2} \\ &= \beta + \sum_{i=1}^n \omega_i u_i\end{aligned}$$

$u_i \sim N(0,\sigma^2)$ で，かつ，それぞれ独立に分布する。また，$\widehat{\beta}$ の平均，分散はそれぞれ，

$$\begin{aligned}&\mathrm{E}(\widehat{\beta}) = \beta, \\ &\mathrm{V}(\widehat{\beta}) = \frac{\sigma^2}{\sum_{i=1}^n (X_i-\overline{X})^2},\end{aligned}$$

となるので，

$$\widehat{\beta} \sim N(\beta, \frac{\sigma^2}{\sum_{i=1}^n (X_i-\overline{X})^2}),$$

を得る。変形すると，

$$\frac{\widehat{\beta}-\beta}{\sigma/\sqrt{\sum_{i=1}^n (X_i-\overline{X})^2}} \sim N(0,1),$$

となる。

さらに，

$$\frac{(n-2)s^2}{\sigma^2} \sim \chi^2(n-2),$$

となり (証明略)，$\widehat{\beta}$ とは独立なので (証明略)，

$$\begin{aligned}&\frac{\dfrac{\widehat{\beta}-\beta}{\sigma/\sqrt{\sum_{i=1}^n (X_i-\overline{X})^2}}}{\sqrt{\dfrac{(n-2)s^2}{\sigma^2}/(n-2)}} \\ &= \frac{\widehat{\beta}-\beta}{s/\sqrt{\sum_{i=1}^n (X_i-\overline{X})^2}} \sim t(n-2)\end{aligned}$$

4.5.3 $\widehat{\alpha}$ について：

また，$\widehat{\alpha}$ の平均，分散はそれぞれ，

$$\begin{aligned}&\mathrm{E}(\widehat{\alpha}) = \alpha, \\ &\mathrm{V}(\widehat{\alpha}) = \frac{\sigma^2 \sum_{i=1}^n X_i^2}{n\sum_{i=1}^n (X_i-\overline{X})^2},\end{aligned}$$

となるので，

$$\widehat{\alpha} \sim N(\alpha, \frac{\sigma^2 \sum_{i=1}^n X_i^2}{n\sum_{i=1}^n (X_i-\overline{X})^2}),$$

を得る。変形すると，

$$\frac{\widehat{\alpha}-\alpha}{\sigma\sqrt{\sum_{i=1}^n X_i^2/n\sum_{i=1}^n (X_i-\overline{X})^2}} \sim N(0,1),$$

となる。

さらに，σ を s で置き換えると，

$$\frac{\widehat{\alpha}-\alpha}{s\sqrt{\sum_{i=1}^n X_i^2/n\sum_{i=1}^n (X_i-\overline{X})^2}} \sim t(n-2),$$

となる。

4.5.4 まとめ：

$$\frac{\widehat{\beta}-\beta}{s_{\widehat{\beta}}} = \frac{\widehat{\beta}-\beta}{s/\sqrt{\sum_{i=1}^n (X_i-\overline{X})^2}} \sim t(n-2),$$

$$\frac{\widehat{\alpha}-\alpha}{s_{\widehat{\alpha}}} = \frac{\widehat{\alpha}-\alpha}{s\sqrt{\dfrac{\sum_{i=1}^n X_i^2}{n\sum_{i=1}^n (X_i-\overline{X})^2}}} \sim t(n-2),$$

4.6 α，β の区間推定 (信頼区間)

4.6.1 統計学の復習： 区間推定 (信頼区間)

$\overline{X}$ の分布を利用して，μ の信頼区間を求める。

1. $\overline{X}$ の分布は以下の通り。

$$\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$$

となる。

2. $t_{\alpha/2}(n-1)$，$t_{1-\alpha/2}(n-1)$ を自由度 $n-1$ の t 分布の上から $100\times\dfrac{\alpha}{2}$ % 点，$100\times(1-\dfrac{\alpha}{2})$ % 点の値とする。このとき，

$$\mathrm{Prob}\Big(t_{1-\alpha/2}(n-1) < \frac{\overline{X}-\mu}{S/\sqrt{n}} < t_{\alpha/2}(n-1)\Big) = 1-\alpha$$

となる。ただし，自由度と α が決まれば，$t_{\alpha/2}(n-1)$，$t_{1-\alpha/2}(n-1)$ は t 分布表から得られる。

3. t 分布は左右対称なので，

$$t_{1-\alpha/2}(n-1) = -t_{\alpha/2}(n-1)$$
$$t_{\alpha/2}(n-1) = |t_{1-\alpha/2}(n-1)|$$
$$t_{1-\alpha/2}(n-1) = -|t_{\alpha/2}(n-1)|$$

となる。

4. 書き直して，

$$\mathrm{Prob}\Big(\ \overline{X}-t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}} < \mu < \overline{X}+t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}}\ \Big) = 1-\alpha$$

となる。

5. μ が区間 $(\overline{X}-t_{\alpha/2}(n-1)\dfrac{S}{\sqrt{n}}, \overline{X}+t_{\alpha/2}(n-1)\dfrac{S}{\sqrt{n}})$ にある確率は $1-\alpha$ である。

6. 推定量 $\overline{X}$，S^2 をその推定値 $\overline{x}$，s^2 で置き換える。ただし，$\overline{x} = \dfrac{1}{n}\sum_{i=1}^{n} x_i$，$s^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})^2$ とする。

7. 区間 $(\overline{x}-t_{\alpha/2}(n-1)\dfrac{s}{\sqrt{n}}, \overline{x}+t_{\alpha/2}(n-1)\dfrac{s}{\sqrt{n}})$ を信頼係数 $1-\alpha$ の信頼区間といい，$\overline{x}-t_{\alpha/2}(n-1)\dfrac{s}{\sqrt{n}}$ を信頼下限，$\overline{x}+t_{\alpha/2}(n-1)\dfrac{s}{\sqrt{n}}$ を信頼上限と呼ぶ。

4.6.2 α，β の区間推定 (信頼区間)

$\widehat{\alpha}$，$\widehat{\beta}$ の分布は，以下のように得られた。

$$\frac{\widehat{\beta}-\beta}{s_{\widehat{\beta}}} \sim t(n-2),$$

$$\frac{\widehat{\alpha}-\alpha}{s_{\widehat{\alpha}}} \sim t(n-2),$$

$t_{\alpha/2}(n-2)$，$t_{1-\alpha/2}(n-2)$ をそれぞれ自由度 $n-2$ の t 分布の上側から $100\times\dfrac{\alpha}{2}$ % 点， $100\times(1-\dfrac{\alpha}{2})$ % 点の値とする。このとき，

$$\mathrm{Prob}\Big(t_{1-\alpha/2}(n-2) < \frac{\widehat{\beta}-\beta}{s_{\widehat{\beta}}} < t_{\alpha/2}(n-2)\Big) = 1-\alpha,$$

すなわち，$t_{1-\alpha/2}(n-2) = -t_{\alpha/2}(n-2)$ により，

$$\mathrm{Prob}\Big(-t_{\alpha/2}(n-2) < \frac{\widehat{\beta}-\beta}{s_{\widehat{\beta}}} < t_{\alpha/2}(n-2)\Big) = 1-\alpha,$$

となる。ただし，自由度と α が決まれば，$t_{\alpha/2}(n-2)$ は t 分布表から得られる。
書き直して，

$$\mathrm{Prob}\Big(\ \widehat{\beta}-t_{\alpha/2}(n-2)s_{\widehat{\beta}} < \beta < \widehat{\beta}+t_{\alpha/2}(n-2)s_{\widehat{\beta}}\ \Big) = 1-\alpha,$$

と表される。
したがって，$\widehat{\beta}$，$s_{\widehat{\beta}}$ を推定値で置き換えて，信頼係数 $1-\alpha$ の β の信頼区間は，

$$\Big(\widehat{\beta}-t_{\alpha/2}(n-2)s_{\widehat{\beta}},\ \widehat{\beta}+t_{\alpha/2}(n-2)s_{\widehat{\beta}}\Big)$$

となる。
同様に，信頼係数 $1-\alpha$ の α の信頼区間は，

$$\Big(\widehat{\alpha}-t_{\alpha/2}(n-2)s_{\widehat{\alpha}},\ \widehat{\alpha}+t_{\alpha/2}(n-2)s_{\widehat{\alpha}}\Big)$$

となる。

数値例： 今までと同様に，以下の数値例をとりあげる。

i	Y_i	X_i
1	6	10
2	9	12
3	10	14
4	10	16

回帰モデル $Y_i = \alpha + \beta X_i + u_i$ を推定した結果，以下の推定値を得た。

$$\widehat{\beta} = 0.65,$$
$$\widehat{\alpha} = 0.3,$$
$$s_{\widehat{\beta}} = \sqrt{0.0575} = 0.240,$$
$$s_{\widehat{\alpha}} = \sqrt{10.005} = 3.163,$$

$t_{0.025}(2) = 4.303$ なので，信頼係数 0.95 の β の信頼区間は，

$$\big(0.65 - 4.303 \times 0.240,\ 0.65 + 4.303 \times 0.240,\big)$$

となり (すなわち，$(-0.383,\ 1.683)$)，信頼係数 0.95 の α の信頼区間は，

$$\big(0.3 - 4.303 \times 3.163,\ 0.3 + 4.303 \times 3.163,\big)$$

となる (すなわち，$(-13.31,\ 13.91)$)。
同様にして，信頼係数 0.90 の β の信頼区間は，

$$\big(0.65 - 2.920 \times 0.240,\ 0.65 + 2.920 \times 0.240,\big)$$

となり (すなわち，$(-0.051,\ 1.051)$)，信頼係数 0.95 の α の信頼区間は，

$$\big(0.3 - 2.920 \times 3.163,\ 0.3 + 2.920 \times 3.163,\big)$$

となる (すなわち，$(-8.94,\ 9.24)$)。

4.7 α，β の仮説検定

4.7.1 統計学の復習： 仮説検定

$\overline{X}$ の分布を利用して，μ の仮説検定を行う。

1. 帰無仮説 H_0： $\mu = \mu_0$

   対立仮説 H_1： $\mu \neq \mu_0$

2. 帰無仮説 H_0： $\mu = \mu_0$ が正しいもとでの分布は，

$$\frac{\overline{X} - \mu_0}{S/\sqrt{n}} \sim t(n-1)$$

   となる。

3. $\mathrm{Prob}\Big(t_{1-\alpha/2}(n-1) < \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}} < t_{\alpha/2}(n-1)\Big) = 1-\alpha$

   $t_{\alpha/2}(n-1)$，$t_{1-\alpha/2}(n-1)$ をそれぞれ自由度 $n-1$ の t 分布の上から $100 \times \dfrac{\alpha}{2}$ % 点，$100 \times \dfrac{1-\alpha}{2}$ % 点の値とする。

   自由度と α が決まれば，$t_{\alpha/2}(n-1)$，$t_{1-\alpha/2}(n-1)$ は t 分布表から得られる。

4. α を有意水準と呼ぶ。慣習的に $\alpha = 0.01, 0.05$ が使われる。

5. $-t_{\alpha/2}(n-1) > \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}}$，または，

   $\dfrac{\overline{X} - \mu_0}{S/\sqrt{n}} > t_{\alpha/2}(n-1)$ ならば，帰無仮説 H_0： $\mu = \mu_0$ は，分布の端にあり，起こりにくいと考える。

   $\Longrightarrow$ 有意水準 α で帰無仮説 H_0： $\mu = \mu_0$ を棄却する。

6. 実際の検定手続：

   (a) $\overline{X}$, S^2 を実績値で置き換えて，

$$\frac{\overline{x} - \mu_0}{s/\sqrt{n}}$$

   を得る。

   ただし，$\overline{x} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} x_i$，$s^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}(x_i - \overline{x})^2$ とする。

   (b) $-t_{\alpha/2}(n-1) > \dfrac{\overline{x} - \mu_0}{s/\sqrt{n}}$，または，

   $\dfrac{\overline{x} - \mu_0}{s/\sqrt{n}} > t_{\alpha/2}(n-1)$ ならば，有意水準 α で帰無仮説 H_0： $\mu = \mu_0$ を棄却する。

4.7.2 α，β の仮説検定

「帰無仮説： $H_0 : \beta = \beta_0$，対立仮説： $H_1 : \beta \neq \beta_0$」の検定を行う。
帰無仮説： $H_0 : \beta = \beta_0$ が正しいとするもとでは，

$$\frac{\widehat{\beta} - \beta_0}{s_{\widehat{\beta}}} = \frac{\widehat{\beta} - \beta_0}{s/\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2}} \sim t(n-2),$$

の分布に従う。
よって，検定の手順は，

1. 検定統計値

$$\frac{\widehat{\beta}-\beta_0}{s/\sqrt{\sum_{i=1}^n (X_i-\overline{X})^2}}$$

を計算する。

ただし，

$$s^2=\frac{1}{n-2}\sum_{i=1}^n (Y_i-\widehat{\alpha}-\widehat{\beta}X_i)^2,$$
$$\widehat{\beta}=\frac{\sum_{i=1}^n (X_i-\overline{X})(Y_i-\overline{Y})}{\sum_{i=1}^n (X_i-\overline{X})^2},$$
$$\widehat{\alpha}=\overline{Y}-\widehat{\beta}\overline{X},$$

とする。

2. 有意水準 α を決めて (この α は，回帰式の定数項の α とは異なることに注意)，t 分布表から $100\times\alpha$ % 点の値を求める。通常，$\alpha=0.01, 0.05$ とする。

3. t 分布表から得られた $100\times\alpha$ % 点の値と検定統計値の大小を比較する。すなわち，

$$-t_{\alpha/2}(n-2)>\frac{\widehat{\beta}-\beta_0}{s/\sqrt{\sum_{i=1}^n (X_i-\overline{X})^2}},$$

または，

$$\frac{\widehat{\beta}-\beta_0}{s/\sqrt{\sum_{i=1}^n (X_i-\overline{X})^2}}>t_{\alpha/2}(n-2)$$

ならば，有意水準 α で帰無仮説 H_0： $\beta=\beta_0$ を棄却する。

帰無仮説 H_0： $\beta=\beta_0$ を中心とした分布を考えるとき，データから得られた検定統計量は，分布の端にあり，確率的に起こりにくいと考える。

となる。
定数項の推定量 $\widehat{\alpha}$ についても同様。

4.7.3 t 値について

特に,「帰無仮説： $H_0:\beta=0$，対立仮説： $H_1:\beta\neq 0$」の検定は特別の意味を持つ。

帰無仮説のもとでは，

$$\frac{\widehat{\beta}}{s/\sqrt{\sum_{i=1}^n (X_i-\overline{X})^2}}\sim t(n-2)$$

の分布に従う。
このときの t 統計量の値 (すなわち，$\dfrac{\widehat{\beta}}{s/\sqrt{\sum_{i=1}^n (X_i-\overline{X})^2}}$) を t 値と呼ぶ。
回帰式

$$Y_i=\alpha+\beta X_i+u_i,$$

において，帰無仮説 $H_0:\beta=0$ は,「X_i が Y_i に何の影響も与えない」ということを意味する。
有意水準 α のもとで (この α は，回帰式の定数項の α とは異なることに注意)，

$$-t_{\alpha/2}(n-2)>\frac{\widehat{\beta}}{s/\sqrt{\sum_{i=1}^n (X_i-\overline{X})^2}},$$

または，

$$\frac{\widehat{\beta}}{s/\sqrt{\sum_{i=1}^n (X_i-\overline{X})^2}}>t_{\alpha/2}(n-2)$$

ならば，有意水準 α で帰無仮説 H_0： $\beta=0$ を棄却する。

1. 回帰モデル

$$Y_i=\alpha+\beta X_i+u_i$$

において，経済理論から $\beta>0$ という符号条件が想定されたとする。

2. 有意水準 α を決める。(例えば，$\alpha=0.05, 0.01$)

3. 実際のデータから，$\widehat{\beta}>0$ が得られた場合

   (a) t 値が，

$$\frac{\widehat{\beta}}{s/\sqrt{\sum_{i=1}^n (X_i-\overline{X})^2}}>t_{\alpha/2}(n-2)$$

   となった場合，帰無仮説 H_0： $\beta=0$ が棄却され，$\beta>0$ が統計的にも証明され，経済理論は現実経済をサポートする結果。

(b) t 値が，

$$\frac{\widehat{\beta}}{s/\sqrt{\sum_{i=1}^n (X_i - \overline{X})^2}} < t_{\alpha/2}(n-2)$$

となった場合，帰無仮説 H_0： $\beta = 0$ を棄却できず，$\widehat{\beta} > 0$ にもかかわらず，$\beta < 0$ の可能性もあるため，経済理論が現実経済を積極的には支持しない。

4. 実際のデータから，$\widehat{\beta} < 0$ が得られた場合

(a) t 値が，

$$-t_{\alpha/2}(n-2) < \frac{\widehat{\beta}}{s/\sqrt{\sum_{i=1}^n (X_i - \overline{X})^2}},$$

となった場合，帰無仮説 H_0： $\beta = 0$ を棄却できず，$\widehat{\beta} < 0$ にもかかわらず，$\beta > 0$ の可能性もあるため，得られた推定結果と経済理論は完全に矛盾しているとは言えない。

(b) t 値が，

$$-t_{\alpha/2}(n-2) > \frac{\widehat{\beta}}{s/\sqrt{\sum_{i=1}^n (X_i - \overline{X})^2}},$$

となった場合，帰無仮説 H_0： $\beta = 0$ は棄却され，統計的には $\beta < 0$ となり，得られた推定結果と経済理論は完全に矛盾している。すなわち，この場合，経済理論の立て直しが必要。

数値例：　今までと同様に，以下の数値例をとりあげる。

i	Y_i	X_i
1	6	10
2	9	12
3	10	14
4	10	16

回帰モデル $Y_i = \alpha + \beta X_i + u_i$ を推定した結果，以下の推定値を得た。

$$\widehat{\beta} = 0.65,$$

$$s_{\widehat{\beta}} = \sqrt{0.0575} = 0.240,$$

帰無仮説 H_0： $\beta = 0$，対立仮説 H_0： $\beta \neq 0$ の検定を行う。t 値は $0.65/0.240 = 2.711$，有意水準 5 % の $t_{\alpha/2}(n-2)$ の値は 4.303 となり ($\alpha = 0.5$, $n = 4$)，

$$\frac{\widehat{\beta}}{s_{\widehat{\beta}}} = \frac{0.65}{\sqrt{0.0575}} = 2.708 < t_{\alpha/2}(n-2) = 4.303$$

を得る。このように，有意水準 5 % で帰無仮説 H_0： $\beta = 0$ は棄却されない。よって，β の符号は統計学的に確定できない。

また，α についても同様に，t 値を計算できる。

$$\widehat{\alpha} = 0.3,$$

$$s_{\widehat{\alpha}} = \sqrt{10.005} = 3.163,$$

なので，t 値は，

$$\frac{\widehat{\alpha}}{s_{\widehat{\alpha}}} = \frac{0.3}{\sqrt{10.005}} = 0.095 < t_{\alpha/2}(n-2) = 4.303,$$

となり，有意水準 5 % で H_0： $\alpha = 0$ を棄却できない。しかし，定数項については，経済学的意味が無い場合が多い。

推定結果の表記方法：　回帰モデル：

$$Y_i = \alpha + \beta X_i + u_i,$$

の推定の結果，$\widehat{\alpha} = 0.3$, $\widehat{\beta} = 0.65$, $s_{\widehat{\alpha}} = \sqrt{10.0005} = 3.163$, $s_{\widehat{\beta}} = \sqrt{0.0575} = 0.240$, $\dfrac{\widehat{\alpha}}{s_{\widehat{\alpha}}} = 0.095$, $\dfrac{\widehat{\beta}}{s_{\widehat{\beta}}} = 2.708$, $s^2 = 1.15$ (すなわち，$s = 1.07$), $R^2 = 0.786$ を得た。これらをまとめて，

$$Y_i = \underset{(0.095)}{0.3} + \underset{(2.708)}{0.65} X_i,$$

$$R^2 = 0.786, \quad s = 1.07,$$

ただし，係数の推定値の下の括弧内は t 値を表すものとする。

または，

$$Y_i = \underset{(3.163)}{0.3} + \underset{(0.240)}{0.65} X_i,$$

$$R^2 = 0.786, \quad s = 1.07,$$

ただし，係数の推定値の下の括弧内は標準誤差を表すものとする。

のように書く。

5　多重回帰

n 組のデータ $(Y_i, X_{1i}, X_{2i}, \cdots, X_{ki})$, $i = 1, 2, \cdots, n$ を用いて，k 変数の多重回帰モデルを考える。

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + u_i,$$

ただし，X_{ji} は j 番目の説明変数の第 i 番目の観測値を表す。u_i は誤差項 (または，攪乱項) で，同じ仮定を用いる (すなわち，$u_1, u_2, \cdots, u_n$ は互いに独立に，平均ゼロ，分散 σ^2 の正規分布に従う)。

$\beta_1, \beta_2, \cdots, \beta_k$ は推定されるべきパラメータである。

すべての i について，$X_{1i}=1$ とすれば，β_1 は定数項として表される。

次のような関数 $S(\beta_1, \beta_2, \cdots, \beta_k)$ を定義する。

$$\begin{aligned}
&S(\beta_1, \beta_2, \cdots, \beta_k) \\
&= \sum_{i=1}^{n} u_i^2 \\
&= \sum_{i=1}^{n} (Y_i - \beta_1 X_{1i} - \beta_2 X_{2i} - \cdots - \beta_k X_{ki})^2
\end{aligned}$$

このとき，

$$\min_{\beta_1, \beta_2, \cdots, \beta_k} S(\beta_1, \beta_2, \cdots, \beta_k)$$

となるような $\beta_1, \beta_2, \cdots, \beta_k$ を求める。$\Longrightarrow$ 最小自乗法

このときの解を $\widehat{\beta}_1, \widehat{\beta}_2, \cdots, \widehat{\beta}_k$ とする。

最小化のためには，

$$\begin{aligned}
\frac{\partial S(\beta_1, \beta_2, \cdots, \beta_k)}{\partial \beta_1} &= 0 \\
\frac{\partial S(\beta_1, \beta_2, \cdots, \beta_k)}{\partial \beta_2} &= 0 \\
&\vdots \\
\frac{\partial S(\beta_1, \beta_2, \cdots, \beta_k)}{\partial \beta_k} &= 0
\end{aligned}$$

を満たす $\beta_1, \beta_2, \cdots, \beta_k$ が $\widehat{\beta}_1, \widehat{\beta}_2, \cdots, \widehat{\beta}_k$ となる。

すなわち，$\widehat{\beta}_1, \widehat{\beta}_2, \cdots, \widehat{\beta}_k$ は，

$$\begin{aligned}
&\sum_{i=1}^{n} (Y_i - \widehat{\beta}_1 X_{1i} - \widehat{\beta}_2 X_{2i} - \cdots - \widehat{\beta}_k X_{ki}) X_{1i} = 0, \\
&\sum_{i=1}^{n} (Y_i - \widehat{\beta}_1 X_{1i} - \widehat{\beta}_2 X_{2i} - \cdots - \widehat{\beta}_k X_{ki}) X_{2i} = 0, \\
&\qquad\vdots \\
&\sum_{i=1}^{n} (Y_i - \widehat{\beta}_1 X_{1i} - \widehat{\beta}_2 X_{2i} - \cdots - \widehat{\beta}_k X_{ki}) X_{ki} = 0,
\end{aligned}$$

を満たす。

さらに，

$$\begin{aligned}
\sum_{i=1}^{n} X_{1i} Y_i &= \widehat{\beta}_1 \sum_{i=1}^{n} X_{1i}^2 + \widehat{\beta}_2 \sum_{i=1}^{n} X_{1i} X_{2i} + \cdots \\
&\qquad + \widehat{\beta}_k \sum_{i=1}^{n} X_{1i} X_{ki}, \\
\sum_{i=1}^{n} X_{2i} Y_i &= \widehat{\beta}_1 \sum_{i=1}^{n} X_{1i} X_{2i} + \widehat{\beta}_2 \sum_{i=1}^{n} X_{2i}^2 + \cdots \\
&\qquad + \widehat{\beta}_k \sum_{i=1}^{n} X_{2i} X_{ki}, \\
&\vdots \\
\sum_{i=1}^{n} X_{ki} Y_i &= \widehat{\beta}_1 \sum_{i=1}^{n} X_{1i} X_{ki} + \widehat{\beta}_2 \sum_{i=1}^{n} X_{2i} X_{ki} + \cdots \\
&\qquad + \widehat{\beta}_k \sum_{i=1}^{n} X_{ki}^2,
\end{aligned}$$

行列表示によって，

$$\begin{pmatrix} \sum X_{1i} Y_i \\ \sum X_{2i} Y_i \\ \vdots \\ \sum X_{ki} Y_i \end{pmatrix} =$$

$$\begin{pmatrix}
\sum X_{1i}^2 & \sum X_{1i} X_{2i} & \cdots & \sum X_{1i} X_{ki} \\
\sum X_{1i} X_{2i} & \sum X_{2i}^2 & \cdots & \sum X_{2i} X_{ki} \\
\vdots & \vdots & \ddots & \vdots \\
\sum X_{1i} X_{ki} & \sum X_{2i} X_{ki} & \cdots & \sum X_{ki}^2
\end{pmatrix}
\begin{pmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \\ \vdots \\ \widehat{\beta}_k \end{pmatrix},$$

が得られ，$\widehat{\beta}_1, \widehat{\beta}_2, \cdots, \widehat{\beta}_k$ についてまとめると，

$$\begin{pmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \\ \vdots \\ \widehat{\beta}_k \end{pmatrix} =$$

$$\begin{pmatrix}
\sum X_{1i}^2 & \sum X_{1i} X_{2i} & \cdots & \sum X_{1i} X_{ki} \\
\sum X_{1i} X_{2i} & \sum X_{2i}^2 & \cdots & \sum X_{2i} X_{ki} \\
\vdots & \vdots & \ddots & \vdots \\
\sum X_{1i} X_{ki} & \sum X_{2i} X_{ki} & \cdots & \sum X_{ki}^2
\end{pmatrix}^{-1}$$

$$\begin{pmatrix} \sum X_{1i} Y_i \\ \sum X_{2i} Y_i \\ \vdots \\ \sum X_{ki} Y_i \end{pmatrix},$$

を解くことになる。$\Longrightarrow$ コンピュータによって計算

5.1 重回帰モデルにおける回帰係数の意味

結論： 他の変数の影響を取り除いての被説明変数への影響を表す。

$k=2$ の単純なモデル：

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + u_i, \qquad i = 1, 2, \cdots, n$$

β_1, β_2 の最小二乗推定量は，

$$\min_{\beta_1,\beta_2} \sum_{i=1}^{n} (Y_i - \beta_1 X_{1i} - \beta_2 X_{2i})^2$$

を解いて，$\widehat{\beta}_1$，$\widehat{\beta}_2$ が次のように得られる。

$$\begin{pmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \end{pmatrix} = \begin{pmatrix} \sum_i X_{1i}^2 & \sum_i X_{1i}X_{2i} \\ \sum_i X_{1i}X_{2i} & \sum_i X_{2i}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum_i X_{1i}Y_i \\ \sum_i X_{2i}Y_i \end{pmatrix}$$
$$= \frac{1}{(\sum_i X_{1i}^2)(\sum_i X_{2i}^2) - (\sum_i X_{1i}X_{2i})^2}$$
$$\times \begin{pmatrix} \sum_i X_{2i}^2 & -\sum_i X_{1i}X_{2i} \\ -\sum_i X_{1i}X_{2i} & \sum_i X_{1i}^2 \end{pmatrix} \begin{pmatrix} \sum_i X_{1i}Y \\ \sum_i X_{2i}Y \end{pmatrix}$$
$$= \begin{pmatrix} \dfrac{(\sum_i X_{2i}^2)(\sum_i X_{1i}Y_i) - (\sum_i X_{1i}X_{2i})(\sum_i X_{2i}Y_i)}{(\sum_i X_{1i}^2)(\sum_i X_{2i}^2) - (\sum_i X_{1i}X_{2i})^2} \\ \dfrac{-(\sum_i X_{1i}X_{2i})(\sum_i X_{1i}Y_i) + (\sum_i X_{1i}^2)(\sum_i X_{2i}Y_i)}{(\sum_i X_{1i}^2)(\sum_i X_{2i}^2) - (\sum_i X_{1i}X_{2i})^2} \end{pmatrix}$$

$\sum_{i=1}^n X_{ji}X_{li}$，$\sum_{i=1}^n X_{ji}Y_i$ を $\sum_i X_{ji}X_{li}$，$\sum_i X_{ji}Y_i$ と表記する。
ただし，$j=1,2$，$l=1,2$ とする。

一方，次の 2 つの回帰式を考える。

$$Y_i = \alpha_1 X_{2i} + v_i$$
$$X_{1i} = \alpha_2 X_{2i} + w_i$$

α_1，α_2 のそれぞれの最小二乗推定量を求めると，

$$\widehat{\alpha}_1 = \frac{\sum_i X_{2i}Y_i}{\sum_i X_{2i}^2}, \qquad \widehat{\alpha}_2 = \frac{\sum_i X_{2i}X_{1i}}{\sum_i X_{2i}^2}$$

となる。
$\widehat{\alpha}_1$，$\widehat{\alpha}_2$ を用いて，残差 $\widehat{v}_i$，$\widehat{w}_i$ を下記のようにそれぞれ求める。

$$\widehat{v}_i = Y_i - \widehat{\alpha}_1 X_{2i}, \qquad \widehat{w}_i = X_{1i} - \widehat{\alpha}_2 X_{2i}$$

$\widehat{v}_i$，$\widehat{w}_i$ は Y_i，X_{1i} から X_{2i} の影響を取り除いたものと解釈できる。

更に，次の回帰式を考える。

$$\widehat{v}_i = \gamma \widehat{w}_i + \epsilon_i$$

γ の最小二乗推定量 $\widehat{\gamma}$ は $\widehat{\beta}_1$ に一致することを示す。

$$\begin{aligned}
\widehat{\gamma} &= \frac{\sum_i \widehat{w}_i \widehat{v}_i}{\sum_i \widehat{w}_i^2} \\
&= \frac{\sum_i (X_{1i} - \widehat{\alpha}_2 X_{2i})(Y_i - \widehat{\alpha}_1 X_{2i})}{\sum_i (X_{1i} - \widehat{\alpha}_2 X_{2i})^2} \\
&= \frac{\sum_i X_{1i}Y_i - \widehat{\alpha}_1 \sum_i X_{1i}X_{2i} - \widehat{\alpha}_2 \sum_i X_{2i}Y_i + \widehat{\alpha}_1\widehat{\alpha}_2 \sum_i X_{2i}^2}{\sum_i X_{1i}^2 - 2\widehat{\alpha}_2 \sum_i X_{1i}X_{2i} + \widehat{\alpha}_2^2 \sum_i X_{2i}^2} \\
&= \frac{\sum_i X_{1i}Y_i - \dfrac{(\sum_i X_{2i}Y_i)(\sum_i X_{1i}X_{2i})}{\sum_i X_{2i}^2}}{\sum_i X_{1i}^2 - \dfrac{(\sum_i X_{1i}X_{2i})^2}{\sum_i X_{2i}^2}} \\
&= \frac{(\sum_i X_{2i}^2)(\sum_i X_{1i}Y_i) - (\sum_i X_{1i}X_{2i})(\sum_i X_{2i}Y_i)}{(\sum_i X_{1i}^2)(\sum_i X_{2i}^2) - (\sum_i X_{1i}X_{2i})^2} \\
&= \widehat{\beta}_1,
\end{aligned}$$

「Y_i から X_{2i} の影響を取り除いた変数」を被説明変数,「X_{1i} から X_{2i} の影響を取り除いた変数」を説明変数とした回帰係数が β_1 に等しい。

一般化： 次の回帰モデルを考える。

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + u_i$$

j 番目の回帰係数 β_j の意味は,「Y_i から X_{1i}, $\cdots$, $X_{j-1,i}$, $X_{j+1,i}$, $\cdots$, X_{ki} (すなわち，X_{ji} 以外の説明変数) の影響を取り除いた変数」を被説明変数,「X_{ji} から X_{1i}, $\cdots$, $X_{j-1,i}$, $X_{j+1,i}$, $\cdots$, X_{ki} (すなわち，X_{ji} 以外の説明変数) の影響を取り除いた変数」を説明変数とした回帰係数となる。

5.2 推定量の性質

$\beta_1, \beta_2, \cdots, \beta_k$ の最小二乗推定量は $\widehat{\beta}_1, \widehat{\beta}_2, \cdots, \widehat{\beta}_k$ とする。
誤差項 (または，撹乱項) u_i の分散 σ^2 の推定量 s^2 は，

$$\begin{aligned}
s^2 &= \frac{1}{n-k} \sum_{i=1}^{n} \widehat{u}_i^2 \\
&= \frac{1}{n-k} \sum_{i=1}^{n} (Y_i - \widehat{\beta}_1 X_{1i} - \widehat{\beta}_2 X_{2i} - \cdots - \widehat{\beta}_k X_{ki})^2
\end{aligned}$$

として表される。

このとき，

$$\begin{aligned}
&\mathrm{E}(\widehat{\beta}_j) = \beta_j,\\
&\mathrm{plim}\widehat{\beta}_j = \beta_j,\\
&\mathrm{E}(s^2) = \sigma^2,\\
&\mathrm{plim}s^2 = \sigma^2,
\end{aligned}$$

を証明することが出来る。(証明略)

分布について： $\widehat{\beta}_1, \widehat{\beta}_2, \cdots, \widehat{\beta}_k$ の分散は以下のように表される。

$$\begin{aligned}
&\mathrm{V}\begin{pmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \\ \vdots \\ \widehat{\beta}_k \end{pmatrix}\\
&= \begin{pmatrix}
\mathrm{V}(\widehat{\beta}_1) & \mathrm{Cov}(\widehat{\beta}_1, \widehat{\beta}_2) & \cdots & \mathrm{Cov}(\widehat{\beta}_1, \widehat{\beta}_k) \\
\mathrm{Cov}(\widehat{\beta}_2, \widehat{\beta}_1) & \mathrm{V}(\widehat{\beta}_2) & \cdots & \mathrm{Cov}(\widehat{\beta}_2, \widehat{\beta}_k) \\
\vdots & \vdots & \ddots & \vdots \\
\mathrm{Cov}(\widehat{\beta}_k, \widehat{\beta}_1) & \mathrm{Cov}(\widehat{\beta}_k, \widehat{\beta}_2) & \cdots & \mathrm{V}(\widehat{\beta}_k)
\end{pmatrix}\\
&= \sigma^2 \begin{pmatrix}
\sum X_{1i}^2 & \sum X_{1i}X_{2i} & \cdots & \sum X_{1i}X_{ki} \\
\sum X_{1i}X_{2i} & \sum X_{2i}^2 & \cdots & \sum X_{2i}X_{ki} \\
\vdots & \vdots & \ddots & \vdots \\
\sum X_{1i}X_{ki} & \sum X_{2i}X_{ki} & \cdots & \sum X_{ki}^2
\end{pmatrix}^{-1}
\end{aligned}$$

$\widehat{\beta}_j$ の分散 (すなわち，上の逆行列の j 番目の対角要素) を，

$$\mathrm{V}(\widehat{\beta}_j) = \sigma^2_{\widehat{\beta}_j},$$

として，その推定量を $s^2_{\widehat{\beta}_j}$ とする。
このとき，

$$\widehat{\beta}_j \sim N(\beta_j, \sigma^2_{\widehat{\beta}_j}),$$

となり，標準化すると，

$$\frac{\widehat{\beta}_j - \beta_j}{\sigma_{\widehat{\beta}_j}} \sim N(0,1),$$

が得られる。さらに，

$$\frac{(n-k)s^2}{\sigma^2} \sim \chi^2(n-k),$$

となり (証明略)，しかも，$\widehat{\beta}_j$ と s^2 の独立性から (証明略)，

$$\frac{\widehat{\beta}_j - \beta_j}{s_{\widehat{\beta}_j}} \sim t(n-k)$$

となる。
よって，通常の区間推定や仮説検定を行うことが出来る。

決定係数について： また，決定係数 R^2 についても同様に表される。

$$\begin{aligned}
R^2 &= \frac{\sum_{i=1}^n (\widehat{Y}_i - \overline{Y})^2}{\sum_{i=1}^n (Y_i - \overline{Y})^2}\\
&= 1 - \frac{\sum_{i=1}^n \widehat{u}_i^2}{\sum_{i=1}^n (Y_i - \overline{Y})^2}
\end{aligned}$$

ただし，$\widehat{Y}_i = \widehat{\beta}_1 X_{1i} + \widehat{\beta}_2 X_{2i} + \cdots + \widehat{\beta}_k X_{ki}$，$Y_i = \widehat{Y}_i + \widehat{u}_i$ である。
R^2 は，説明変数を増やすことによって，必ず大きくなる。なぜなら，説明変数が増えることによって，$\sum_{i=1}^n \widehat{u}_i^2$ が必ず減少するからである。
R^2 を基準にすると，被説明変数にとって意味のない変数でも，説明変数が多いほど，よりよいモデルということになる。この点を改善するために，自由度修正済み決定係数 $\overline{R}^2$ を用いる。

$$\overline{R}^2 = 1 - \frac{\sum_{i=1}^n \widehat{u}_i^2/(n-k)}{\sum_{i=1}^n (Y_i - \overline{Y})^2/(n-1)},$$

$\sum_{i=1}^n \widehat{u}_i^2/(n-k)$ は u_i の分散 σ^2 の不偏推定量であり，$\sum_{i=1}^n (Y_i - \overline{Y})^2/(n-1)$ は Y_i の分散の不偏推定量である。
R^2 と $\overline{R}^2$ との関係は，

$$\overline{R}^2 = 1 - (1-R^2)\frac{n-1}{n-k},$$

となる。さらに，

$$\frac{1-\overline{R}^2}{1-R^2} = \frac{n-1}{n-k} \geq 1,$$

という関係から，$\overline{R}^2 \leq R^2$ という結果を得る。($k=1$ のときのみに，等号が成り立つ。)

数値例： 今までと同じ数値例で，$\overline{R}^2$ を計算する。

i	Y_i	X_i	X_iY_i	X_i^2	$\widehat{Y}_i$	$\widehat{u}_i$
1	6	10	60	100	6.8	-0.8
2	9	12	108	144	8.1	0.9
3	10	14	140	196	9.4	0.6
4	10	16	160	256	10.7	-0.7
合計	$\sum Y_i$	$\sum X_i$	$\sum X_iY_i$	$\sum X_i^2$	$\sum \widehat{Y}_i$	$\sum \widehat{u}_i$
	35	52	468	696	35	0
平均	$\overline{Y}$	$\overline{X}$				
	8.75	13				

まず R^2 は，

$$R^2 = 1 - \frac{\sum \widehat{u}_i^2}{\sum Y_i^2 - n\overline{Y}^2}$$

$$= 1 - \frac{(-0.8)^2 + 0.9^2 + 0.6^2 + (-0.7)^2}{35 - 4 \times 8.75^2}$$
$$= 1 - \frac{2.30}{10.75}$$
$$= 0.786$$

となり，$\overline{R}^2$ は，

$$\overline{R}^2 = 1 - \frac{\sum \widehat{u}_i^2/(n-k)}{(\sum Y_i^2 - n\overline{Y}^2)/(n-1)}$$
$$= 1 - \frac{2.30/(4-2)}{10.75/(4-1)}$$
$$= 0.679$$

となる。

注意： R^2 や $\overline{R}^2$ を比較する場合，被説明変数が同じことが必要である。被説明変数が異なる場合 (例えば，被説明変数を上昇率とするかそのままの値を用いるかによって，被説明変数が異なる)，誤差項 u_i の標準誤差で比較すべきである (標準誤差の小さいモデルを採用する)。
$\Longrightarrow$ 関数型の選択

5.3 ダミー変数について

5.3.1 異常値

データに異常値が含まれている場合，経済構造がある時期から変化した場合，ダミー変数を使う。
ダミー変数とは，0 と 1 から成る変数のことである。
例えば，データが 20 期間あるとして，9 期目のデータが，回帰直線から離れている場合 (異常値の場合) を考える。

$$D_i = \begin{cases} 0, & i \neq 9 \text{ のとき} \\ 1, & i = 9 \text{ のとき} \end{cases}$$

という変数を作り，

$$Y_i = \alpha + \delta D_i + \beta X_i + u_i$$

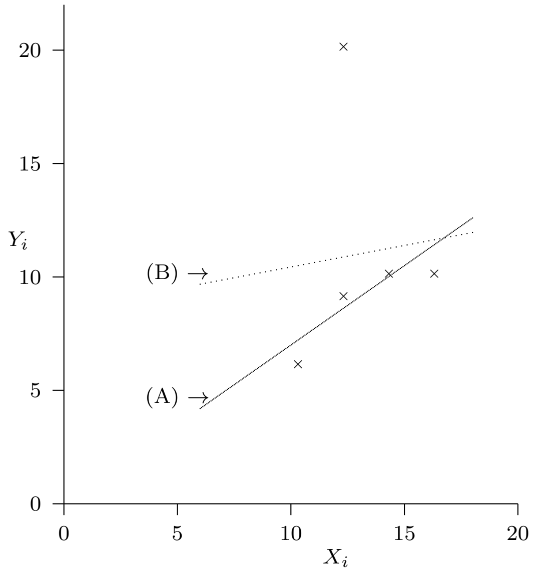
を推定する。δ の推定値 $\widehat{\delta}$ の有意性を調べることによって，異常値かどうかの検定ができる。

数値例： 今までと同様に，以下の数値例をとりあげる。

i	Y_i	X_i	D_i
1	6	10	0
2	9	12	0
3	10	14	0
4	10	16	0
5	20	12	1

第 5 期目が異常値である。

図 3： 異常値



(A) は $i = 1, 2, 3, 4$ のデータを使って，推定した回帰直線である。(B) は $i = 1, 2, 3, 4, 5$ のデータを使って，推定した回帰直線である。(A), (B) の推定結果は以下のとおりである。

$$\text{(A)}:\ Y_i = \underset{(0.095)}{0.3} + \underset{(2.708)}{0.65}\ X_i,$$
$$R^2 = 0.786, \quad s^2 = 1.072^2,$$

$$\text{(B)}:\ Y_i = \underset{(0.49)}{8.54} + \underset{(0.14)}{0.19}\ X_i,$$
$$R^2 = 0.007, \quad s^2 = 6.09^2,$$

ただし，係数の推定値の下の括弧内は t 値を表すものとする。

このように，結果が大幅に変わる。第 5 期は異常値なので，ダミー変数を用いて，

$$Y_i = \alpha + \beta X_i + \gamma D_i + u_i,$$

として推定を行う。$i = 1,2,3,4$ について，$D_i = 0$ とし，$i = 5$ について，$D_i = 1$ とする変数である。この回帰式の意味は，

$$Y_i = \begin{cases} \alpha + \beta X_i + u_i, & i = 1,2,3,4 \text{ のとき}, \\ (\alpha + \gamma) + \beta X_i + u_i, & i = 5 \text{ のとき}, \end{cases}$$

となる。推定結果は，

$$Y_i = \underset{(0.095)}{0.3} + \underset{(2.708)}{0.65} X_i + \underset{(9.73)}{11.9} D_i,$$
$$R^2 = 0.979, \quad s^2 = 1.072^2,$$

となる。この場合，$\widehat{Y}_5 = Y_5$，すなわち，$\widehat{u}_5 = 0$ となることに注意。

5.3.2 構造変化

次に，9 期目以前と以降とで，経済構造が変化している場合を考える。

$$D_i = \begin{cases} 0, & i = 1,2,\cdots,8 \text{ のとき} \\ 1, & i = 9,10,\cdots,20 \text{ のとき} \end{cases}$$

という変数を作り，

$$Y_i = \alpha + \delta D_i + \beta X_i + u_i$$

を推定する (定数項だけが変化したと考えた場合)。または，

$$Y_i = \alpha + \delta D_i + \beta X_i + \gamma D_i X_i + u_i$$

を推定する (定数項も係数も変化)。
δ や γ の推定値の有意性を調べることによって，構造変化の検定を行うことができる。
上の例でデータを示すと，

i	Y_i	X_i	D_i	D_iX_i
1	Y_1	X_1	0	0
2	Y_2	X_2	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
8	Y_8	X_8	0	0
9	Y_9	X_9	1	X_9
10	Y_{10}	X_{10}	1	X_{10}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
20	Y_{20}	X_{20}	1	X_{20}

となる。

6 関数型について

線型：

$$Y_i = \alpha + \beta X_i + u_i,$$

この場合，

$$\beta = \frac{\mathrm{d}Y_i}{\mathrm{d}X_i}$$

なので，β は，X_i が一単位上昇 (下落) したとき，Y_i は何単位上昇 (下落) するのかを表す。すなわち，β は限界係数と呼ばれる。

成長率：

$$100 \times \frac{Y_i - Y_{i-1}}{Y_{i-1}} = \alpha + \beta X_i + u_i,$$

として，成長率を被説明変数として用いる場合もある。$100 \times \dfrac{Y_i - Y_{i-1}}{Y_{i-1}}$ という変数をあらかじめ作っておき，これをこれまでの Y_i として扱う。
注意：
$Y_i = \alpha + \beta X_i + u_i$ と $100 \times \dfrac{Y_i - Y_{i-1}}{Y_{i-1}} = \alpha + \beta X_i + u_i$ では，得られる決定係数の大きさが全く異なる。単純に，R^2 や $\overline{R}^2$ による比較はこの場合出来ない。
$\Longrightarrow s^2$ で比較すればよい。

対数線型：

$$\log(Y_i) = \alpha + \beta \log(X_i) + u_i,$$

この場合，

$$\beta = \frac{\mathrm{d}\log(Y_i)}{\mathrm{d}\log(X_i)} = \frac{\dfrac{\mathrm{d}Y_i}{Y_i}}{\dfrac{\mathrm{d}X_i}{X_i}} = \frac{100\dfrac{\mathrm{d}Y_i}{Y_i}}{100\dfrac{\mathrm{d}X_i}{X_i}}$$

となる。
2 つ目の等号では，$\dfrac{\mathrm{d}\log(Y_i)}{\mathrm{d}Y_i} = \dfrac{1}{Y_i}$ が利用される。
3 つ目の等号の分子 $100\dfrac{\mathrm{d}Y_i}{Y_i}$ や分母 $100\dfrac{\mathrm{d}X_i}{X_i}$ は上昇率を表す。
したがって，β は，X_i が 1%上昇 (下落) したとき，Y_i は何%上昇 (下落) するのかを表す。β は弾力性と呼ばれる。

例： コブ＝ダグラス型生産関数：

$$Q_i = \beta_1 K_i^{\beta_2} L_i^{\beta_3}$$

ただし，Q_i は生産量，K_i は資本，L_i は労働である。この場合，対数変換によって，

$$\log(Q_i) = \beta_1' + \beta_2 \log(K_i) + \beta_3 \log(L_i) + u_i,$$

として，$\log(Q_i)$, $\log(K_i)$, $\log(L_i)$ のデータをあらかじめ変換しておき，最小二乗法で β_1', β_2, β_3 を推定する。また，生産関数には一次同次の制約 $\beta_2 + \beta_3 = 1$ を置く場合が多い。この場合は，

$$\begin{aligned}\log(Q_i) &= \beta_1' + \beta_2 \log(K_i) + \beta_3 \log(L_i) \\ &= \beta_1' + \beta_2 \log(K_i) + (1-\beta_2)\log(L_i) + u_i \\ &= \beta_1' + \beta_2\big(\log(K_i) - \log(L_i)\big) + \log(L_i) + u_i,\end{aligned}$$

となるので，

$$\log(Q_i) - \log(L_i) = \beta_1' + \beta_2\big(\log(K_i) - \log(L_i)\big) + u_i,$$

を最小二乗法で推定し，β_1', β_2 を求めることになる。この場合も同様に，各変数をあらかじめ，$\log(Q_i) - \log(L_i)$, $\log(K_i) - \log(L_i)$ としてデータを作っておく必要がある。

二次式：

$$Y_i = \alpha + \beta X_i + \gamma X_i^2 + u_i,$$

$\Longrightarrow$ 平均費用と生産量との関係等

逆数：

$$Y_i = \alpha + \beta \frac{1}{X_i} + u_i,$$

$\Longrightarrow$ 賃金上昇率と失業率との関係 (フィリップス曲線)

遅れのある変数： 習慣的効果を考慮に入れたモデル：

$$Y_i = \alpha + \beta X_i + \gamma Y_{i-1} + u_i,$$

ラグ付き内生変数が説明変数に用いられる。
X_i の Y_i への効果は，短期効果，長期効果の 2 つある。β は短期効果を表す係数である。長期効果とは，$Y_i = Y_{i-1}$ となるときの，X_i から Y_i への影響を示す効果である。

$$Y_i = \alpha + \beta X_i + \gamma Y_i + u_i,$$

として，Y_i について解くと，

$$Y_i = \frac{\alpha}{1-\gamma} + \frac{\beta}{1-\gamma} X_i + \frac{1}{1-\gamma} u_i,$$

となり，$\dfrac{\beta}{1-\gamma}$ が X_i の Y_i への長期効果を表す係数となる。
問題点：

1. 最小二乗法の仮定の一つに，説明変数は確率変数ではないという仮定がある。ラグ付き内生変数を説明変数に加えることによって，この仮定が満たされなくなる。最小二乗推定量は最小分散線型不偏推定量ではなくなる。

2. Y_i と X_i とは，経済理論的に考えると，相関が高いはず。Y_i と Y_{i-1} は相関が高い。当然，Y_{i-1} と X_i も高い相関を示す。

   $\Longrightarrow$ 多重共線性の可能性が高い。

3. DW 統計量は意味をなさない。(DW については，後述)

遅れのある変数の解釈 (部分調整モデル)： X_i が与えられたときの Y の最適水準を Y_i^* とする。

$$Y_i^* = \alpha + \beta X_i,$$

現実の水準 Y_i は，最適水準 Y_i^* と前期の水準 Y_{i-1} との差の一定割合と前期の水準 Y_{i-1} との和で与えられるとする。調整関数を考える。

$$Y_i - Y_{i-1} = \lambda(Y_i^* - Y_{i-1}) + u_i,$$

ただし，u_i は互いに独立で同一な分布の誤差項，$0 < \lambda < 1$ とする。
よって，

$$Y_i = \lambda\alpha + \lambda\beta X_i + (1-\lambda) Y_{i-1} + u_i,$$

を得る。
Y_{i-1} と u_i との相関はない。
しかし，Y_{i-1} が説明変数の一つに入っている (説明変数間が確率変数でないという仮定に反する)。
推定量は不偏推定量ではないが，一致推定量である (証明略)。

7　系列相関：DW について

7.1　DW について

最小自乗法の仮定の一つに,「攪乱項 $u_1, u_2, \cdots, u_n$ はそれぞれ独立に分布する」というものがあった。ダービン・ワトソン比 (DW) とは，誤差項の系列相関，すなわち，u_i と u_{i-1} との間の相関の有無を検定するために考案された。
$\Longrightarrow$ 時系列データのときのみ有効
$u_1, u_2, \cdots, u_n$ の系列について，それぞれの符号が，+ + + - - - - + + - - - + + のように，プラスが連続で続いた後で，マイナスが連続で続くというような場合，$u_1, u_2, \cdots, u_n$ は正の系列相関があると言う。また，+ - + - + - + - + のように交互にプラス，マイナスになる場合，$u_1, u_2, \cdots, u_n$ 負の系列相関があると言う。
特徴：$u_1, u_2, \cdots, u_i$ から u_{i+1} の符号が予想できる。$\Longrightarrow$「$u_1, u_2, \cdots, u_n$ はそれぞれ独立に分布する」という仮定に反する。
すなわち，ダービン・ワトソン比とは，回帰式が

$$Y_i = \alpha + \beta X_i + u_i,$$
$$u_i = \rho u_{i-1} + \epsilon_i,$$

のときに，H_0： $\rho = 0$, H_1： $\rho \neq 0$ の検定である。ただし，$\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ は互いに独立とする。
ダービン・ワトソン比の定義は次の通りである。

$$DW = \frac{\sum_{i=2}^n (\widehat{u}_i - \widehat{u}_{i-1})^2}{\sum_{i=1}^n \widehat{u}_i^2}$$

DW は近似的に，次のように表される。

$$\begin{aligned}
DW &= \frac{\sum_{i=2}^n (\widehat{u}_i - \widehat{u}_{i-1})^2}{\sum_{i=1}^n \widehat{u}_i^2} \\
&= \frac{\sum_{i=2}^n \widehat{u}_i^2 - 2\sum_{i=2}^n \widehat{u}_i\widehat{u}_{i-1} + \sum_{i=2}^n \widehat{u}_{i-1}^2}{\sum_{i=1}^n \widehat{u}_i^2} \\
&= \frac{2\sum_{i=1}^n \widehat{u}_i^2 - (\widehat{u}_1^2 + \widehat{u}_n^2)}{\sum_{i=1}^n \widehat{u}_i^2} - 2\frac{\sum_{i=2}^n \widehat{u}_i\widehat{u}_{i-1}}{\sum_{i=1}^n \widehat{u}_i^2} \\
&\approx 2(1-\widehat{\rho}),
\end{aligned}$$

以下の 2 つの近似が用いられる。

$$\frac{\widehat{u}_1^2 + \widehat{u}_n^2}{\sum_{i=1}^n \widehat{u}_i^2} \approx 0,$$

$$\begin{aligned}
\frac{\sum_{i=2}^n \widehat{u}_i\widehat{u}_{i-1}}{\sum_{i=1}^n \widehat{u}_i^2} &= \frac{\sum_{i=2}^n \widehat{u}_i\widehat{u}_{i-1}}{\sum_{i=2}^n \widehat{u}_{i-1}^2 + \widehat{u}_n^2} \\
&\approx \frac{\sum_{i=2}^n \widehat{u}_i\widehat{u}_{i-1}}{\sum_{i=2}^n \widehat{u}_{i-1}^2} \\
&= \widehat{\rho},
\end{aligned}$$

すなわち，$\widehat{\rho}$ は $\widehat{u}_i$ と $\widehat{u}_{i-1}$ の回帰係数である。$u_i = \rho u_{i-1} + \epsilon_i$ において，u_i, u_{i-1} の代わりに $\widehat{u}_i, \widehat{u}_{i-1}$ に置き換えて，ρ の推定値 $\widehat{\rho}$ を求める。

1. DW の値が 2 前後のとき，系列相関なし ($\widehat{\rho} = 0$ のとき，$DW \approx 2$)。
2. DW が 2 より十分に小さいとき，正の系列相関と判定される。
3. DW が 2 より十分に大きいとき，負の系列相関と判定される。

正確な判定には，データ数 n とパラメータ数 k に依存する。表 1 と表 2 を参照せよ。
表 1 と表 2 で，k' は定数項を除くパラメータ数を表すものとする。

数値例：　今までと同じ数値例で，DW を計算する。

i	Y_i	X_i	X_iY_i	X_i^2	$\widehat{Y}_i$	$\widehat{u}_i$
1	6	10	60	100	6.8	-0.8
2	9	12	108	144	8.1	0.9
3	10	14	140	196	9.4	0.6
4	10	16	160	256	10.7	-0.7
合計	$\sum Y_i$	$\sum X_i$	$\sum X_iY_i$	$\sum X_i^2$	$\sum \widehat{Y}_i$	$\sum \widehat{u}_i$
	35	52	468	696	35	0
平均	$\overline{Y}$	$\overline{X}$				
	8.75	13				

$$\begin{aligned}
DW &= \frac{\sum_{i=2}^n (\widehat{u}_i - \widehat{u}_{i-1})^2}{\sum_{i=1}^n \widehat{u}_i^2} \\
&= \frac{(-0.8-0.9)^2 + (0.9-0.6)^2 + (0.6-(-0.7))^2}{(-0.8)^2 + 0.9^2 + 0.6^2 + (-0.7)^2} \\
&= \frac{4.67}{2.30} = 2.03
\end{aligned}$$

推定結果の表記方法：　回帰モデル：

$$Y_i = \alpha + \beta X_i + u_i,$$

の推定の結果，$\widehat{\alpha} = 0.3$, $\widehat{\beta} = 0.65$, $s_{\widehat{\alpha}} = \sqrt{10.0005} = 3.163$, $s_{\widehat{\beta}} = \sqrt{0.0575} = 0.240$, $\dfrac{\widehat{\alpha}}{s_{\widehat{\alpha}}} = 0.095$, $\dfrac{\widehat{\beta}}{s_{\widehat{\beta}}} = 2.708$,

$s^2 = 1.15$ (すなわち，$s = 1.07$), $R^2 = 0.786$, $\overline{R}^2 = 0.679$, $DW = 2.03$ を得た。これらをまとめて，

$$Y_i = \underset{(0.095)}{0.3} + \underset{(2.708)}{0.65} X_i,$$
$$R^2 = 0.786, \quad \overline{R}^2 = 0.679,$$
$$s = 1.07, \quad DW = 2.03,$$

ただし，係数の推定値の下の括弧内は t 値を表すものとする。

または，

$$Y_i = \underset{(3.163)}{0.3} + \underset{(0.240)}{0.65} X_i,$$
$$R^2 = 0.786, \quad \overline{R}^2 = 0.679,$$
$$s = 1.07, \quad DW = 2.03,$$

ただし，係数の推定値の下の括弧内は標準誤差を表すものとする。

のように書く。$s = \sqrt{1.15} = 1.07$ に注意。

図 4： 正の系列相関

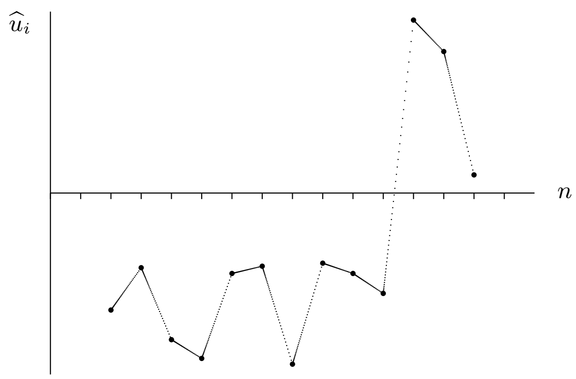
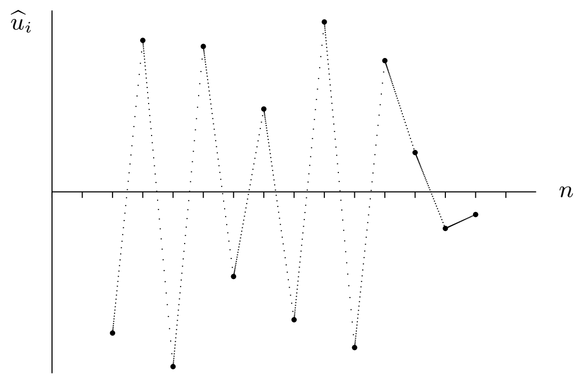


図 5： 負の系列相関



7.2 最小二乗推定量の分散について

単回帰

$$Y_i = \alpha + \beta X_i + u_i, \qquad n = 1, 2, \cdots, n$$

仮定：$\mathrm{E}(u_i) = 0$

$\mathrm{V}(u_i) = \mathrm{E}(u_i^2) = \sigma^2$

$i \neq j$ について，$\mathrm{Cov}(u_i, u_j) = \mathrm{E}(u_i u_j) = \sigma_{ij} \longleftarrow$ この仮定追加

系列相関を無視して，通常の最小二乗推定量は，

$$\widehat{\beta} = \sum_i \omega_i Y_i = \beta + \sum_i \omega_i u_i$$

である。ただし，$\omega_i = \dfrac{X_i - \overline{X}}{\sum_j (X_j - \overline{X})^2}$

$\mathrm{E}(\widehat{\beta})$ について，

$$\mathrm{E}(\widehat{\beta}) = \mathrm{E}(\beta + \sum_i \omega_i u_i)$$
$$= \beta + \sum_i \omega_i \mathrm{E}(u_i) = \beta$$

u_1, u_2, $\cdots$, u_n に系列相関があっても，$\widehat{\beta}$ は不偏推定量となる。

表 1: ダービン・ワトソン統計量の 5 % 点の上限と下限

(1) $k' = 1$

n	A		B		C		D		E	
	下限	上限	下限	上限	下限	上限	下限	上限	下限	上限
	0	dl	dl	du	du	$4-du$	$4-du$	$4-dl$	$4-dl$	4
15	0	1.08	1.08	1.36	1.36	2.64	2.64	2.92	2.92	4
20	0	1.20	1.20	1.41	1.41	2.59	2.59	2.80	2.80	4
25	0	1.29	1.29	1.45	1.45	2.55	2.55	2.71	2.71	4
30	0	1.35	1.35	1.49	1.49	2.51	2.51	2.65	2.65	4

(2) $k' = 2$

n	A		B		C		D		E	
	下限	上限	下限	上限	下限	上限	下限	上限	下限	上限
	0	dl	dl	du	du	$4-du$	$4-du$	$4-dl$	$4-dl$	4
15	0	0.95	0.95	1.54	1.54	2.46	2.46	3.05	3.05	4
20	0	1.10	1.10	1.54	1.54	2.46	2.46	2.90	2.90	4
25	0	1.21	1.21	1.55	1.55	2.45	2.45	2.79	2.79	4
30	0	1.28	1.28	1.57	1.57	2.43	2.43	2.72	2.72	4

(3) $k' = 3$

n	A		B		C		D		E	
	下限	上限	下限	上限	下限	上限	下限	上限	下限	上限
	0	dl	dl	du	du	$4-du$	$4-du$	$4-dl$	$4-dl$	4
15	0	0.82	0.82	1.75	1.75	2.25	2.25	2.25	3.18	4
20	0	1.00	1.00	1.68	1.68	2.32	2.32	2.32	3.00	4
25	0	1.12	1.12	1.66	1.66	2.34	2.34	2.34	2.88	4
30	0	1.21	1.21	1.65	1.65	2.35	2.35	2.35	2.79	4

(4) $k' = 4$

n	A		B		C		D		E	
	下限	上限	下限	上限	下限	上限	下限	上限	下限	上限
	0	dl	dl	du	du	$4-du$	$4-du$	$4-dl$	$4-dl$	4
15	0	0.69	0.69	1.97	1.97	2.03	2.03	3.31	3.31	4
20	0	0.90	0.90	1.83	1.83	2.17	2.17	3.10	3.10	4
25	0	1.04	1.04	1.77	1.77	2.23	2.23	2.96	2.96	4
30	0	1.14	1.14	1.74	1.74	2.26	2.26	2.86	2.86	4

(5) $k' = 5$

n	A		B		C		D		E	
	下限	上限	下限	上限	下限	上限	下限	上限	下限	上限
	0	dl	dl	du	du	$4-du$	$4-du$	$4-dl$	$4-dl$	4
15	0	0.56	0.56	2.21	—	—	2.21	3.44	3.44	4
20	0	0.79	0.79	1.99	1.99	2.01	2.01	3.21	3.21	4
25	0	0.95	0.95	1.89	1.89	2.11	2.11	3.05	3.05	4
30	0	1.07	1.07	1.83	1.83	2.17	2.17	2.93	2.93	4

A: 正の系列相関あり
B: 系列相関の有無を判定不能
C: 系列相関なし
D: 系列相関の有無を判定不能
E: 負の系列相関あり

表 2: ダービン・ワトソン統計量の 5 % 点の上限と下限

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$		$k'=6$		$k'=7$	
	dl	du	dl	du	dl	du	dl	du	dl	du	dl	du	dl	du
6	0.610	1.400	—	—	—	—	—	—	—	—	—	—	—	—
7	0.700	1.356	0.467	1.896	—	—	—	—	—	—	—	—	—	—
8	0.763	1.332	0.559	1.777	0.367	2.287	—	—	—	—	—	—	—	—
9	0.824	1.320	0.629	1.699	0.455	2.128	0.296	2.588	—	—	—	—	—	—
10	0.879	1.320	0.697	1.641	0.525	2.016	0.376	2.414	0.243	2.822	—	—	—	—
11	0.927	1.324	0.758	1.604	0.595	1.928	0.444	2.283	0.315	2.645	0.203	3.004	—	—
12	0.971	1.331	0.812	1.579	0.658	1.864	0.512	2.177	0.380	2.506	0.268	2.832	0.171	3.149
13	1.010	1.340	0.861	1.562	0.715	1.816	0.574	2.094	0.444	2.390	0.328	2.692	0.230	2.985
14	1.045	1.350	0.905	1.551	0.767	1.779	0.632	2.030	0.505	2.296	0.389	2.572	0.286	2.848
15	1.077	1.361	0.946	1.543	0.814	1.750	0.685	1.977	0.562	2.220	0.447	2.471	0.343	2.727
16	1.106	1.371	0.982	1.539	0.857	1.728	0.734	1.935	0.615	2.157	0.502	2.388	0.398	2.624
17	1.133	1.381	1.015	1.536	0.897	1.710	0.779	1.900	0.664	2.104	0.554	2.318	0.451	2.537
18	1.158	1.391	1.046	1.535	0.933	1.696	0.820	1.872	0.710	2.060	0.603	2.257	0.502	2.461
19	1.180	1.401	1.074	1.536	0.967	1.685	0.859	1.848	0.752	2.023	0.649	2.206	0.549	2.396
20	1.201	1.411	1.100	1.537	0.998	1.676	0.894	1.828	0.792	1.991	0.691	2.162	0.595	2.339
21	1.221	1.420	1.125	1.538	1.026	1.669	0.927	1.812	0.829	1.964	0.731	2.124	0.637	2.290
22	1.239	1.429	1.147	1.541	1.053	1.664	0.958	1.797	0.863	1.940	0.769	2.090	0.677	2.246
23	1.257	1.437	1.168	1.543	1.078	1.660	0.986	1.785	0.895	1.920	0.804	2.061	0.715	2.208
24	1.273	1.446	1.188	1.546	1.101	1.656	1.013	1.775	0.925	1.902	0.837	2.035	0.750	2.174
25	1.288	1.454	1.206	1.550	1.123	1.654	1.038	1.767	0.953	1.886	0.868	2.013	0.784	2.144
26	1.302	1.461	1.224	1.553	1.143	1.652	1.062	1.759	0.979	1.873	0.897	1.992	0.816	2.117
27	1.316	1.469	1.240	1.556	1.162	1.651	1.084	1.753	1.004	1.861	0.925	1.974	0.845	2.093
28	1.328	1.476	1.255	1.560	1.181	1.650	1.104	1.747	1.028	1.850	0.951	1.959	0.874	2.071
29	1.341	1.483	1.270	1.563	1.198	1.650	1.124	1.743	1.050	1.841	0.975	1.944	0.900	2.052
30	1.352	1.489	1.284	1.567	1.214	1.650	1.143	1.739	1.071	1.833	0.998	1.931	0.926	2.034
31	1.363	1.496	1.297	1.570	1.229	1.650	1.160	1.735	1.090	1.825	1.020	1.920	0.950	2.018
32	1.373	1.502	1.309	1.574	1.244	1.650	1.177	1.732	1.109	1.819	1.041	1.909	0.972	2.004
33	1.383	1.508	1.321	1.577	1.258	1.651	1.193	1.730	1.127	1.813	1.061	1.900	0.994	1.991
34	1.393	1.514	1.333	1.580	1.271	1.652	1.208	1.728	1.144	1.808	1.079	1.891	1.015	1.978
35	1.402	1.519	1.343	1.584	1.283	1.653	1.222	1.726	1.160	1.803	1.097	1.884	1.034	1.967
36	1.411	1.525	1.354	1.587	1.295	1.654	1.236	1.724	1.175	1.799	1.114	1.876	1.053	1.957
37	1.419	1.530	1.364	1.590	1.307	1.655	1.249	1.723	1.190	1.795	1.131	1.870	1.071	1.948
38	1.427	1.535	1.373	1.594	1.318	1.656	1.261	1.722	1.204	1.792	1.146	1.864	1.088	1.939
39	1.435	1.540	1.382	1.597	1.328	1.658	1.273	1.722	1.218	1.789	1.161	1.859	1.104	1.932
40	1.442	1.544	1.391	1.600	1.338	1.659	1.285	1.721	1.230	1.786	1.175	1.854	1.120	1.924
45	1.475	1.566	1.430	1.615	1.383	1.666	1.336	1.720	1.287	1.776	1.238	1.835	1.189	1.895
50	1.503	1.585	1.462	1.628	1.421	1.674	1.378	1.721	1.335	1.771	1.291	1.822	1.246	1.875
55	1.528	1.601	1.490	1.641	1.452	1.681	1.414	1.724	1.374	1.768	1.334	1.814	1.294	1.861
60	1.549	1.616	1.514	1.652	1.480	1.689	1.444	1.727	1.408	1.767	1.372	1.808	1.335	1.850
65	1.567	1.629	1.536	1.662	1.503	1.696	1.471	1.731	1.438	1.767	1.404	1.805	1.370	1.843
70	1.583	1.641	1.554	1.672	1.525	1.703	1.494	1.735	1.464	1.768	1.433	1.802	1.401	1.838
75	1.598	1.652	1.571	1.680	1.543	1.709	1.515	1.739	1.487	1.770	1.458	1.801	1.428	1.834
80	1.611	1.662	1.586	1.688	1.560	1.715	1.534	1.743	1.507	1.772	1.480	1.801	1.453	1.831
85	1.623	1.671	1.600	1.696	1.575	1.721	1.550	1.747	1.525	1.774	1.500	1.801	1.474	1.829
90	1.635	1.679	1.612	1.703	1.589	1.726	1.566	1.751	1.542	1.776	1.518	1.801	1.494	1.827
95	1.645	1.687	1.623	1.709	1.602	1.732	1.579	1.755	1.557	1.778	1.535	1.802	1.512	1.827
100	1.654	1.694	1.634	1.715	1.613	1.736	1.592	1.758	1.571	1.780	1.550	1.803	1.528	1.826
150	1.720	1.747	1.706	1.760	1.693	1.774	1.679	1.788	1.665	1.802	1.651	1.817	1.637	1.832
200	1.758	1.779	1.748	1.789	1.738	1.799	1.728	1.809	1.718	1.820	1.707	1.831	1.697	1.841

n	$k'=8$		$k'=9$		$k'=10$		$k'=11$		$k'=12$		$k'=13$	
	dl	du	dl	du	dl	du	dl	du	dl	du	dl	du
6	—	—	—	—	—	—	—	—	—	—	—	—
7	—	—	—	—	—	—	—	—	—	—	—	—
8	—	—	—	—	—	—	—	—	—	—	—	—
9	—	—	—	—	—	—	—	—	—	—	—	—
10	—	—	—	—	—	—	—	—	—	—	—	—
11	—	—	—	—	—	—	—	—	—	—	—	—
12	—	—	—	—	—	—	—	—	—	—	—	—
13	0.147	3.266	—	—	—	—	—	—	—	—	—	—
14	0.200	3.111	0.127	3.360	—	—	—	—	—	—	—	—
15	0.251	2.979	0.175	3.216	0.111	3.438	—	—	—	—	—	—
16	0.304	2.860	0.222	3.090	0.155	3.304	0.098	3.503	—	—	—	—
17	0.356	2.757	0.272	2.975	0.198	3.184	0.138	3.378	0.087	3.557	—	—
18	0.407	2.668	0.321	2.873	0.244	3.073	0.177	3.265	0.123	3.441	0.078	3.603
19	0.456	2.589	0.369	2.783	0.290	2.974	0.220	3.159	0.160	3.335	0.111	3.496
20	0.502	2.521	0.416	2.704	0.336	2.885	0.263	3.063	0.200	3.234	0.145	3.395
21	0.546	2.461	0.461	2.633	0.380	2.806	0.307	2.976	0.240	3.141	0.182	3.300
22	0.588	2.407	0.504	2.571	0.424	2.735	0.349	2.897	0.281	3.057	0.220	3.211
23	0.628	2.360	0.545	2.514	0.465	2.670	0.391	2.826	0.322	2.979	0.259	3.129
24	0.666	2.318	0.584	2.464	0.506	2.613	0.431	2.761	0.362	2.908	0.297	3.053
25	0.702	2.280	0.621	2.419	0.544	2.560	0.470	2.702	0.400	2.844	0.335	2.983
26	0.735	2.246	0.657	2.379	0.581	2.513	0.508	2.649	0.438	2.784	0.373	2.919
27	0.767	2.216	0.691	2.342	0.616	2.470	0.544	2.600	0.475	2.730	0.409	2.860
28	0.798	2.188	0.723	2.309	0.649	2.431	0.578	2.555	0.510	2.680	0.445	2.805
29	0.826	2.164	0.753	2.278	0.681	2.396	0.612	2.515	0.544	2.634	0.479	2.754
30	0.854	2.141	0.782	2.251	0.712	2.363	0.643	2.477	0.577	2.592	0.513	2.708
31	0.879	2.120	0.810	2.226	0.741	2.333	0.674	2.443	0.608	2.553	0.545	2.665
32	0.904	2.102	0.836	2.203	0.769	2.306	0.703	2.411	0.638	2.518	0.576	2.625
33	0.927	2.085	0.861	2.181	0.796	2.281	0.731	2.382	0.667	2.484	0.606	2.588
34	0.950	2.069	0.885	2.162	0.821	2.257	0.758	2.355	0.695	2.454	0.634	2.553
35	0.971	2.054	0.908	2.144	0.845	2.236	0.783	2.330	0.722	2.425	0.662	2.521
36	0.991	2.041	0.930	2.127	0.868	2.216	0.808	2.306	0.748	2.398	0.689	2.492
37	1.011	2.029	0.951	2.112	0.891	2.197	0.831	2.285	0.772	2.374	0.714	2.464
38	1.029	2.017	0.970	2.098	0.912	2.180	0.854	2.265	0.796	2.351	0.739	2.438
39	1.047	2.007	0.990	2.085	0.932	2.164	0.875	2.246	0.819	2.329	0.763	2.413
40	1.064	1.997	1.008	2.072	0.952	2.150	0.896	2.228	0.840	2.309	0.785	2.391
45	1.139	1.958	1.089	2.022	1.038	2.088	0.988	2.156	0.938	2.225	0.887	2.296
50	1.201	1.930	1.156	1.986	1.110	2.044	1.064	2.103	1.019	2.163	0.973	2.225
55	1.253	1.909	1.212	1.959	1.170	2.010	1.129	2.062	1.087	2.116	1.045	2.170
60	1.298	1.894	1.260	1.939	1.222	1.984	1.184	2.031	1.145	2.079	1.106	2.127
65	1.336	1.882	1.301	1.923	1.266	1.964	1.231	2.006	1.195	2.049	1.160	2.093
70	1.369	1.874	1.337	1.910	1.305	1.948	1.272	1.987	1.239	2.026	1.206	2.066
75	1.399	1.867	1.369	1.901	1.339	1.935	1.308	1.970	1.277	2.006	1.247	2.043
80	1.425	1.861	1.397	1.893	1.369	1.925	1.340	1.957	1.312	1.990	1.283	2.024
85	1.448	1.857	1.422	1.886	1.396	1.916	1.369	1.946	1.342	1.977	1.315	2.008
90	1.469	1.854	1.445	1.881	1.420	1.909	1.395	1.937	1.369	1.966	1.344	1.995
95	1.489	1.852	1.465	1.877	1.442	1.903	1.418	1.930	1.394	1.956	1.370	1.984
100	1.506	1.850	1.484	1.874	1.462	1.898	1.439	1.923	1.416	1.948	1.393	1.974
150	1.622	1.846	1.608	1.862	1.593	1.877	1.579	1.892	1.564	1.908	1.549	1.924
200	1.686	1.852	1.675	1.863	1.665	1.874	1.654	1.885	1.643	1.897	1.632	1.908

n は標本数, k' は定数項を除く説明変数の数とする。

（出所）http://www.stanford.edu/~clint/bench/dwcrit.htm

$\mathrm{V}(\widehat{\beta})$ について，

$$\begin{aligned}
\mathrm{V}(\widehat{\beta}) &= \mathrm{V}(\beta+\sum_i \omega_i u_i) = \mathrm{V}(\sum_i \omega_i u_i) \\
&= \mathrm{E}((\sum_i \omega_i u_i)^2) \quad \longleftarrow \text{分散の定義} \\
&= \mathrm{E}((\sum_i \omega_i u_i)(\sum_j \omega_j u_j)) \quad \longleftarrow \text{添字一つ変更} \\
&= \mathrm{E}(\sum_i \sum_j \omega_i \omega_j u_i u_j) \\
&= \sum_i \sum_j \omega_i \omega_j \mathrm{E}(u_i u_j) \\
&= \sum_i \omega_i^2 \mathrm{E}(u_i^2) + \sum_{\substack{i \\ i \neq j}} \sum_j \omega_i \omega_j \mathrm{E}(u_i u_j) \\
&= \sigma^2 \sum_i \omega_i^2 + \sum_{\substack{i \\ i \neq j}} \sum_j \sigma_{ij} \omega_i \omega_j \\
&\neq \sigma^2 \sum_i \omega_i^2
\end{aligned}$$

したがって，u_1，u_2，$\cdots$，u_n に系列相関があるとき，通常の最小二乗推定量 $\widehat{\beta}$ の分散の推定量は，$s^2 \sum_i \omega_i^2 + \sum_{\substack{i \\ i \neq j}} \sum_j s_{ij} \omega_i \omega_j$ とならなければならない。
s^2, s_{ij} は σ^2, σ_{ij} の推定量とする。
しかし，計量ソフトは $s^2 \sum_i \omega_i^2$ と計算する。

7.3 系列相関のもとで回帰式の推定

回帰式が

$$\begin{aligned}
Y_i &= \alpha + \beta X_i + u_i, \\
u_i &= \rho u_{i-1} + \epsilon_i,
\end{aligned}$$

のときの推定を考える。ただし，$\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ は互いに独立とする。
u_i を消去すると，

$$(Y_i - \alpha - \beta X_i) = \rho(Y_{i-1} - \alpha - \beta X_{i-1}) + \epsilon_i,$$

また，少し整理すると，

$$(Y_i - \rho Y_{i-1}) = \alpha(1-\rho) + \beta(X_i - \rho X_{i-1}) + \epsilon_i,$$

となる。
$\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ は互いに独立とするので，最小二乗法を適用が可能となる。しかし，通常の最小二乗法と同様に，残差平方和を最小にするような推定量 $\hat{\alpha}$，$\hat{\beta}$，$\hat{\rho}$ を求めたいが，$\hat{\alpha}$，$\hat{\beta}$，$\hat{\rho}$ の解を (X_i, Y_i) の陽関数の形で書き表すことは不可能である。したがって，少し工夫が必要となる。
残差 $\hat{\epsilon}_i$ を次のように二通りの表し方をする。

- $\hat{\epsilon}_i = \hat{u}_i - \hat{\rho}\hat{u}_{i-1}$，ただし，$\hat{u}_i = Y_i - \hat{\alpha} - \hat{\beta} X_i$
- $\hat{\epsilon}_i = Y_i^* - \hat{\alpha}' - \hat{\beta} X_i^*$，ただし，$Y_i^* = Y_i - \hat{\rho} Y_{i-1}$，$X_i^* = X_i - \hat{\rho} X_{i-1}$，$\hat{\alpha}' = \hat{\alpha}(1-\hat{\rho})$

残差平方和 $\sum_{i=2}^n \hat{\epsilon}_i$ を $S(\hat{\alpha},\ \hat{\beta},\ \hat{\rho})$ とおく。$S(\hat{\alpha},\ \hat{\beta},\ \hat{\rho})$ を最小にするような $\hat{\alpha}$，$\hat{\beta}$，$\hat{\rho}$ を求める。すなわち，次の最小化問題を解く。

$$\min_{\hat{\alpha},\ \hat{\beta},\ \hat{\rho}} S(\hat{\alpha},\ \hat{\beta},\ \hat{\rho})$$

$\hat{\alpha}$，$\hat{\beta}$，$\hat{\rho}$ について $S(\hat{\alpha},\ \hat{\beta},\ \hat{\rho})$ を微分してゼロとおいて，

$$\begin{aligned}
\frac{\partial S(\hat{\alpha},\hat{\beta},\hat{\rho})}{\partial \hat{\alpha}} &= \frac{\partial S(\hat{\alpha},\hat{\beta},\hat{\rho})}{\partial \hat{\alpha}'} \frac{\partial \hat{\alpha}'}{\partial \hat{\alpha}} \\
&= -2(1-\hat{\rho}) \sum_{i=2}^n (Y_i^* - \hat{\alpha}' - \hat{\beta} X_i^*) = 0 \\
\frac{\partial S(\hat{\alpha},\hat{\beta},\hat{\rho})}{\partial \hat{\beta}} &= -2 \sum_{i=2}^n X_i^* (Y_i^* - \hat{\alpha}' - \hat{\beta} X_i^*) = 0 \\
\frac{\partial S(\hat{\alpha},\hat{\beta},\hat{\rho})}{\partial \hat{\rho}} &= -2 \sum_{i=2}^n \hat{u}_{i-1}(\hat{u}_i - \hat{\rho}\hat{u}_{i-1}) = 0
\end{aligned}$$

の 3 つの連立方程式を解く。すなわち，

$$\begin{aligned}
\hat{\beta} &= \frac{\sum_{i=2}^n (X_i^* - \overline{X}_i^*)(Y_i^* - \overline{Y}_i^*)}{\sum_{i=2}^n (X_i^* - \overline{X}_i^*)^2} \\
\hat{\alpha}' &= \hat{\alpha}(1-\hat{\rho}) = \overline{Y}_i^* - \hat{\beta}\overline{X}_i^* \\
\hat{\rho} &= \frac{\sum_{i=2}^n \hat{u}_{i-1}\hat{u}_i}{\sum_{i=2}^n \hat{u}_{i-1}^2}
\end{aligned}$$

を解くことになる。ただし，$\overline{X}_i^* = \dfrac{1}{n-1}\sum_{i=2}^n X_i^*$，$\overline{Y}_i^* = \dfrac{1}{n-1}\sum_{i=2}^n Y_i^*$ とする（$n-1$ 個のデータの平均)。
計算の手順として，

(i) $\hat{\rho}$ に与えたもとで（最初は $\hat{\rho} = 0$），上記最初の 2 つの式を用いて，$(\hat{\alpha},\ \hat{\beta})$ を求める。

(ii) $(\hat{\alpha},\ \hat{\beta})$ を与えたもとで，上記 3 つ目の式を用いて，$\hat{\rho}$ を求める。

(iii) 上記手順 (i) と (ii) を交互に，$(\hat{\alpha},\ \hat{\beta},\ \hat{\rho})$ が収束するまで繰り返す。

とする。この計算手順はコクラン＝オーカット法 (Cochrane-Orcutt estimation) と呼ばれる。

簡便法：$\boldsymbol{\rho}$ の求め方： より簡単なもう一つの方法として，DW は近似的に $DW \approx 2(1-\widehat{\rho})$ と表されるので，DW から $\widehat{\rho}$ を逆算して求める。そして，$Y_i^* = Y_i - \widehat{\rho}Y_{i-1}$，$X_i^* = X_i - \widehat{\rho}X_{i-1}$) を新たな変数として，

$$Y_i^* = \alpha' + \beta X_i^* + \epsilon_i,$$

に最小二乗法を適用する。ただし，$\alpha' = \alpha(1-\hat{\rho})$ とする。

より一般的に，回帰式が

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + u_i,$$
$$u_i = \rho u_{i-1} + \epsilon_i,$$

の重回帰ときの推定を考える。ただし，$\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ は互いに独立とする。
u_i を消去すると，

$$\begin{aligned}(Y_i - \rho Y_{i-1}) &= \beta_1(X_{1i} - \rho X_{1,i-1}) \\ &\quad + \beta_2(X_{2i} - \rho X_{2,i-1}) + \cdots \\ &\quad + \beta_k(X_{ki} - \rho X_{k,i-1}) + \epsilon_i,\end{aligned}$$

となる。残差平方和 $S(\hat{\beta}_1,\ \hat{\beta}_2,\ \cdots,\ \hat{\beta}_k,\ \hat{\rho}) = \sum_{i=2}^n \hat{\epsilon}_i^2$ を最小にする $\hat{\beta}_1,\ \hat{\beta}_2,\ \cdots,\ \hat{\beta}_k,\ \hat{\rho}$ を求める。ただし，$\hat{\epsilon}_i = Y_i^* - \hat{\beta}_1 X_{1i}^* - \hat{\beta}_2 X_{2i}^* - \ \cdots\ - \hat{\beta}_k X_{ki}^*$，または，$\hat{\epsilon}_i = \hat{u}_i - \hat{\rho}\hat{u}_{i-1}$ のどちらかで残差が表される。また，式中の記号は，$Y_i^* = Y_i - \hat{\rho}Y_{i-1}$，$X_{ji}^* = X_{ji} - \hat{\rho}X_{j,i-1}$，$j = 1, 2, \cdots, k$，$\hat{u}_i = Y_i - \beta_1 X_{1i} - \beta_2 X_{2i} - \ \cdots\ - \beta_k X_{ki}$ である。
最小化のための一階の条件は，

$$\frac{S(\hat{\beta}_1,\ \hat{\beta}_2,\ \cdots,\ \hat{\beta}_k,\ \hat{\rho})}{\hat{\beta}_j} = 0, \qquad j = 1, 2, \cdots, k$$
$$\frac{S(\hat{\beta}_1,\ \hat{\beta}_2,\ \cdots,\ \hat{\beta}_k,\ \hat{\rho})}{\hat{\rho}} = 0$$

であり，具体的に計算すると，

$$\begin{aligned}&\frac{S(\hat{\beta}_1,\ \hat{\beta}_2,\ \cdots,\ \hat{\beta}_k,\ \hat{\rho})}{\hat{\beta}_j} \\ &= -2\sum_{i=2}^n X_{ji}^*(Y_i^* - \hat{\beta}_1 X_{1i}^* - \hat{\beta}_2 X_{2i}^* - \ \cdots\ - \hat{\beta}_k X_{ki}^*) \\ &= 0, \qquad\qquad j = 1, 2, \cdots, k\end{aligned}$$

$$\begin{aligned}&\frac{S(\hat{\beta}_1,\ \hat{\beta}_2,\ \cdots,\ \hat{\beta}_k,\ \hat{\rho})}{\hat{\rho}} \\ &= -2\sum_{i=2}^n \hat{u}_{i-1}(\hat{u}_i - \hat{\rho}\hat{u}_{i-1}) = 0\end{aligned}$$

となる。$(k+1)$ 本の連立方程式から，$\hat{\beta}_1$，$\hat{\beta}_2$，$\cdots$，$\hat{\beta}_k$，$\hat{\rho}$ が得られる。単回帰のときと同様に，収束計算によって求めることになる。$\longrightarrow$ コクラン＝オーカット法 (Cochrane-Orcutt estimation)

簡便法：$\boldsymbol{\rho}$ の求め方： より簡単なもう一つの方法として，DW は近似的に $DW \approx 2(1-\widehat{\rho})$ と表されるので，DW は近似的に $DW \approx 2(1-\widehat{\rho})$ と表されるので，DW から $\widehat{\rho}$ を逆算して求める。そして，$Y_i^* = Y_i - \widehat{\rho}Y_{i-1}$，$X_{1i}^* = X_{1i} - \widehat{\rho}X_{1,i-1}$，$X_{2i}^* = X_{2i} - \widehat{\rho}X_{2,i-1}$，$\cdots$，$X_{ki}^* = X_{ki} - \widehat{\rho}X_{k,i-1}$ を新たな変数として，

$$Y_i^* = \beta_1 X_{1i}^* + \beta_2 X_{2i}^* + \cdots + \beta_k X_{ki}^* + \epsilon_i,$$

に最小二乗法を適用する。

8 不均一分散 (不等分散)

8.1 不均一分散 (不等分散) の意味と推定方法

回帰式が

$$Y_i = \alpha + \beta X_i + u_i$$

の場合を考える。X_i が外生変数，Y_i は内生変数，u_i は互いに独立な同一の分布を持つ攪乱項 (最小二乗法に必要な仮定) とする。「独立な同一の分布」の意味は「攪乱項 u_1，$u_2, \cdots, u_n$ はそれぞれ独立に平均ゼロ，分散 σ^2 の分布する」である。
分散が時点に依存する場合，代表的には，分散が他の変数 (例えば，z_i) に依存する場合，すなわち，u_i の平均はゼロ，分散は $\sigma_*^2 z_i^2$ の場合は，最小二乗法の仮定に反する。そのため，単純には，$Y_i = \alpha + \beta X_i + u_i$ に最小二乗法を適用できない。以下のような修正が必要となる。

$$\frac{Y_i}{z_i} = \alpha\frac{1}{z_i} + \beta\frac{X_i}{z_i} + \frac{u_i}{z_i}$$

$$= \alpha \frac{1}{z_i} + \beta \frac{X_i}{z_i} + u_i^*$$

このとき，新たな攪乱項 u_i^* は平均ゼロ，分散 σ_*^2 の分布となる (すなわち,「同一の」分布)。

$$\mathrm{E}(u_i^*) = \mathrm{E}\left(\frac{u_i}{z_i}\right) = \left(\frac{1}{z_i}\right)\mathrm{E}(u_i) = 0$$

u_i の仮定 $\mathrm{E}(u_i) = 0$ が使われている。

$$\mathrm{V}(u_i^*) = \mathrm{V}\left(\frac{u_i}{z_i}\right) = \left(\frac{1}{z_i}\right)^2 \mathrm{V}(u_i) = \sigma_*^2$$

u_i の仮定 $\mathrm{V}(u_i) = \sigma_*^2 z_i^2$ が最後に使われている。
よって，$\dfrac{Y_i}{z_i}$, $\dfrac{1}{z_i}$, $\dfrac{X_i}{z_i}$ を新たな変数として，最小二乗法を適用することができる。

不均一分散の検定について

$$\widehat{u}_i^2 = \gamma z_i^2 + \epsilon_i$$

を推定し，γ の推定値 $\widehat{\gamma}$ の有意性の検定を行う (通常の t 検定)。

z_i は回帰式に含まれる変数でもよい。例えば，u_i の平均はゼロ，分散は $\sigma_*^2 X_i^2$ の場合，各変数を X_i で割って，

$$\frac{Y_i}{X_i} = \alpha \frac{1}{X_i} + \beta + \frac{u_i}{X_i}$$
$$= \alpha \frac{1}{X_i} + \beta + u_i^*$$

を推定すればよい。β は定数項として推定されるが，意味は限界係数 (すなわち，傾き) と同じなので注意すること。

8.2 最小二乗推定量の分散について

単回帰

$$Y_i = \alpha + \beta X_i + u_i, \qquad n = 1, 2, \cdots, n$$

仮定：$\mathrm{E}(u_i) = 0$
$\mathrm{V}(u_i) = \mathrm{E}(u_i^2) = \sigma_i^2 \longleftarrow$ この仮定追加
$i \neq j$ について，$\mathrm{Cov}(u_i, u_j) = \mathrm{E}(u_i u_j) = 0$

不均一分散を無視して，通常の最小二乗推定量は，

$$\widehat{\beta} = \sum_i \omega_i Y_i = \beta + \sum_i \omega_i u_i$$

である。ただし，$\omega_i = \dfrac{X_i - \overline{X}}{\sum_j (X_j - \overline{X})^2}$

$\mathrm{E}(\widehat{\beta})$ について，

$$\mathrm{E}(\widehat{\beta}) = \mathrm{E}(\beta + \sum_i \omega_i u_i)$$
$$= \beta + \sum_i \omega_i \mathrm{E}(u_i) = \beta$$

$u_1, u_2, \cdots, u_n$ の分散が不均一であっても，$\widehat{\beta}$ は不偏推定量となる。

$\mathrm{V}(\widehat{\beta})$ について，

$$\begin{aligned}
\mathrm{V}(\widehat{\beta}) &= \mathrm{V}(\beta + \sum_i \omega_i u_i) = \mathrm{V}(\sum_i \omega_i u_i) \\
&= \mathrm{E}((\sum_i \omega_i u_i)^2) \quad \longleftarrow \text{分散の定義} \\
&= \mathrm{E}((\sum_i \omega_i u_i)(\sum_j \omega_j u_j)) \quad \longleftarrow \text{添字一つ変更} \\
&= \mathrm{E}(\sum_i \sum_j \omega_i \omega_j u_i u_j) \\
&= \sum_i \sum_j \omega_i \omega_j \mathrm{E}(u_i u_j) \\
&= \sum_i \omega_i^2 \mathrm{E}(u_i^2) + \sum_i \sum_{\substack{j \\ i \neq j}} \omega_i \omega_j \mathrm{E}(u_i u_j) \\
&= \sum_i \sigma_i^2 \omega_i^2 \neq \sigma^2 \sum_i \omega_i^2
\end{aligned}$$

したがって，$u_1, u_2, \cdots, u_n$ の分散が不均一であるとき，通常の最小二乗推定量 $\widehat{\beta}$ の分散の推定量は，$s_i^2 \sum_i \omega_i^2$ とならなければならない。
s_i^2 は σ_i^2 の推定量とする。
しかし，計量ソフトは $s^2 \sum_i \omega_i^2$ と計算する。

9 多重共線性について

回帰式が

$$Y_i = \alpha W_i + \beta X_i + u_i$$

の場合を考える。W_i, X_i が外生変数，Y_i は内生変数，u_i は互いに独立な攪乱項とする。$W_i = 1$ のとき，α は定数項となる。

W_i と X_i の相関が大きいことを多重共線性が強いと言う。
W_i と X_i の相関が大きい場合は，α, β の推定値は不安定になる。
極端な場合，W_i と X_i の相関が 1 の場合 (完全相関の場合) は，すべての i について，$W_i = \gamma X_i$ となる。この場合，回帰式は

$$\begin{aligned} Y_i &= \alpha W_i + \beta X_i + u_i \\ &= (\alpha\gamma + \beta) X_i + u_i \end{aligned}$$

となり，$\alpha\gamma + \beta$ を推定することは可能だが，α, β を別々に推定することはできなくなる。$Y_i = \alpha W_i + \beta X_i + u_i$ を推定した場合，$\alpha\gamma + \beta$ の推定値が一定値となる $\widehat{\alpha}, \widehat{\beta}$ の組み合わせは無数に存在する。この意味で，$\widehat{\alpha}, \widehat{\beta}$ は不安定であると言える。
厳密には，最小二乗法によると，

$$\sum_{i=1}^{n} u_i^2 = \sum_{i=1}^{n} (Y_i - \alpha W_i - \beta X_i)^2$$

を最小にする α, β をその推定値 $\widehat{\alpha}, \widehat{\beta}$ とする。
すなわち，

$$\frac{\partial \sum u_i^2}{\partial \alpha} = -2 \sum_{i=1}^{n} (Y_i - \alpha W_i - \beta X_i) W_i = 0$$

$$\frac{\partial \sum u_i^2}{\partial \beta} = -2 \sum_{i=1}^{n} (Y_i - \alpha W_i - \beta X_i) X_i = 0$$

の連立方程式を解くことになる。

$$\begin{aligned} &\sum_{i=1}^{n} Y_i W_i - \widehat{\alpha} \sum_{i=1}^{n} W_i^2 - \widehat{\beta} \sum_{i=1}^{n} X_i W_i = 0 \\ &\sum_{i=1}^{n} Y_i X_i - \widehat{\alpha} \sum_{i=1}^{n} W_i X_i - \widehat{\beta} \sum_{i=1}^{n} X_i^2 = 0 \end{aligned}$$

行列表示により，

$$\begin{pmatrix} \sum Y_i W_i \\ \sum Y_i X_i \end{pmatrix} = \begin{pmatrix} \sum W_i^2 & \sum X_i W_i \\ \sum W_i X_i & \sum X_i^2 \end{pmatrix} \begin{pmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{pmatrix}$$

$\widehat{\alpha}, \widehat{\beta}$ について表すと，

$$\begin{pmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{pmatrix} = \begin{pmatrix} \sum W_i^2 & \sum X_i W_i \\ \sum W_i X_i & \sum X_i^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum Y_i W_i \\ \sum Y_i X_i \end{pmatrix}$$

逆行列を計算して，

$$\begin{aligned} \begin{pmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{pmatrix} = &\frac{1}{(\sum W_i^2)(\sum X_i^2) - (\sum W_i X_i)^2} \\ &\times \begin{pmatrix} \sum X_i^2 & -\sum X_i W_i \\ -\sum W_i X_i & \sum W_i^2 \end{pmatrix} \begin{pmatrix} \sum Y_i W_i \\ \sum Y_i X_i \end{pmatrix} \end{aligned}$$

完全な多重共線性の場合 ($W_i = \gamma X_i$ の場合)，

$$(\sum W_i^2)(\sum X_i^2) - (\sum W_i X_i) = 0$$

となる。
また，

$$\begin{aligned} \mathrm{V}\begin{pmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{pmatrix} &= \begin{pmatrix} \mathrm{V}(\widehat{\alpha}) & \mathrm{Cov}(\widehat{\alpha}, \widehat{\beta}) \\ \mathrm{Cov}(\widehat{\alpha}, \widehat{\beta}) & \mathrm{V}(\widehat{\beta}) \end{pmatrix} \\ &= \sigma^2 \begin{pmatrix} \sum W_i^2 & \sum X_i W_i \\ \sum W_i X_i & \sum X_i^2 \end{pmatrix}^{-1} \\ &= \frac{\sigma^2}{(\sum W_i^2)(\sum X_i^2) - (\sum W_i X_i)^2} \\ &\quad \times \begin{pmatrix} \sum X_i^2 & -\sum X_i W_i \\ -\sum W_i X_i & \sum W_i^2 \end{pmatrix} \end{aligned}$$

となるので，完全な多重共線性の場合は，推定値の分散が無限大となる。推定値の分散が無限大という意味は，どこにパラメータがあるか分からないということを意味する。
簡単化のため，$\overline{W} = \dfrac{1}{n} \sum W_i = 0$，$\overline{X} = \dfrac{1}{n} \sum X_i = 0$ とする。W_i と X_i との相関係数を r とすると，

$$\begin{aligned} r &= \frac{\sum (W_i - \overline{w})(X_i - \overline{X})}{\sqrt{\sum (W_i - \overline{W})^2 \sum (X_i - \overline{X})^2}} \\ &= \frac{\sum W_i X_i}{\sqrt{\sum W_i^2 \sum X_i^2}} \end{aligned}$$

となる。さらに，r を用いて，$\mathrm{V}(\widehat{\alpha})$，$\mathrm{V}(\widehat{\beta})$ を求めると，

$$\begin{aligned} \mathrm{V}(\widehat{\alpha}) &= \frac{\sigma^2 \sum X_i^2}{(\sum W_i^2)(\sum X_i^2) - (\sum W_i X_i)^2} \\ &= \frac{\sigma^2}{(1 - r^2) \sum W_i^2} \end{aligned}$$

$$\begin{aligned} \mathrm{V}(\widehat{\beta}) &= \frac{\sigma^2 \sum W_i^2}{(\sum W_i^2)(\sum X_i^2) - (\sum W_i X_i)^2} \\ &= \frac{\sigma^2}{(1 - r^2) \sum X_i^2} \end{aligned}$$

が得られる。
これは，r が 1 または -1 に近づくにつれて (または，r^2 が 1 に近づくにつれて)，$\mathrm{V}(\widehat{\alpha})$，$\mathrm{V}(\widehat{\beta})$ は大きくなるということを意味する。
$\Longrightarrow$ 係数の推定値の有意性が低くなる。
$\Longrightarrow$ 本来は W_i や X_i が Y_i に影響を与えているにもかかわらず，統計的に有意な推定値は得られなくなるので，回帰分析によって理論モデルを立証しようという試みは成功しなくなる。

多重共線性の症状： 多重共線性が起こっていると考えられるケースは，

1. 推定値の符号が理論と合わない。
2. 決定係数 (R^2 や $\overline{R}^2$) は大きいのに，個々の t 値は小さい。
3. 観測値の数 (データ数) を少し増やすと，推定値が大きく変わる。
4. 説明変数を増減すると，推定値が大きく変動する。

等である。

10 $\boldsymbol{F}$ 検定について

複数の線形制約の検定を行う場合に F 検定が用いられる。

10.1 いくつかの例

例 1：コブ＝ダグラス型生産関数： Q_i は生産量，K_i は資本，L_i は労働とする。生産関数を推定する。

$$\log(Q_i) = \beta_1' + \beta_2 \log(K_i) + \beta_3 \log(L_i) + u_i,$$

において，一次同時の制約 $\beta_2 + \beta_3 = 1$ を検定したい。すなわち，帰無仮説，対立仮説は以下のように表される。

帰無仮説 H_0： $\beta_2 + \beta_3 = 1$,
対立仮説 H_1： $\beta_2 + \beta_3 \neq 1$,

例 2：構造変化の検定： n_0 期以前と $n_0 + 1$ 期以降とで経済構造が変化したと考えて推定を行う。しかも，定数項，傾き共に変化したと想定した場合，回帰式は以下のようになる。

$$Y_i = \alpha + \beta X_i + \gamma D_i + \delta D_i X_i + u_i,$$

ただし，

$$D_i = \begin{cases} 0, & i = 1, 2, \cdots, n_0 \text{ のとき}, \\ 1, & i = n_0 + 1, n_0 + 2, \cdots, n \text{ のとき}, \end{cases}$$

とする。構造変化が $n_0 + 1$ 期で起こったかどうかを検定したい。すなわち，帰無仮説，対立仮説は以下のように表される。

帰無仮説 H_0： $\gamma = \delta = 0$,
対立仮説 H_1： $\gamma \neq 0$，または，$\delta \neq 0$,

例 3：多重回帰モデルの係数の同時検定： 2 つの説明変数が含まれる場合を考える。

$$Y_i = \alpha + \beta X_i + \gamma Z_i + u_i,$$

のモデルにおいて，X_i と Z_i のどちらも，Y_i に影響を与えていないという仮説を検定したい。この場合，帰無仮説，対立仮説は以下のように表される。

帰無仮説 H_0： $\beta = \gamma = 0$,
対立仮説 H_1： $\beta \neq 0$，または，$\gamma \neq 0$,

10.2 統計学の復習

$U \sim \chi^2(n)$，$V \sim \chi^2(m)$，U と V は独立とする。
このとき，

$$F = \frac{U/n}{V/m} \sim F(n, m)$$

となる。

10.3 検定の方法

多重回帰モデル

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + u_i,$$

において，パラメータ $\beta_1, \beta_2, \cdots, \beta_k$ に何らかの制約が妥当かどうかを検定する。
制約の数を G 個とする。
全く制約の無い場合に得られた残差を $\widehat{u}_i$ とする。
制約を含めて推定されたときの残差を $\widetilde{u}_i$ とする。
すなわち，

帰無仮説 H_0： $\beta_{k-G+1} = \cdots = \beta_k = 0$,
対立仮説 H_1：H_0 でない。

を検定する場合，

$$\begin{aligned} Y_i = {} & \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_{k-G} X_{k-G,i} \\ & + \beta_{k-G+1} X_{k-G+1,i} + \cdots + \beta_k X_{ki} + u_i, \end{aligned}$$

の推定によって得られた残差を $\widehat{u}_i$ (制約なし残差) とおき，

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_{k-G} X_{k-G,i} + u_i,$$

の推定によって得られた残差を $\widetilde{u}_i$ (制約付き残差) とする。

1. H_0 が真のとき，$\dfrac{\sum \widetilde{u}_i^2 - \sum \widehat{u}_i^2}{\sigma^2} \sim \chi^2(G)$ となる。(証明略)

2. また，$\dfrac{\sum \widehat{u}_i^2}{\sigma^2} \sim \chi^2(n-k)$ となる。(証明略)

3. さらに，$\dfrac{\sum \widetilde{u}_i^2 - \sum \widehat{u}_i^2}{\sigma^2}$ と $\dfrac{\sum \widehat{u}_i^2}{\sigma^2}$ とは独立に分布する。(証明略)

4. したがって，この場合，

$$\frac{(\sum \widetilde{u}_i^2 - \sum \widehat{u}_i^2)/G}{\sum \widehat{u}_i^2/(n-k)} \sim F(G, n-k),$$

となる。(証明略)

例 1： コブ＝ダグラス型生産関数：
制約なしの場合：

$$\log(Q_i) = \beta_1' + \beta_2 \log(K_i) + \beta_3 \log(L_i) + u_i,$$

$\beta_2 + \beta_3 = 1$ の制約ありの場合：

$$\log(\frac{Q_i}{L_i}) = \beta_1' + \beta_2 \log(\frac{K_i}{L_i}) + u_i,$$

例 2： 構造変化の検定：
制約なしの場合：

$$Y_i = \alpha + \beta X_i + \gamma D_i + \delta D_i X_i + u_i,$$

$\gamma = \delta = 0$ の制約ありの場合：

$$Y_i = \alpha + \beta X_i + u_i,$$

例 3： 多重回帰モデルの係数の同時検定：
制約なしの場合：

$$Y_i = \alpha + \beta X_i + \gamma Z_i + u_i,$$

$\beta = \gamma = 0$ の制約ありの場合：

$$Y_i = \alpha + u_i,$$

11 応用例

11.1 マクロの消費関数

1. 所得から税金を差し引いたものを可処分所得と呼ぶが，可処分所得 (Y) が増えれば消費 (C) も増える。

2. この関数を $C = \alpha + \beta Y$ という線形 (一次式) によって表されると仮定しよう。

3. この場合，経済学では，α は基礎消費，β は限界消費性向と呼ばれる。

4. α で表される基礎消費とは所得がなくても日常生活に最低限必要な消費 (すなわち，衣食住宅費等) であり，β の限界消費性向とは所得が 1 円増えれば消費はいくら増えるのかという指標である。

5. α, β はパラメータと呼ばれ，未知である。

6. C や Y は『国民経済計算年報』(経済企画庁編) から「国内家計最終消費支出」「家計国民可処分所得」という項目で，それぞれデータは公表される。

7. ここでは，平成 10 年版の『国民経済計算年報』のデータを扱う。

1. 『国民経済計算年報』から「国内家計最終消費支出」と「家計国民可処分所得」の 1970 年～1996 年の年次データ (時系列データの種類は年次データ，四半期データ，月次データ等がある) を取ってくる。

2. 計量分析で重要なことは，**名目値**でなく**実質値**をとるということである。

3. 実質値とはある基準となる年を定めて，その年の物価で生のデータ (名目値) を変換するということである。

4. 実質値 $= \dfrac{\text{名目値}}{\text{物価指数}}$ という関係が成り立つ。(ここでの「物価指数」は基準年次を 1 とした場合のもので，もし基準年次が 100 ならば，実質値 $= \dfrac{\text{名目値}}{\text{物価指数}/100}$ とする必要がある。)

5. 異なる時点間でデータを比較する場合，それぞれの時点で物価が異なるので，物価の変動を取り除いたデータで比較する必要がある。

6. ここで用いられる国内家計最終消費支出 C_i と家計国民可処分所得 Y_i のデータは名目データであり，実質データに変換する必要がある。

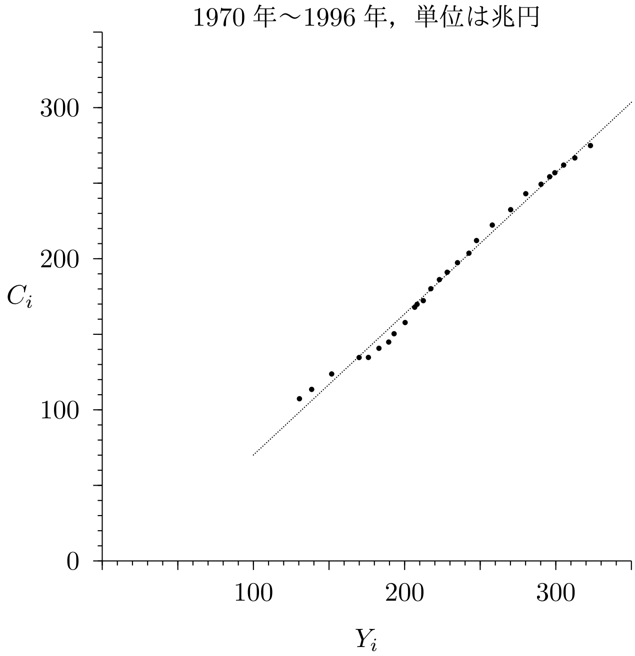
7. 1990 年の「国内家計最終消費支出デフレータ」は 100 なので，基準年次は 1990 年となる。

表 3: 所得と消費のデータ

暦年	国内家計最終支出	家計可処分所得	国内家計最終支出デレータ
1970	37784.1	45913.2	35.2
1971	42571.6	51944.3	37.5
1972	49124.1	60245.4	39.7
1973	59366.1	74924.8	44.1
1974	71782.1	93833.2	53.3
1975	83591.1	108712.8	59.4
1976	94443.7	123540.9	65.2
1977	105397.8	135318.4	70.1
1978	115960.3	147244.2	73.5
1979	127600.9	157071.1	76.0
1980	138585.0	169931.5	81.6
1981	147103.4	181349.2	85.4
1982	157994.0	190611.5	87.7
1983	166631.6	199587.8	89.5
1984	175383.4	209451.9	91.8
1985	185335.1	220655.6	93.9
1986	193069.6	229938.8	94.8
1987	202072.8	235924.0	95.3
1988	212939.9	247159.7	95.8
1989	227122.2	263940.5	97.7
1990	243035.7	280133.0	100.0
1991	255531.8	297512.9	102.5
1992	265701.6	309256.6	104.5
1993	272075.3	317021.6	105.9
1994	279538.7	325655.7	106.7
1995	283245.4	331967.5	106.2
1996	291374.8	342303.0	106.0

図 1: 実質消費 (縦軸) と実質所得 (横軸)

1970 年〜1996 年，単位は兆円



8. 1990 年の貨幣価値に変換することを，1990 年価格で実質化すると言う。

まず最初に，

$$C_i = \alpha + \beta Y_i + u_i, \qquad u_i \sim N(0, \sigma^2),$$

の推定を行う。ただし，u_i は互いに独立に正規分布するものと仮定する。

$$C_i = \underset{(3844.54)}{-23216.7} + \underset{(.016333)}{.933542}\ Y_i,$$

$$R^2 = .992406, \quad \overline{R}^2 = .992102,$$

$$s = 4557.04, \quad DW = .289838,$$

ただし，係数の推定値の下の括弧内は，係数推定値の標準誤差を表すものとする。

1. 推定値の符号条件について，基礎消費の推定値 $\widehat{\alpha}$ は -23216.7 で負，限界消費性向の推定値 $\widehat{\beta}$ は 0.933542 で正となっている。基礎消費については符号条件を満たさないが，限界消費性向は符号条件を満たす。

2. $\widehat{\alpha}, \widehat{\beta}$ から，α, β の符号を統計的に調べる。

   (a) データ数は 27，推定すべきパラメータ数は 2 なので，自由度は $27 - 2 = 25$ となる。

(b) $\alpha = 0.05$ のとき $t_{\alpha/2}(25) = 2.060$，$\alpha = 0.01$ のとき $t_{\alpha/2}(25) = 2.787$ である。

(c) 有意水準 0.05 のとき，$H_0:\ \alpha = 0$，$H_1:\ \alpha \neq 0$ の結果は，

$$\frac{-23216.7}{3844.54} = -6.039 < -t_{0.025}(25) = -2.060,$$

有意水準 0.05 のとき，$H_0:\ \beta = 0$，$H_1:\ \beta \neq 0$ の結果は，

$$\frac{.933542}{.016333} = 57.16 > t_{0.025}(25) = 2.060,$$

となり，共に統計的に有意である。

(d) したがって，実証結果から，真の基礎消費 α は負，真の限界消費性向 β は正という結論になる。

(e) $\beta > 0$ を統計的に示したが，本当は $\beta < 1$ も示すべきである。すなわち，$H_0:\ \beta = 1$，$H_1:\ \beta \neq 1$ も検定すべきである。(省略)

(f) 基礎消費は正となるべきなので，経済理論と矛盾する。

(g) 次に行うべき分析は，なぜ矛盾したかを追求すること。

$\Longrightarrow$ 最初に考えた理論が間違っていた，構造変化のためだった，推定式が最小二乗法の仮定を満たしていなかった，$\cdots\cdots$

3. $s = 4557.04$ は，誤差項 u_i の標準偏差 σ^2 の推定値 (すなわち，回帰の標準誤差) である。

4. 自由度修正済み決定係数は $\overline{R}^2 = 0.992102$ であり，非常に 1 に近い値が得られたことから，消費と所得の間の関係を表す回帰式の当てはまりは非常に良いと言える。

5. DW について，$n = 27$, $k = 2$ の 5% 点の値は $dl = 1.32$, $du = 1.47$ であるので，5% で，

(a) $DW < 1.32$ のとき，誤差項に正の系列相関がある。

(b) $1.32 \leq DW < 1.47$ のとき，判定不能。

(c) $1.47 \leq DW < 2.53$ のとき，誤差項に系列相関はない。

(d) $2.53 \leq DW < 2.68$ のとき，判定不能。

(e) $2.68 \leq DW$ のとき，誤差項に負の系列相関がある。

となる。

この場合，$DW = 0.289838$ なので，誤差項に正の系列相関が見られる。

$\Longrightarrow$ 最小二乗法の仮定を満たしていない。

$\Longrightarrow$ $t(25)$ 分布を検定に使うことができないので，今までの検定結果は間違っている可能性がある。

次に，誤差項に系列相関 (一階の自己相関) があるモデル

$$C_i = \alpha + \beta Y_i + u_i,$$
$$u_i = \rho u_{i-1} + \epsilon_i,$$

の推定を行う。ただし，ϵ_i は互いに独立に正規分布するものと仮定する。

ρ の推定値 $\widehat{\rho}$ は，最小二乗法の推定結果の DW を用いて，$\widehat{\rho} = 1 - \dfrac{DW}{2} = 1 - \dfrac{.2898375}{2} = .855081$ を得る。データを，次のように変換する。

$$C_i^* = C_i - \widehat{\rho} C_{i-1},$$
$$Y_i^* = Y_i - \widehat{\rho} Y_{i-1},$$

そして，

$$C_i^* = \alpha' + \beta Y_i^* + \epsilon_i, \qquad \epsilon_i \sim N(0, \sigma_\epsilon^2),$$

を推定する。ただし，$\alpha' = \alpha(1-\rho)$ に注意。

結果は，以下の通りとなる。

$$C_i^* = \underset{(2183.63)}{-3679.78} + \underset{(.054012)}{.930350}\ Y_i^*,$$
$$R^2 = .922288, \quad \overline{R}^2 = .919180,$$
$$s = 2232.08, \quad DW = 1.35684,$$

ただし，係数の推定値の下の括弧内は，係数推定値の標準誤差を表すものとする。

1. この場合，データ数は 26，推定すべきパラメータ数は 2 なので，自由度は $26 - 2 = 24$ となる。

2. 基礎消費の推定値 $\widehat{\alpha}$ は，$\widehat{\alpha} = \dfrac{\widehat{\alpha}'}{1-\widehat{\rho}}$ によって求められ，$\dfrac{-3679.78}{1-.855081} = -25392.0$ で負，限界消費性向の推定値 $\widehat{\beta}$ は .930350 で正となる。

3. 真の基礎消費の符号を統計的に調べる。基礎消費 α の符号は α' の符号と同じなので，α' の符号を $\widehat{\alpha}'$ から調べる。

有意水準 0.05 のとき，

$$\frac{-3679.78}{2183.63} = -1.685 > -t_{0.025}(24) = -2.064,$$

となり，α' が負だとは言えない。

4. α は正の可能性もある。
5. したがって，推定結果から，最初に考えた理論モデルが間違っているとは言えない。
6. β について，有意水準 0.05 のとき，

   $$\frac{.930350}{.054012} = 17.225 > t_{0.025}(24) = 2.064,$$

   となり，統計的に有意である。
7. $s = 2232.08$ は，誤差項 ϵ_i の標準偏差 σ_ϵ^2 の推定値 (すなわち，回帰の標準誤差) である。
8. 自由度修正済み決定係数 $\overline{R}^2$ について，最初の推定結果では 0.992102 となり，今回の推定結果では .919180 と小さくなっている。

   $\Longrightarrow$ 「回帰式の当てはまりが悪くなった」と考えるのは間違い。

   被説明変数の値自体が，最初の推定結果と今回のものとでは異なるため，$\overline{R}^2$ の比較は不適当。

   $\Longrightarrow$ 回帰の標準誤差 (4557.04 と 2232.08) で比較すべき。後者の方が小さいため，後者の方が回帰式の当てはまりは良いと言える。
9. DW も 1.35684 となり，$1.32 \le DW < 1.47$ であるので，系列相関の有無を判定できない。

   $\Longrightarrow$ 「明らかに誤差項に系列相関がある」とは言えない。

   $\Longrightarrow$ この式で，信頼区間，仮説検定を行うべき。

● Stata による出力結果

データは，

http://www2.econ.osaka-u.ac.jp/~tanizaki/class/2011/econome/cons.csv

からダウンロード可

```
------------------------------------------------------------


. tsset year
        time variable:  year, 1970 to 1996
                delta:  1 unit

. gen ryd=yd/(pcons/100)

. gen rcons=cons/(pcons/100)

. reg rcons ryd

      Source |       SS       df       MS              Number of obs =      27
-------------+------------------------------           F(  1,    25) = 3267.05
       Model |  6.7845e+10     1  6.7845e+10           Prob > F      =  0.0000
    Residual |   519164248    25  20766569.9           R-squared     =  0.9924
-------------+------------------------------           Adj R-squared =  0.9921
       Total |  6.8365e+10    26  2.6294e+09           Root MSE      =    4557

------------------------------------------------------------------------------
       rcons |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         ryd |   .9335422   .0163326    57.16   0.000     .8999045    .9671799
       _cons |  -23216.75   3844.539    -6.04   0.000    -31134.72   -15298.77
------------------------------------------------------------------------------

. dwstat

Durbin-Watson d-statistic(  2,    27) =  .2898375

. gen rho=1-0.5*.2898375

. gen drcons=rcons-rho*l.rcons
(1 missing value generated)

. gen dryd=ryd-rho*l.ryd
(1 missing value generated)

. reg drcons dryd

      Source |       SS       df       MS              Number of obs =      26
-------------+------------------------------           F(  1,    24) =  261.29
       Model |  1.3558e+09     1  1.3558e+09           Prob > F      =  0.0000
    Residual |   124530197    24  5188758.21           R-squared     =  0.9159
-------------+------------------------------           Adj R-squared =  0.9124
       Total |  1.4803e+09    25    59212957           Root MSE      =  2277.9

------------------------------------------------------------------------------
      drcons |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        dryd |   .9315055   .0576262    16.16   0.000     .8125708     1.05044
       _cons |  -3731.757   2353.261    -1.59   0.126    -8588.649    1125.134
------------------------------------------------------------------------------
```

途中で，

$\mathtt{DRCONS}_i = \mathtt{RCONS}_i - \widehat{\rho}\,\mathtt{RCONS}_{i-1}$

$\mathtt{DRY}_i = \mathtt{RY}_i - \widehat{\rho}\,\mathtt{RY}_{i-1}$

として，データをの変換 (ただし，$\widehat{\rho} = 1 - .5DW = 1 - .5 \times .2898375$)

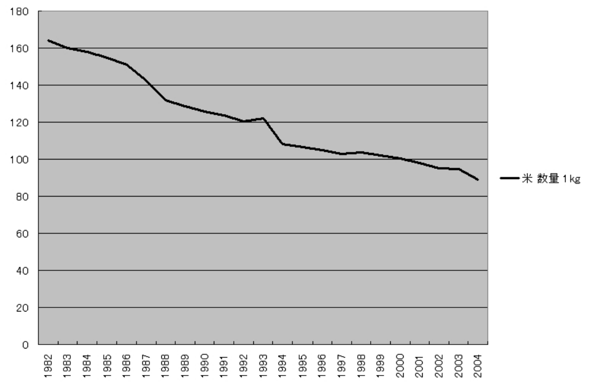
11.2 ミクロの消費関数（需要関数）

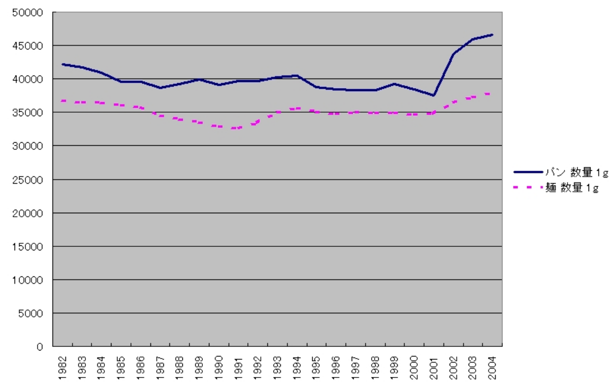
米，パン，麺の選択を考える。

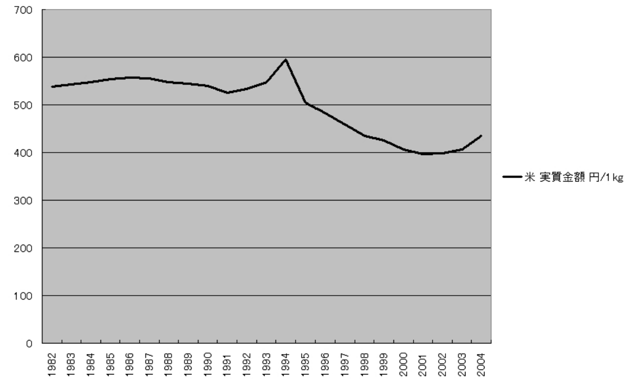
1 世帯当たり年間の品目別支出金額，購入数量及び平均価格（全世帯・勤労者世帯）

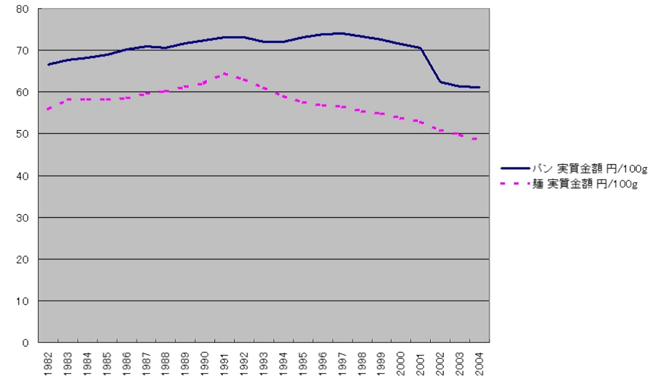
年	米		パン		麺		総支出額	消費者物価指数
	数量	金額	数量	金額	数量	金額		
	1kg	1kg	1g	100g	1g	100g		(全国総合)
i	Q_{1i}	P_{1i}	Q_{2i}	P_{2i}	Q_{3i}	P_{3i}	E_i	P_i
1982	164.22	435.98	42161	54.04	36854	45.32	3038024	0.811
1983	160.14	448.20	41745	55.88	36492	47.97	3114247	0.825
1984	158.06	461.69	40890	57.62	36500	49.21	3195829	0.844
1985	154.51	477.41	39545	59.42	36099	50.20	3277373	0.861
1986	150.96	482.80	39532	60.86	35859	50.74	3316493	0.867
1987	142.60	482.67	38710	61.53	34576	51.83	3371326	0.868
1988	132.04	478.40	39218	61.75	33971	52.65	3493468	0.874
1989	128.40	486.37	39927	63.99	33603	54.71	3592205	0.893
1990	125.78	497.33	39157	66.71	32890	57.14	3734084	0.921
1991	123.82	499.36	39659	69.57	32615	61.44	3925358	0.951
1992	120.58	516.05	39697	70.75	33401	61.06	4003931	0.967
1993	121.93	536.85	40209	70.51	35085	59.80	4022955	0.979
1994	107.99	587.50	40458	71.08	35760	58.37	4006086	0.986
1995	106.42	496.64	38766	71.97	35096	56.77	3948741	0.985
1996	104.91	476.26	38436	72.74	34804	55.90	3946187	0.986
1997	102.81	460.70	38333	74.39	35061	56.77	3999759	1.004
1998	103.53	439.24	38287	74.10	34956	55.98	3938235	1.010
1999	101.99	427.60	39246	73.00	34963	55.37	3876091	1.007
2000	100.40	406.82	38480	71.47	34722	53.83	3805600	1.000
2001	97.83	394.67	37554	70.12	34753	52.52	3704298	0.993
2002	95.15	391.28	43727	61.34	36493	50.11	3673550	0.984
2003	94.83	398.37	45876	60.12	37302	48.93	3631473	0.981
2004	89.02	426.12	46653	59.92	37957	47.72	3650436	0.981

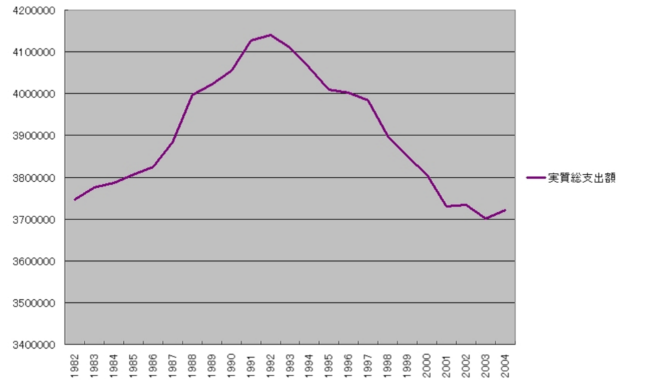
出所：『消費者物価指数年報 (平成 16 年)』(日本銀行) ← P_i
『家計調査年報 (平成 16 年)』(総務省統計局) ← その他











●線型関数

$$Q_{1i} = \underset{(8.89)}{397.0} - \underset{(9.35)}{0.00018}\frac{E_i}{P_i} + \underset{(4.39)}{0.194}\frac{P_{1i}}{P_i} + \underset{(0.29)}{0.184}\frac{P_{2i}}{P_i} + \underset{(5.57)}{5.614}\frac{P_{3i}}{P_i}$$

$$s = 6.86, \quad R^2 = 0.932, \quad \overline{R}^2 = 0.917, \quad DW = 1.532$$

$$Q_{2i} = \underset{(13.9)}{61946.7} + \underset{(4.62)}{0.00893}\frac{E_i}{P_i} + \underset{(0.91)}{4.014}\frac{P_{1i}}{P_i} - \underset{(11.4)}{710.4}\frac{P_{2i}}{P_i} - \underset{(1.51)}{152.5}\frac{P_{3i}}{P_i}$$

$$s = 685.1, \quad R^2 = 0.931, \quad \overline{R}^2 = 0.915, \quad DW = 1.334$$

$$\begin{aligned} Q_{3i} = &\ \underset{(16.9)}{56642.5} - \underset{(0.67)}{0.00098}\ \frac{E_i}{P_i} + \underset{(5.40)}{17.9}\ \frac{P_{1i}}{P_i} \\ &- \underset{(1.04)}{48.8}\ \frac{P_{2i}}{P_i} - \underset{(5.32)}{403.6}\ \frac{P_{3i}}{P_i} \end{aligned}$$

$$s = 515.7, \quad R^2 = 0.885, \quad \overline{R}^2 = 0.860, \quad DW = 1.044$$

●対数線型

$$\begin{aligned} \log Q_{1i} = &\ \underset{(10.4)}{70.9} - \underset{(10.4)}{5.35}\ \log\frac{E_i}{P_i} + \underset{(4.53)}{0.705}\ \log\frac{P_{1i}}{P_i} \\ &- \underset{(0.06)}{0.018}\ \log\frac{P_{2i}}{P_i} + \underset{(6.46)}{2.68}\ \log\frac{P_{3i}}{P_i} \end{aligned}$$

$$s = 0.048, \quad R^2 = 0.949, \quad \overline{R}^2 = 0.937, \quad DW = 1.713$$

$$\begin{aligned} \log Q_{2i} = &\ \underset{(1.88)}{4.35} + \underset{(4.33)}{0.753}\ \log\frac{E_i}{P_i} + \underset{(1.21)}{0.063}\ \log\frac{P_{1i}}{P_i} \\ &- \underset{(10.9)}{1.135}\ \log\frac{P_{2i}}{P_i} - \underset{(1.31)}{0.184}\ \log\frac{P_{3i}}{P_i} \end{aligned}$$

$$s = 0.016, \quad R^2 = 0.931, \quad \overline{R}^2 = 0.916, \quad DW = 1.317$$

$$\begin{aligned} \log Q_{3i} = &\ \underset{(6.87)}{14.5} - \underset{(1.19)}{0.188}\ \log\frac{E_i}{P_i} + \underset{(5.51)}{0.266}\ \log\frac{P_{1i}}{P_i} \\ &- \underset{(0.27)}{0.025}\ \log\frac{P_{2i}}{P_i} - \underset{(5.30)}{0.682}\ \log\frac{P_{3i}}{P_i} \end{aligned}$$

$$s = 0.015, \quad R^2 = 0.883, \quad \overline{R}^2 = 0.856, \quad DW = 1.067$$

●対数線型 (変数を減らす)

$$\begin{aligned} \log Q_{1i} = &\ \underset{(10.9)}{70.9} - \underset{(11.2)}{5.36}\ \log\frac{E_i}{P_i} + \underset{(5.23)}{0.709}\ \log\frac{P_{1i}}{P_i} \\ &+ \underset{(7.69)}{2.67}\ \log\frac{P_{3i}}{P_i} \end{aligned}$$

$$s = 0.047, \quad R^2 = 0.949, \quad \overline{R}^2 = 0.941, \quad DW = 1.724$$

$$\begin{aligned} \log Q_{3i} = &\ \underset{(7.24)}{14.7} - \underset{(1.35)}{0.200}\ \log\frac{E_i}{P_i} + \underset{(6.45)}{0.271}\ \log\frac{P_{1i}}{P_i} \\ &- \underset{(6.49)}{0.700}\ \log\frac{P_{3i}}{P_i} \end{aligned}$$

$$s = 0.014, \quad R^2 = 0.882, \quad \overline{R}^2 = 0.863, \quad DW = 1.122$$

● Stata による出力結果

データは,

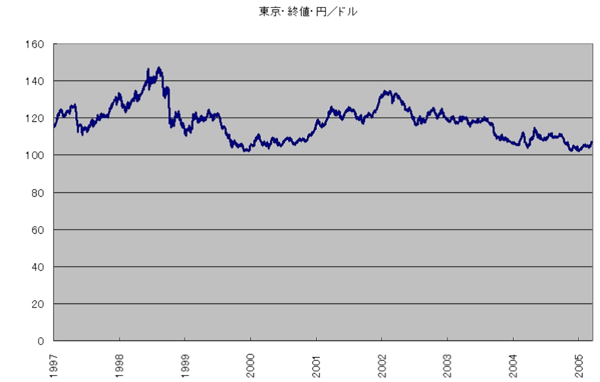
`http://www2.econ.osaka-u.ac.jp/~tanizaki/class/2011/econome/demand.csv`

からダウンロード可

```
-----------------------------------------------------------


. tsset year
        time variable:  year, 1982 to 2004
                delta:  1 unit

. gen re=e/p

. gen rp1=p1/p

. gen rp2=p2/p

. gen rp3=p3/p

. reg q1 re rp1 rp2 rp3

      Source |       SS       df       MS              Number of obs =      23
-------------+------------------------------           F(  4,    18) =   62.32
       Model |  11605.1437     4  2901.28592           Prob > F      =  0.0000
    Residual |  837.923294    18  46.5512941           R-squared     =  0.9327
-------------+------------------------------           Adj R-squared =  0.9177
       Total |   12443.067    22  565.593953           Root MSE      =  6.8229

------------------------------------------------------------------------------
          q1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          re |  -.0001813   .0000192    -9.43   0.000    -.0002217   -.0001409
         rp1 |   .1935249    .043943     4.40   0.000      .101204    .2858457
         rp2 |   .1825145    .621349     0.29   0.772    -1.122891     1.48792
         rp3 |   5.623917   1.003036     5.61   0.000     3.516618    7.731217
       _cons |   397.3071   44.31083     8.97   0.000     304.2135    490.4007
------------------------------------------------------------------------------

. dwstat

Durbin-Watson d-statistic(  5,    23) =  1.538178

. reg q2 re rp1 rp2 rp3

      Source |       SS       df       MS              Number of obs =      23
-------------+------------------------------           F(  4,    18) =   60.62
       Model |   113479303     4  28369825.8           Prob > F      =  0.0000
    Residual |  8424108.48    18  468006.027           R-squared     =  0.9309
-------------+------------------------------           Adj R-squared =  0.9155
       Total |   121903412    22  5541064.17           Root MSE      =  684.11

------------------------------------------------------------------------------
          q2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          re |   .0089294   .0019288     4.63   0.000      .004877    .0129817
         rp1 |   3.975941   4.406052     0.90   0.379     -5.28083    13.23271
         rp2 |  -710.4687   62.30106   -11.40   0.000    -841.3584   -579.5791
         rp3 |  -152.1237   100.5718    -1.51   0.148    -363.4172    59.16979
       _cons |   61930.62   4442.933    13.94   0.000     52596.36    71264.87
------------------------------------------------------------------------------

. dwstat

Durbin-Watson d-statistic(  5,    23) =  1.340373

. reg q3 re rp1 rp2 rp3

      Source |       SS       df       MS              Number of obs =      23
-------------+------------------------------           F(  4,    18) =   34.82
       Model |  37006848.3     4  9251712.07           Prob > F      =  0.0000
    Residual |  4782689.62    18  265704.979           R-squared     =  0.8856
-------------+------------------------------           Adj R-squared =  0.8601
       Total |  41789537.9    22  1899524.45           Root MSE      =  515.47

------------------------------------------------------------------------------
          q3 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          re |    -.00097   .0014534    -0.67   0.513    -.0040234    .0020834
         rp1 |   17.89258    3.31989     5.39   0.000     10.91775    24.86741
         rp2 |  -48.87487   46.94286    -1.04   0.312    -147.4982    49.74843
         rp3 |  -403.4164   75.77925    -5.32   0.000    -562.6227   -244.2101
       _cons |   56606.16   3347.679    16.91   0.000     49572.95    63639.37
------------------------------------------------------------------------------

. dwstat

Durbin-Watson d-statistic(  5,    23) =  1.047824

. gen lq1=log(q1)

. gen lq2=log(q2)

. gen lq3=log(q3)

. gen lre=log(re)

. gen lrp1=log(rp1)

. gen lrp2=log(rp2)

. gen lrp3=log(rp3)
```

```
. reg lq1 lre lrp1 lrp2 lrp3

      Source |       SS       df       MS              Number of obs =      23
-------------+------------------------------           F(  4,    18) =   84.42
       Model |  .766574852     4  .191643713           Prob > F      =  0.0000
    Residual |   .04086074    18  .002270041           R-squared     =  0.9494
-------------+------------------------------           Adj R-squared =  0.9381
       Total |  .807435592    22  .036701618           Root MSE      =  .04764

------------------------------------------------------------------------------
         lq1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         lre |  -5.355623   .5092073   -10.52   0.000    -6.425428   -4.285818
        lrp1 |   .7030891   .1547336     4.54   0.000     .3780058    1.028172
        lrp2 |    -.01919   .3056177    -0.06   0.951    -.6612689     .622889
        lrp3 |   2.687511   .4126367     6.51   0.000     1.820594    3.554429
       _cons |   70.90628    6.78413    10.45   0.000     56.65335    85.15921
------------------------------------------------------------------------------

. dwstat

Durbin-Watson d-statistic(  5,    23) =  1.720916

. reg lq2 lre lrp1 lrp2 lrp3

      Source |       SS       df       MS              Number of obs =      23
-------------+------------------------------           F(  4,    18) =   61.41
       Model |  .064720875     4  .016180219           Prob > F      =  0.0000
    Residual |  .004742568    18  .000263476           R-squared     =  0.9317
-------------+------------------------------           Adj R-squared =  0.9166
       Total |  .069463442    22  .003157429           Root MSE      =  .01623

------------------------------------------------------------------------------
         lq2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         lre |   .7530252   .1734795     4.34   0.000     .3885582    1.117492
        lrp1 |   .0633603   .0527155     1.20   0.245    -.0473909    .1741114
        lrp2 |  -1.135159   .1041195   -10.90   0.000    -1.353906   -.9164119
        lrp3 |  -.1838809   .1405793    -1.31   0.207    -.4792271    .1114654
       _cons |   4.346289   2.311255     1.88   0.076    -.5094776    9.202055
------------------------------------------------------------------------------

. dwstat

Durbin-Watson d-statistic(  5,    23) =  1.323931

. reg lq3 lre lrp1 lrp2 lrp3

      Source |       SS       df       MS              Number of obs =      23
-------------+------------------------------           F(  4,    18) =   33.83
       Model |  .029851208     4  .007462802           Prob > F      =  0.0000
    Residual |  .003971008    18  .000220612           R-squared     =  0.8826
-------------+------------------------------           Adj R-squared =  0.8565
       Total |  .033822216    22  .001537373           Root MSE      =  .01485

------------------------------------------------------------------------------
         lq3 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         lre |  -.1868154    .158742    -1.18   0.255      -.52032    .1466892
        lrp1 |   .2652044   .0482372     5.50   0.000     .1638618     .366547
        lrp2 |  -.0255097   .0952743    -0.27   0.792    -.2256736    .1746542
        lrp3 |  -.6819342   .1286368    -5.30   0.000      -.95219   -.4116784
       _cons |   14.52537   2.114908     6.87   0.000     10.08212    18.96863
------------------------------------------------------------------------------

. dwstat

Durbin-Watson d-statistic(  5,    23) =  1.071649

. reg lq1 lre lrp1 lrp3

      Source |       SS       df       MS              Number of obs =      23
-------------+------------------------------           F(  3,    19) =  118.79
       Model |  .766565902     3  .255521967           Prob > F      =  0.0000
    Residual |  .040869691    19  .002151036           R-squared     =  0.9494
-------------+------------------------------           Adj R-squared =  0.9414
       Total |  .807435592    22  .036701618           Root MSE      =  .04638

------------------------------------------------------------------------------
         lq1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         lre |  -5.364744   .4750821   -11.29   0.000    -6.359103   -4.370386
        lrp1 |   .7074307   .1347484     5.25   0.000      .425399    .9894623
        lrp3 |    2.67427    .345262     7.75   0.000     1.951629    3.396912
       _cons |   70.98983    6.47565    10.96   0.000     57.43613    84.54352
------------------------------------------------------------------------------

. dwstat

Durbin-Watson d-statistic(  4,    23) =  1.733124

. reg lq3 lre lrp1 lrp3

      Source |       SS       df       MS              Number of obs =      23
-------------+------------------------------           F(  3,    19) =   47.40
       Model |  .029835392     3  .009945131           Prob > F      =  0.0000
    Residual |  .003986824    19  .000209833           R-squared     =  0.8821
-------------+------------------------------           Adj R-squared =  0.8635
       Total |  .033822216    22  .001537373           Root MSE      =  .01449

------------------------------------------------------------------------------
         lq3 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         lre |  -.1989407   .1483821    -1.34   0.196     -.509508    .1116267
        lrp1 |   .2709758   .0420859     6.44   0.000      .182889    .3590625
        lrp3 |  -.6995357   .1078355    -6.49   0.000     -.925238   -.4738335
       _cons |   14.63644   2.022536     7.24   0.000     10.40322    18.86965
------------------------------------------------------------------------------

. dwstat

Durbin-Watson d-statistic(  4,    23) =  1.127541
```

11.3　株価，金利，為替レート



$$\mathrm{Kabu}_i = -\underset{(2.58)}{1883.1} + \underset{(8.27)}{46.3}\ \mathrm{ExRate}_i + \underset{(51.7)}{6802.6}\ R_i$$

$s = 2309.9, \quad R^2 = 0.570, \quad \overline{R}^2 = 0.570, \quad DW = 0.021$

$$\mathrm{Kabu}_i \;=\; -\underset{(0.03)}{2.36} \;+\; \underset{(0.31)}{0.164}\ \mathrm{ExRate}_i \;+\; \underset{(1.96)}{36.5}\ R_i \;+\; \underset{(482)}{0.995}\ \mathrm{Kabu}_{i-1}$$

$s = 214.8, \quad R^2 = 0.996, \quad \overline{R}^2 = 0.996, \quad DW = 2.100$

$$\Delta\mathrm{Kabu}_i = \underset{(0.119)}{8.11} \;-\; \underset{(0.173)}{0.091}\ \mathrm{ExRate}_i \;-\; \underset{(0.070)}{0.862}\ R_i$$

$s = 215.1, \quad R^2 = 0.000015, \quad \overline{R}^2 = -0.00097, \quad DW = 2.106$

● Stata による出力結果

データは,

http://www2.econ.osaka-u.ac.jp/~tanizaki/class/2011/econome/nikkei.csv

からダウンロード可

```
--------------------------------------------------------


. gen time=_n

. tsset time
        time variable:  time, 1 to 2028
                delta:  1 unit

. reg kabu exrate r

      Source |       SS       df       MS              Number of obs =    2028
-------------+------------------------------           F(  2,  2025) = 1341.74
       Model |  1.4318e+10     2  7.1590e+09           Prob > F      =  0.0000
    Residual |  1.0805e+10  2025  5335630.99           R-squared     =  0.5699
-------------+------------------------------           Adj R-squared =  0.5695
       Total |  2.5123e+10  2027  12394036.8           Root MSE      =  2309.9

------------------------------------------------------------------------------
        kabu |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      exrate |   46.30522   5.601817     8.27   0.000      35.3193    57.29115
           r |    6802.61   131.6017    51.69   0.000     6544.521    7060.699
       _cons |  -1883.071    730.508    -2.58   0.010    -3315.696   -450.4449
------------------------------------------------------------------------------

. dwstat

Durbin-Watson d-statistic(  3,  2028) =   .021141

. reg kabu exrate r l.kabu

      Source |       SS       df       MS              Number of obs =    2027
-------------+------------------------------           F(  3,  2023) =       .
       Model |  2.4997e+10     3  8.3323e+09           Prob > F      =  0.0000
    Residual |  93298782.6  2023  46119.0225           R-squared     =  0.9963
-------------+------------------------------           Adj R-squared =  0.9963
       Total |  2.5090e+10  2026  12384078.2           Root MSE      =  214.75

------------------------------------------------------------------------------
        kabu |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      exrate |   .1643596   .5295945     0.31   0.756    -.8742481    1.202967
           r |   36.51834    18.6521     1.96   0.050     -.060989    73.09767
             |
        kabu |
         L1. |   .9945113    .002064   481.84   0.000     .9904636     .998559
             |
       _cons |  -2.359124   68.05383    -0.03   0.972     -135.822    131.1038
------------------------------------------------------------------------------

. dwstat

Durbin-Watson d-statistic(  4,  2027) =  2.100248

. reg dkabu exrate r if tin(2,2028)

      Source |       SS       df       MS              Number of obs =    2027
-------------+------------------------------           F(  2,  2024) =    0.02
       Model |  1443.23541     2  721.617703           Prob > F      =  0.9845
    Residual |  93624927.9  2024  46257.3754           R-squared     =  0.0000
-------------+------------------------------           Adj R-squared = -0.0010
       Total |  93626371.1  2026   46212.424           Root MSE      =  215.08

------------------------------------------------------------------------------
       dkabu |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      exrate |  -.0906022   .5216244    -0.17   0.862    -1.113579    .9323746
           r |  -.8624553   12.27849    -0.07   0.944    -24.94225    23.21734
       _cons |   8.109931    68.0417     0.12   0.905    -125.3291     141.549
------------------------------------------------------------------------------

. dwstat

Durbin-Watson d-statistic(  3,  2027) =  2.106308
```

12 推定量の求め方

12.1 最小二乗法

・n 個のデータ (実現値)： $x_1, x_2, \cdots, x_n$
・背後に対応する確率変数を仮定： $X_1, X_2, \cdots, X_n$
・$\mathrm{E}(X_i) = \mu$，$\mathrm{V}(X_i) = \sigma^2$ を仮定
母数 (μ, σ^2) を推定する。

観測データ $x_1, x_2, \cdots, x_n$ をもとにして，μ の最小二乗推定値を求める。

$$\min_{\mu} \sum_{i=1}^{n} (x_i - \mu)^2$$

μ の解を $\widehat{\mu}$ とすると，

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

となり，$\widehat{\mu} \equiv \overline{x}$ を得る。

すなわち，

$$\frac{\mathrm{d} \sum_{i=1}^{n} (x_i - \mu)^2}{\mathrm{d}\mu} = 0$$

を μ について解く。

μ の最小二乗推定量はデータ x_i を対応する確率変数 X_i で置き換えて，

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

となり，$\widehat{\mu} \equiv \overline{X}$ を得る ($\widehat{\mu}$ について，推定値と推定量は同じ記号を使っている)。

以上を回帰分析に応用すると，

$$\min_{\alpha,\beta}\sum_{i=1}^{n}(Y_i-\alpha-\beta X_i)^2$$

を解くことになる。

すなわち，

$$\frac{\partial\sum_{i=1}^{n}(Y_i-\alpha-\beta X_i)^2}{\partial\alpha}=0$$
$$\frac{\partial\sum_{i=1}^{n}(Y_i-\alpha-\beta X_i)^2}{\partial\beta}=0$$

の連立方程式を α, β について解く。

12.2 最尤法

n 個の確率変数 $X_1, X_2, \cdots, X_n$ は互いに独立で，同じ確率分布 $f(x)\equiv f(x;\theta)$ とする。ただし，θ は母数で，例えば，$\theta=(\mu,\sigma^2)$ である。

$X_1, X_2, \cdots, X_n$ の結合分布は，互いに独立なので，

$$f(x_1,x_2,\cdots,x_n;\theta)\equiv\prod_{i=1}^{n}f(x_i;\theta)$$

と表される。

観測データ $x_1, x_2, \cdots, x_n$ を与えたもとで，$\prod_{i=1}^{n}f(x_i;\theta)$ は θ の関数として表される。すなわち，

$$l(\theta)=\prod_{i=1}^{n}f(x_i;\theta)$$

となる。
$l(\theta)$ を尤度関数と呼ぶ。

$$\max_{\theta}\ l(\theta)$$

となる θ を最尤推定値 $\widehat{\theta}=\widehat{\theta}(x_1,x_2,\cdots,x_n)$ と呼ぶ。

データ $x_1, x_2, \cdots, x_n$ を確率変数 $X_1, X_2, \cdots, X_n$ で置き換えて，$\widehat{\theta}=\widehat{\theta}(X_1,X_2,\cdots,X_n)$ を最尤推定量と呼ぶ。

$$\max_{\theta}\ l(\theta)$$

と

$$\max_{\theta}\ \log l(\theta)$$

の θ の解はともに同じものであることに注意。$\log l(\theta)$ を対数尤度関数と呼ぶ。

最尤推定量の性質： n が大きいとき，

$$\widehat{\theta}\sim N(\theta,\sigma_\theta^2)$$

ただし，

$$\sigma_\theta^2=\frac{1}{\sum_{i=1}^{n}\mathrm{E}\Big[\Big(\frac{\mathrm{d}\log f(X_i;\theta)}{\mathrm{d}\theta}\Big)^2\Big]}=-\frac{1}{\sum_{i=1}^{n}\mathrm{E}\Big[\frac{\mathrm{d}^2\log f(X_i;\theta)}{\mathrm{d}\theta^2}\Big]}$$

θ がベクトル $(k\times 1)$ の場合，n が大きいとき，

$$\widehat{\theta}\sim N(\theta,\Sigma_\theta)$$

ただし，

$$\Sigma_\theta=\Big(\sum_{i=1}^{n}\mathrm{E}\Big[\Big(\frac{\partial\log f(X_i;\theta)}{\partial\theta}\Big)\Big(\frac{\partial\log f(X_i;\theta)}{\partial\theta}\Big)'\Big]\Big)^{-1}=-\Big(\sum_{i=1}^{n}\mathrm{E}\Big[\frac{\partial^2\log f(X_i;\theta)}{\partial\theta\partial\theta'}\Big]\Big)^{-1}$$

例 1： 正規母集団 $\mathrm{N}(\mu,\sigma^2)$ からの標本値 $x_1, x_2, \cdots, x_n$ を用いて，
(1) σ^2 が既知のとき，μ の最尤推定値と最尤推定量
(2) σ^2 が未知のとき，μ と σ^2 の最尤推定値と最尤推定量
をそれぞれ求める。

［解］$\mathrm{N}(\mu,\sigma^2)$ の密度関数は，

$$f(x;\mu,\sigma^2)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\Big(-\frac{1}{2\sigma^2}(x-\mu)^2\Big)$$

となる。したがって，互いに独立な $X_1, X_2, \cdots, X_n$ の結合分布は，

$$f(x_1,x_2,\cdots,x_n;\mu,\sigma^2)\equiv\prod_{i=1}^{n}f(x_i;\mu,\sigma^2)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\Big(-\frac{1}{2\sigma^2}(x_i-\mu)^2\Big)=(2\pi\sigma^2)^{-\frac{n}{2}}\exp\Big(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\Big)$$

となる。

(1) σ^2 が既知のとき，尤度関数 $l(\mu)$ は，

$$l(\mu)=(2\pi\sigma^2)^{-\frac{n}{2}}\exp\Big(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\Big)$$

となる。
$l(\mu)$ を最大にする μ と $\log l(\mu)$ を最大にする μ は同じになる。

したがって，対数尤度関数は，

$$\log l(\mu) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

となり，

$$\frac{\mathrm{d}\log l(\mu)}{\mathrm{d}\mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) = 0$$

となる μ を求める。μ の解を $\widehat{\mu}$ とすると，μ の最尤推定値は，

$$\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n}x_i \equiv \overline{x}$$

を得る。

さらに，観測値 x_1, x_2, $\cdots$, x_n をその確率変数 X_1, X_2, $\cdots$, X_n で置き換えて，μ の最尤推定量は，

$$\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n}X_i \equiv \overline{X}$$

となる。
$\widehat{\mu}$ の分散を求めるために，

$$\log f(X_i;\mu) = -\frac{1}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}(X_i-\mu)^2$$

$$\frac{\mathrm{d}\log f(X_i;\mu)}{\mathrm{d}\mu} = \frac{1}{\sigma^2}(X_i-\mu)$$

$$\Big(\frac{\mathrm{d}\log f(X_i;\mu)}{\mathrm{d}\mu}\Big)^2 = \frac{1}{\sigma^4}(X_i-\mu)^2$$

$$\mathrm{E}\Big[\Big(\frac{\mathrm{d}\log f(X_i;\mu)}{\mathrm{d}\mu}\Big)^2\Big] = \frac{1}{\sigma^4}\mathrm{E}[(X_i-\mu)^2] = \frac{1}{\sigma^2}$$

と計算される。

最尤推定量の性質から，n が大きいとき，

$$\widehat{\mu} \sim N(\mu, \sigma_\mu^2)$$

ただし，

$$\sigma_\mu^2 = \frac{1}{\sum_{i=1}^{n}\mathrm{E}\Big[\Big(\frac{\mathrm{d}\log f(X_i;\mu)}{\mathrm{d}\mu}\Big)^2\Big]} = \frac{\sigma^2}{n}$$

この場合は，n の大きさに関わらず，$\widehat{\mu} \sim N(\mu, \sigma_\mu^2)$ が成り立つ。

(2) σ^2 が未知のとき，μ と σ^2 の尤度関数は，

$$l(\mu,\sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}}\exp\Big(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\Big)$$

となる。
対数尤度関数は，

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

と表される。
μ と σ^2 について，最大化するためには，

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

$$\frac{\partial\log l(\mu,\sigma^2)}{\partial\sigma^2} = -\frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2 = 0$$

の連立方程式を解く。

μ, σ^2 の解を $\widehat{\mu}$, $\widehat{\sigma}^2$ とすると，最尤推定値は，

$$\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n}x_i \equiv \overline{x}$$

$$\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\widehat{\mu}) \equiv \frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})$$

となる。

観測値 x_1, x_2, $\cdots$, x_n をその確率変数 X_1, X_2, $\cdots$, X_n で置き換えて，μ, σ^2 の最尤推定量は，

$$\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n}X_i \equiv \overline{X}$$

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

となる。

σ^2 の最尤推定量 $\widehat{\sigma}^2$ は，σ^2 の不偏推定量 $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$ とは異なることに注意。

$\theta = (\mu, \sigma^2)'$ とする。n が大きいとき，

$$\widehat{\theta} \sim N(\theta, \Sigma_\theta)$$

ただし，

$$\Sigma_\theta = -\Big(\sum_{i=1}^n \mathrm{E}\Big[\frac{\partial^2 \log f(X_i;\theta)}{\partial\theta\partial\theta'}\Big]\Big)^{-1}$$

$$\log f(X_i;\theta) = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}(X_i-\mu)^2$$

$$\begin{aligned}\frac{\partial \log f(X_i;\theta)}{\partial\theta} &= \begin{pmatrix} \dfrac{\partial \log f(X_i;\theta)}{\partial\mu} \\ \dfrac{\partial \log f(X_i;\theta)}{\partial\sigma^2} \end{pmatrix} \\ &= \begin{pmatrix} \dfrac{1}{\sigma^2}(X_i-\mu) \\ -\dfrac{1}{2\sigma^2} + \dfrac{1}{2\sigma^4}(X_i-\mu)^2 \end{pmatrix}\end{aligned}$$

$$\begin{aligned}&\frac{\partial^2 \log f(X_i;\theta)}{\partial\theta\partial\theta'} \\ &= \begin{pmatrix} \dfrac{\partial^2 \log f(X_i;\theta)}{\partial\mu^2} & \dfrac{\partial^2 \log f(X_i;\theta)}{\partial\mu\partial\sigma^2} \\ \dfrac{\partial^2 \log f(X_i;\theta)}{\partial\sigma^2\partial\mu} & \dfrac{\partial^2 \log f(X_i;\theta)}{\partial(\sigma^2)^2} \end{pmatrix} \\ &= \begin{pmatrix} -\dfrac{1}{\sigma^2} & -\dfrac{1}{\sigma^4}(X_i-\mu) \\ -\dfrac{1}{\sigma^4}(X_i-\mu) & \dfrac{1}{2\sigma^4} - \dfrac{1}{\sigma^6}(X_i-\mu)^2 \end{pmatrix}\end{aligned}$$

$$\begin{aligned}&\mathrm{E}\Big[\frac{\partial^2 \log f(X_i;\theta)}{\partial\theta\partial\theta'}\Big] \\ &= \begin{pmatrix} -\dfrac{1}{\sigma^2} & -\dfrac{1}{\sigma^4}\mathrm{E}(X_i-\mu) \\ -\dfrac{1}{\sigma^4}\mathrm{E}(X_i-\mu) & \dfrac{1}{2\sigma^4} - \dfrac{1}{\sigma^6}\mathrm{E}[(X_i-\mu)^2] \end{pmatrix} \\ &= \begin{pmatrix} -\dfrac{1}{\sigma^2} & 0 \\ 0 & -\dfrac{1}{2\sigma^4} \end{pmatrix}\end{aligned}$$

よって，

$$\begin{aligned}\Sigma_\theta &= -\Big(\sum_{i=1}^n \mathrm{E}\Big[\frac{\partial^2 \log f(X_i;\theta)}{\partial\theta\partial\theta'}\Big]\Big)^{-1} \\ &= \begin{pmatrix} \dfrac{\sigma^2}{n} & 0 \\ 0 & \dfrac{2\sigma^4}{n} \end{pmatrix}\end{aligned}$$

まとめると，μ，σ^2 の最尤推定量 $\widehat{\mu} = (1/n)\sum_{i=1}^n X_i$，$\widehat{\sigma}^2 = (1/n)\sum_{i=1}^n (X_i - \overline{X})^2$ の分布は，n が大きいとき，

$$\begin{pmatrix} \widehat{\mu} \\ \widehat{\sigma}^2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix}, \begin{pmatrix} \dfrac{\sigma^2}{n} & 0 \\ 0 & \dfrac{2\sigma^4}{n} \end{pmatrix}\right)$$

となる。

例 2：　X_1, X_n, $\cdots$, X_n は互いに独立で，それぞれパラメータ p を持ったベルヌイ分布に従うものとする。すなわち，X_i の確率関数は，

$$f(x;p) = p^x(1-p)^{1-x} \qquad\qquad x = 0, 1$$

となる。

このとき尤度関数は，

$$l(p) = \prod_{i=1}^n f(x_i;p) = \prod_{i=1}^n p^{x_i}(1-p)^{1-x_i}$$

となり，対数尤度関数は，

$$\begin{aligned}\log l(p) &= \sum_{i=1}^n \log f(x_i;p) \\ &= \log(p)\sum_{i=1}^n x_i + \log(1-p)\sum_{i=1}^n (1-x_i) \\ &= \log(p)\sum_{i=1}^n x_i + \log(1-p)(n - \sum_{i=1}^n x_i)\end{aligned}$$

となる。
$\log l(p)$ を最大にする p を求める。

$$\frac{\mathrm{d}\log l(p)}{\mathrm{d}p} = \frac{1}{p}\sum_{i=1}^n x_i - \frac{1}{1-p}(n - \sum_{i=1}^n x_i) = 0$$

したがって，p について解くと，p の最尤推定値 $\widehat{p}$ は，

$$\widehat{p} = \frac{1}{n}\sum_{i=1}^n x_i$$

となる。

さらに，x_i を X_i で置き換えて，p の最尤推定量 $\widehat{p}$ は，

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

となる。

$\widehat{p}$ の分布を求める。

$$\log f(X_i;p) = X_i \log(p) + (1-X_i)\log(1-p)$$
$$\frac{\mathrm{d}\log f(X_i;p)}{\mathrm{d}p} = \frac{X_i}{p} - \frac{1-X_i}{1-p} = \frac{X_i - p}{p(1-p)}$$
$$\mathrm{E}\Big[\Big(\frac{\mathrm{d}\log f(X_i;p)}{\mathrm{d}p}\Big)^2\Big] = \frac{\mathrm{E}[(X_i-p)^2]}{p^2(1-p)^2}$$

$$\begin{aligned}\mathrm{E}[(X_i-p)^2] &= \sum_{x_i=0}^{1}(x_i-p)^2 f(x_i;p)\\ &= \sum_{x_i=0}^{1}(x_i-p)^2 p^{x_i}(1-p)^{1-x_i}\\ &= p^2(1-p) + (1-p)^2 p = p(1-p)\end{aligned}$$
$$\frac{1}{\sum_{i=1}^n \mathrm{E}\Big[\Big(\frac{\mathrm{d}\log f(X_i;p)}{\mathrm{d}p}\Big)^2\Big]} = \frac{p(1-p)}{n}$$

したがって，

$$\widehat{p} \sim N\big(p,\ \frac{p(1-p)}{n}\big)$$

を得る。

例 3： $X_1,\, X_n,\, \cdots,\, X_n$ は互いに独立で，それぞれパラメータ λ を持ったポアソン分布に従うものとする。すなわち，X_i の確率関数は，

$$f(x;\lambda) = \frac{\lambda^x e^{-\lambda}}{x!} \qquad\qquad x = 0,1,2,\cdots$$

となる。

このとき尤度関数は，

$$l(\lambda) = \prod_{i=1}^n f(x_i;\lambda) = \prod_{i=1}^n \frac{\lambda^{x_i}e^{-\lambda}}{x_i!}$$

となり，対数尤度関数は，

$$\begin{aligned}\log l(\lambda) &= \sum_{i=1}^n \log f(x_i;\lambda)\\ &= \log(\lambda)\sum_{i=1}^n x_i - n\lambda - \sum_{i=1}^n \log(x_i!)\end{aligned}$$

となる。

$\log l(\lambda)$ を最大にする p を求める。

$$\frac{\mathrm{d}\log l(\lambda)}{\mathrm{d}\lambda} = \frac{1}{\lambda}\sum_{i=1}^n x_i - n = 0$$

したがって，λ について解くと，λ の最尤推定値 $\widehat{\lambda}$ は，

$$\widehat{\lambda} = \frac{1}{n}\sum_{i=1}^n x_i$$

となる。

さらに，x_i を X_i で置き換えて，λ の最尤推定量 $\widehat{\lambda}$ は，

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

となる。

$\widehat{\lambda}$ の分布を求める。

$$\log f(X_i;\lambda) = X_i\log(\lambda) - \lambda - \log(X_i!)$$
$$\frac{\mathrm{d}\log f(X_i;\lambda)}{\mathrm{d}\lambda} = \frac{X_i}{\lambda} - 1$$
$$\frac{\mathrm{d}^2\log f(X_i;\lambda)}{\mathrm{d}\lambda^2} = -\frac{X_i}{\lambda^2}$$
$$\mathrm{E}\Big(\frac{\mathrm{d}^2\log f(X_i;\lambda)}{\mathrm{d}\lambda^2}\Big) = \frac{\mathrm{E}(X_i)}{\lambda^2}$$

$$\begin{aligned}\mathrm{E}(X_i) &= \sum_{x=0}^{\infty} x f(x;\lambda)\\ &= \sum_{x=0}^{\infty} x\frac{\lambda^x e^{-\lambda}}{x!}\\ &= \sum_{x=1}^{\infty} x\frac{\lambda^x e^{-\lambda}}{x!}\\ &= \sum_{x=1}^{\infty} \lambda\frac{\lambda^{x-1} e^{-\lambda}}{(x-1)!}\\ &= \sum_{x=0}^{\infty} \lambda\frac{\lambda^x e^{-\lambda}}{x!}\\ &= \lambda\end{aligned}$$
$$-\frac{1}{\sum_{i=1}^n \mathrm{E}\Big(\frac{\mathrm{d}^2\log f(X_i;\lambda)}{\mathrm{d}\lambda^2}\Big)} = \frac{\lambda}{n}$$

したがって，

$$\widehat{\lambda} \sim N\big(\lambda,\ \frac{\lambda}{n}\big)$$

を得る。

例 4： $X_1, X_n, \cdots, X_n$ は互いに独立で，それぞれパラメータ λ を持った指数分布に従うものとする。すなわち，X_i の密度関数は，

$$f(x;\lambda) = \lambda e^{-\lambda x} \qquad x > 0$$

となる。

このとき尤度関数は，

$$l(\lambda) = \prod_{i=1}^{n} f(x_i;\lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i}$$

となり，対数尤度関数は，

$$\begin{aligned}\log l(\lambda) &= \sum_{i=1}^{n} \log f(x_i;\lambda) \\ &= n\log\lambda - \lambda\sum_{i=1}^{n} x_i\end{aligned}$$

となる。
$\log l(\lambda)$ を最大にする p を求める。

$$\frac{\mathrm{d}\log l(\lambda)}{\mathrm{d}\lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i = 0$$

したがって，λ について解くと，λ の最尤推定値 $\widehat{\lambda}$ は，

$$\widehat{\lambda} = \frac{n}{\sum_{i=1}^{n} x_i}$$

となる。

さらに，x_i を X_i で置き換えて，λ の最尤推定量 $\widehat{\lambda}$ は，

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

となる。

$\widehat{\lambda}$ の分布を求める。

$$\log f(X_i;\lambda) = \log\lambda - \lambda X_i$$

$$\frac{\mathrm{d}\log f(X_i;\lambda)}{\mathrm{d}\lambda} = \frac{1}{\lambda} - X_i$$

$$\frac{\mathrm{d}^2\log f(X_i;\lambda)}{\mathrm{d}\lambda^2} = -\frac{1}{\lambda^2}$$

$$-\frac{1}{\sum_{i=1}^{n} \mathrm{E}\Big(\dfrac{\mathrm{d}^2\log f(X_i;\lambda)}{\mathrm{d}\lambda^2}\Big)} = \frac{\lambda^2}{n}$$

したがって，

$$\widehat{\lambda} \sim N(\lambda,\ \frac{\lambda^2}{n})$$

を得る。

12.2.1　変数変換

確率変数 X の密度関数を $f(x)$，分布関数を $F(x) \equiv \mathrm{P}(X < x)$ とする。$Y = aX + b$ とするとき，Y の密度関数 $g(y)$ を求める。
Y の分布関数を $G(y)$ として，次のように変形できる。

$$\begin{aligned}G(y) &= \mathrm{P}(Y < y) = \mathrm{P}(aX + b < y) \\ &= \begin{cases} \mathrm{P}\Big(X < \dfrac{y-b}{a}\Big), & a > 0 \text{ のとき} \\ \mathrm{P}\Big(X > \dfrac{y-b}{a}\Big), & a < 0 \text{ のとき} \end{cases} \\ &= \begin{cases} \mathrm{P}\Big(X < \dfrac{y-b}{a}\Big), & a > 0 \text{ のとき} \\ 1 - \mathrm{P}\Big(X < \dfrac{y-b}{a}\Big), & a < 0 \text{ のとき} \end{cases} \\ &= \begin{cases} F\Big(\dfrac{y-b}{a}\Big), & a > 0 \text{ のとき} \\ 1 - F\Big(\dfrac{y-b}{a}\Big), & a < 0 \text{ のとき} \end{cases}\end{aligned}$$

分布関数と密度関数との関係は，

$$\frac{\mathrm{d}F(x)}{\mathrm{d}x} = f(x) \qquad \frac{\mathrm{d}G(x)}{\mathrm{d}x} = g(x)$$

であるので，Y の密度関数は，

$$\begin{aligned}g(y) &= \frac{\mathrm{d}G(y)}{\mathrm{d}y} \\ &= \begin{cases} \dfrac{1}{a} f\Big(\dfrac{y-b}{a}\Big), & a > 0 \text{ のとき} \\ -\dfrac{1}{a} f\Big(\dfrac{y-b}{a}\Big), & a < 0 \text{ のとき} \end{cases} \\ &= \Big|\frac{1}{a}\Big| f\Big(\frac{y-b}{a}\Big)\end{aligned}$$

と表される。

一般に，確率変数 X の密度関数を $f(x)$ とする。単調変換 $X = h(Y)$ とするとき，Y の密度関数 $g(y)$ は，

$$g(y) = |h'(y)| f\big(h(y)\big)$$

となる。

12.2.2　回帰分析への応用

回帰モデル

$$Y_i = \alpha + \beta X_i + u_i \qquad i = 1, 2, \cdots, n$$

$u_1, u_2, \cdots, u_n$ は互いに独立で，すべての i について $u_i \sim \mathrm{N}(0, \sigma^2)$ を仮定する。

u_i の密度関数は，

$$f(u_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\Big(-\frac{1}{2\sigma^2}u_i^2\Big)$$

となる。

Y_i の密度関数 $g(Y_i)$ は，

$$g(Y_i) = |h'(Y_i)| f\big(h(Y_i)\big)$$

によって求められる。

この場合，$h(Y_i) = Y_i - \alpha - \beta X_i$ なので，$h'(Y_i) = 1$ となる。

したがって，Y_i の密度関数は，

$$\begin{aligned} g(Y_i) &= |h'(Y_i)| f\big(h(Y_i)\big) \\ &= f\big(h(Y_i)\big) \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\Big(-\frac{1}{2\sigma^2}(Y_i - \alpha - \beta X_i)^2\Big) \end{aligned}$$

となる。

u_1, u_2, $\cdots$, u_n は互いに独立であれば，Y_1, Y_2, $\cdots$, Y_n も互いに独立になるので，Y_1, Y_2, $\cdots$, Y_n の結合密度関数は，

$$\begin{aligned} g(Y_1, Y_2, \cdots, Y_n) &= \prod_{i=1}^n g(Y_i) \\ &= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\Big(-\frac{1}{2\sigma^2}\sum_{i=1}^n (Y_i - \alpha - \beta X_i)^2\Big) \end{aligned}$$

となる。これは α, β, σ^2 の関数となっている。

よって，尤度関数は，

$$l(\alpha, \beta, \sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\Big(-\frac{1}{2\sigma^2}\sum_{i=1}^n (Y_i - \alpha - \beta X_i)^2\Big)$$

となる。

対数尤度関数は，

$$\begin{aligned} \log l(\alpha, \beta, \sigma^2) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) \\ -\frac{1}{2\sigma^2}\sum_{i=1}^n (Y_i - \alpha - \beta X_i)^2 \end{aligned}$$

となる。

$\log l(\alpha, \beta, \sigma^2)$ を最大にするために，

$$\begin{aligned} \frac{\partial \log l(\alpha, \beta, \sigma^2)}{\partial \alpha} &= \frac{1}{\sigma^2}\sum_{i=1}^n (Y_i - \alpha - \beta X_i) = 0 \\ \frac{\partial \log l(\alpha, \beta, \sigma^2)}{\partial \beta} &= \frac{1}{\sigma^2}\sum_{i=1}^n X_i(Y_i - \alpha - \beta X_i) = 0 \\ \frac{\partial \log l(\alpha, \beta, \sigma^2)}{\partial \sigma^2} &= -\frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^n (Y_i - \alpha - \beta X_i)^2 = 0 \end{aligned}$$

の連立方程式を解く。

上 2 つの式は σ^2 に依存していない。α，β の最尤推定量は最小二乗推定量と同じになる。

すなわち，

$$\begin{aligned} \widehat{\beta} &= \frac{\sum_{i=1}^n (X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^n (X_i - \overline{X})^2} \\ \widehat{\alpha} &= \overline{Y} - \widehat{\beta}\overline{X} \end{aligned}$$

σ^2 の最尤推定量は，

$$\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (Y_i - \widehat{\alpha} - \widehat{\beta}X_i)^2$$

となり，s^2 とは異なる。

$\widehat{\theta} = (\widehat{\alpha}, \widehat{\beta}, \widehat{\sigma}^2)'$，$\theta = (\alpha, \beta, \sigma^2)'$ とする。n が大きいとき，

$$\widehat{\theta} \sim N(\theta, \Sigma_\theta)$$

ただし，

$$\begin{aligned} \Sigma_\theta &= \Big(\sum_{i=1}^n \mathrm{E}\Big[\Big(\frac{\partial \log g(Y_i;\theta)}{\partial \theta}\Big)\Big(\frac{\partial \log g(Y_i;\theta)}{\partial \theta}\Big)'\Big]\Big)^{-1} \\ &= -\Big(\sum_{i=1}^n \mathrm{E}\Big[\frac{\partial^2 \log g(Y_i;\theta)}{\partial \theta \partial \theta'}\Big]\Big)^{-1} \end{aligned}$$

Y_i の密度関数 $g(Y_i;\theta)$ の対数は，

$$\begin{aligned} \log g(Y_i;\theta) = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\sigma^2) \\ -\frac{1}{2\sigma^2}(Y_i - \alpha - \beta X_i)^2 \end{aligned}$$

となる。

$$\frac{\partial \log g(Y_i;\theta)}{\partial \theta} = \begin{pmatrix} \dfrac{\partial \log g(Y_i;\theta)}{\partial \alpha} \\ \dfrac{\partial \log g(Y_i;\theta)}{\partial \beta} \\ \dfrac{\partial \log g(Y_i;\theta)}{\partial \sigma^2} \end{pmatrix}$$

$$
= \begin{pmatrix} \frac{1}{\sigma^2}(Y_i - \alpha - \beta X_i) \\ \frac{1}{\sigma^2} X_i (Y_i - \alpha - \beta X_i) \\ -\frac{1}{2\sigma^2} + \frac{1}{2\sigma^4}(Y_i - \alpha - \beta X_i)^2 \end{pmatrix}
$$

$$
\frac{\partial^2 \log g(Y_i;\theta)}{\partial\theta\partial\theta'}
$$

$$
= \begin{pmatrix} \frac{\partial^2 \log g(Y_i;\theta)}{\partial\alpha^2} & \frac{\partial^2 \log g(Y_i;\theta)}{\partial\alpha\partial\beta} & \frac{\partial^2 \log g(Y_i;\theta)}{\partial\alpha\partial\sigma^2} \\ \frac{\partial^2 \log g(Y_i;\theta)}{\partial\beta\partial\alpha} & \frac{\partial^2 \log g(Y_i;\theta)}{\partial\beta^2} & \frac{\partial^2 \log g(Y_i;\theta)}{\partial\beta\partial\sigma^2} \\ \frac{\partial^2 \log g(Y_i;\theta)}{\partial\sigma^2\partial\alpha} & \frac{\partial^2 \log g(Y_i;\theta)}{\partial\sigma^2\partial\beta} & \frac{\partial^2 \log g(Y_i;\theta)}{\partial(\sigma^2)^2} \end{pmatrix}
$$

$$
= \begin{pmatrix} -\frac{1}{\sigma^2} & -\frac{X_i}{\sigma^2} & -\frac{u_i}{\sigma^4} \\ -\frac{X_i}{\sigma^2} & -\frac{X_i^2}{\sigma^2} & -\frac{X_i u_i}{\sigma^4} \\ -\frac{u_i}{\sigma^4} & -\frac{X_i u_i}{\sigma^4} & \frac{1}{2\sigma^4} - \frac{u_i^2}{\sigma^6} \end{pmatrix}
$$

ただし，$u_i = Y_i - \alpha - \beta X_i$
期待値をとると，

$$
\mathrm{E}\Big(\frac{\partial^2 \log g(Y_i;\theta)}{\partial\theta\partial\theta'}\Big) = \mathrm{E}\begin{pmatrix} -\frac{1}{\sigma^2} & -\frac{X_i}{\sigma^2} & -\frac{u_i}{\sigma^4} \\ -\frac{X_i}{\sigma^2} & -\frac{X_i^2}{\sigma^2} & -\frac{X_i u_i}{\sigma^4} \\ -\frac{u_i}{\sigma^4} & -\frac{X_i u_i}{\sigma^4} & \frac{1}{2\sigma^4} - \frac{u_i^2}{\sigma^6} \end{pmatrix}
= \begin{pmatrix} -\frac{1}{\sigma^2} & -\frac{X_i}{\sigma^2} & 0 \\ -\frac{X_i}{\sigma^2} & -\frac{X_i^2}{\sigma^2} & 0 \\ 0 & 0 & -\frac{1}{2\sigma^4} \end{pmatrix}
$$

となる。

$$
\Sigma_\theta = -\Big(\sum_{i=1}^n \mathrm{E}\Big[\frac{\partial^2 \log g(Y_i;\theta)}{\partial\theta\partial\theta'}\Big]\Big)^{-1}
= \begin{pmatrix} \frac{n}{\sigma^2} & \frac{\sum_{i=1}^n X_i}{\sigma^2} & 0 \\ \frac{\sum_{i=1}^n X_i}{\sigma^2} & \frac{\sum_{i=1}^n X_i^2}{\sigma^2} & 0 \\ 0 & 0 & \frac{n}{2\sigma^4} \end{pmatrix}^{-1}
= \begin{pmatrix} \sigma^2 \begin{pmatrix} n & \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i^2 \end{pmatrix}^{-1} & 0 \\ 0 & \frac{2\sigma^4}{n} \end{pmatrix}
$$

したがって，

$$
\begin{pmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{pmatrix} \sim N\Big(\begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \sigma^2 \begin{pmatrix} n & \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i^2 \end{pmatrix}^{-1}\Big)
$$

となる。
$\longrightarrow$ 最小二乗推定量の分布と同じ。

12.2.3 誤差項に系列相関がある場合

回帰モデル

$$
Y_i = \alpha + \beta X_i + u_i
$$
$$
u_i = \rho u_{i-1} + \epsilon_i \qquad i = 2, 3, \cdots, n
$$

ϵ_2, ϵ_3, $\cdots$, ϵ_n は互いに独立で，すべての i について $\epsilon_i \sim \mathrm{N}(0, \sigma^2)$ を仮定する。

u_i を消去すると，

$$
(Y_i - \alpha - \beta X_i) = \rho(Y_{i-1} - \alpha - \beta X_{i-1}) + \epsilon_i
$$

または

$$
(Y_i - \rho Y_{i-1}) = \alpha(1-\rho) + \beta(X_i - \rho X_{i-1}) + \epsilon_i
$$

と書き直すことが出来る。

$\theta = (\alpha, \beta, \sigma^2, \rho)$ とする。

$$
\log f(Y_i;\theta) = -\frac{1}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\Big((Y_i - \rho Y_{i-1}) - \alpha(1-\rho) - \beta(X_i - \rho X_{i-1})\Big)^2
$$

尤度関数は，

$$
\log l(\theta) = \sum_{i=2}^n \log f(Y_i;\theta) = -\frac{n-1}{2}\log(2\pi) - \frac{n-1}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=2}^n \Big((Y_i - \rho Y_{i-1}) - \alpha(1-\rho) - \beta(X_i - \rho X_{i-1})\Big)^2
$$

となる。
尤度関数をそれぞれ α，β，σ^2，ρ について微分し，ゼロとおく。

$$
\frac{\partial \log l(\theta)}{\partial\alpha} = \frac{1-\rho}{\sigma^2}\sum_{i=2}^n \Big((Y_i - \rho Y_{i-1}) - \alpha(1-\rho) - \beta(X_i - \rho X_{i-1})\Big) = 0
$$

$$
\frac{\partial \log l(\theta)}{\partial\beta} = \frac{1}{\sigma^2}\sum_{i=2}^n (X_i - \rho X_{i-1})\Big((Y_i - \rho Y_{i-1}) - \alpha(1-\rho) - \beta(X_i - \rho X_{i-1})\Big) = 0
$$

$$\frac{\partial \log l(\theta)}{\partial \sigma^2} = -\frac{n-1}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=2}^{n}\Big((Y_i - \rho Y_{i-1}) \\ -\alpha(1-\rho) - \beta(X_i - \rho X_{i-1})\Big)^2 = 0$$

$$\frac{\partial \log l(\theta)}{\partial \rho} = \frac{1}{\sigma^2}\sum_{i=2}^{n}(Y_{i-1} - \alpha - \beta X_{i-1})\Big((Y_i - \alpha - \beta X_i) \\ -\rho(Y_{i-1} - \alpha - \beta X_{i-1})\Big) = 0$$

$(Y_i - \alpha - \beta X_i) - \rho(Y_{i-1} - \alpha - \beta X_{i-1})$ は
$(Y_i - \rho Y_{i-1}) - \alpha(1-\rho) - \beta(X_i - \rho X_{i-1})$ を書き直したもの。

4 つの連立方程式を解いて，最尤推定量 $\widehat{\alpha}$，$\widehat{\beta}$，$\widehat{\sigma}^2$，$\widehat{\rho}$ が得られる。
$\longrightarrow$ 下記のように収束計算によって求める。

(i) 初期段階では，$\widehat{\rho} = 0$ とする。

(ii) $X_i^* = X_i - \widehat{\rho}X_{i-1}$

$Y_i^* = Y_i - \widehat{\rho}Y_{i-1}$

(iii) $\begin{pmatrix} \widetilde{\alpha} \\ \widehat{\beta} \end{pmatrix} = \begin{pmatrix} n-1 & \sum_{i=2}^n X_i^* \\ \sum_{i=2}^n X_i^* & \sum_{i=2}^n X_i^{*2} \end{pmatrix}^{-1} \begin{pmatrix} \sum_{i=2}^n Y_i^* \\ \sum_{i=2}^n X_i^* Y_i^* \end{pmatrix}$

(iv) $\widehat{\alpha} = \dfrac{\widetilde{\alpha}}{1-\widehat{\rho}}$

(v) $\widehat{u}_i = Y_i - \widehat{\alpha} - \widehat{\beta}X_i$

(vi) $\widehat{\sigma}^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=2}^{n}(\widehat{u}_i - \widehat{\rho}\widehat{u}_{i-1})^2$

(vii) $\widehat{\rho} = \dfrac{\sum_{i=2}^n \widehat{u}_i\widehat{u}_{i-1}}{\sum_{i=2}^n \widehat{u}_{i-1}^2}$

(viii) ステップ (ii) ～ (vii) を，収束するまで繰り返し計算する。

12.3 尤度比検定

n 個の確率変数 $X_1, X_2, \cdots, X_n$ は互いに独立で，同じ確率分布 $f(x) \equiv f(x;\theta)$ とする。
尤度関数は，

$$l(\theta) = \prod_{i=1}^{n} f(x_i;\theta)$$

となる。
θ の制約つき最尤推定量を $\widetilde{\theta}$，制約無し最尤推定量を $\widehat{\theta}$ とする。
制約の数を G 個とする。

$\dfrac{l(\widetilde{\theta})}{l(\widehat{\theta})}$ を尤度比と呼ぶ

検定方法 1： 尤度比がある値より小さいときに，帰無仮説を棄却する。すなわち，

$$\frac{l(\widetilde{\theta})}{l(\widehat{\theta})} < c$$

となるときに，帰無仮説を棄却する。この場合，c を次のようにして求める必要がある。

$$\int \cdots \int \prod_{i=1}^{n} f(x_i;\widetilde{\theta})\mathrm{d}x_1 \cdots \mathrm{d}x_n = \alpha$$

ただし，α は有意水準（帰無仮説が正しいときに，帰無仮説を棄却する確率）を表す。

検定方法 2（大標本検定）： または，$n \longrightarrow \infty$ のとき，

$$-2\log\frac{l(\widetilde{\theta})}{l(\widehat{\theta})} \longrightarrow \chi^2(G)$$

となる。
この検定を尤度比検定と呼ぶ。

例 1： 正規母集団 $\mathrm{N}(\mu, \sigma^2)$ からの標本値 x_1，x_2，$\cdots$，x_n を用いて，σ^2 が既知のとき，帰無仮説 H_0： $\mu = \mu_0$，H_1： $\mu \neq \mu_0$ の尤度比検定を行う。

σ^2 が既知のとき，尤度関数 $l(\mu)$ は，

$$l(\mu) = (2\pi\sigma^2)^{-\frac{n}{2}}\exp\Big(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\Big)$$

となる。
$l(\mu)$ を最大にする μ と $\log l(\mu)$ を最大にする μ は同じになる。

μ の最尤推定量は，

$$\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i \equiv \overline{X}$$

となる。

尤度比検定統計量は，

$$\frac{l(\mu_0)}{l(\overline{X})} = \frac{\exp\Big(-\dfrac{1}{2\sigma^2}\displaystyle\sum_{i=1}^n (X_i-\mu_0)^2\Big)}{\exp\Big(-\dfrac{1}{2\sigma^2}\displaystyle\sum_{i=1}^n (X_i-\overline{X})^2\Big)}$$
$$= \exp\Big(-\frac{1}{2\sigma^2/n}(\overline{X}-\mu_0)^2\Big) < c$$

となる c を求める。
H_0 が正しいときに，$\sqrt{n}(\overline{X}-\mu_0)/\sigma \sim N(0,1)$ となるので，

$$P\Big(\Big|\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\Big| > z_{\alpha/2}\Big) = \alpha$$

すなわち，

$$P\Big(\exp\Big(-\frac{1}{2\sigma^2/n}(\overline{X}-\mu_0)^2\Big) < \exp\Big(-\frac{1}{2}z_{\alpha/2}^2\Big)\Big) = \alpha$$

と変形できる。したがって，

$$c = \exp\Big(-\frac{1}{2}z_{\alpha/2}^2\Big)$$

とすればよい。

例 2： $X_1, X_n, \cdots, X_n$ は互いに独立で，それぞれパラメータ p を持ったベルヌイ分布に従うものとする。すなわち，X_i の確率関数は，

$$f(x;p) = p^x(1-p)^{1-x} \qquad x = 0,1$$

となる。

このとき尤度関数は，

$$l(p) = \prod_{i=1}^n f(x_i;p) = \prod_{i=1}^n p^{x_i}(1-p)^{1-x_i}$$

となる。
p の最尤推定量 $\widehat{p}$ は，

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

である。

次の仮説検定を考える。

$$H_0:\ p = p_0 \qquad H_1:\ p \neq p_0$$

→ 制約数は 1 つ。$(G=1)$

尤度比は，

$$\frac{l(p_0)}{l(\widehat{p})} = \frac{\prod_{i=1}^n p_0^{X_i}(1-p_0)^{1-X_i}}{\prod_{i=1}^n \widehat{p}^{X_i}(1-\widehat{p})^{1-X_i}}$$

したがって，$n \longrightarrow \infty$ のとき，

$$-2\log\frac{l(p_0)}{l(\widehat{p})} = -2\log\frac{p_0}{\widehat{p}}\sum_{i=1}^n X_i - 2\log\frac{1-p_0}{1-\widehat{p}}\sum_{i=1}^n (1-X_i)$$

$\longrightarrow\ \chi^2(1)$
$\chi^2(1)$ 分布の上側 100 α% 点を $\chi^2_\alpha(1)$ とするとき，

$$-2\log\frac{l(p_0)}{l(\widehat{p})} > \chi^2_\alpha(1)$$

のとき，帰無仮説 $H_0:\ p = p_0$ を棄却する。

例 3： 回帰モデル

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + u_i$$
$$u_i \sim N(0,\sigma^2) \qquad i = 1,2,\cdots,n$$

について，β_1, $\cdots$, β_k に関する仮説の尤度比検定を行う。
例えば，

$$H_0:\ \beta_1 = 0$$
$$H_0:\ \beta_1 + \beta_2 = 1$$
$$H_0:\ \beta_1 = \beta_2 = \beta_3 = 0$$

などのような仮説検定

$\theta = (\beta_1,\cdots,\beta_k,\sigma^2)$ とする。
尤度関数は，

$$l(\theta) = \prod_{i=1}^n f(Y_i;\theta)$$
$$= (2\pi\sigma^2)^{-\frac{n}{2}}\exp\Big(-\frac{1}{2\sigma^2}\sum_{i=1}^n (Y_i - \beta_1 X_{1i} - \cdots - \beta_k X_{ki})^2\Big)$$

となる。
H_0 の制約つき最尤推定量を $\widetilde{\theta} = (\widetilde{\beta}_1,\cdots,\widetilde{\beta}_k,\widetilde{\sigma}^2)$ とする。
この仮設に含まれる制約数を G とする。
制約なし最尤推定量を $\widehat{\theta} = (\widehat{\beta}_1,\cdots,\widehat{\beta}_k,\widehat{\sigma}^2)$ とする。
尤度比

$$\frac{l(\widetilde{\theta})}{l(\widehat{\theta})} = \frac{(2\pi\widetilde{\sigma}^2)^{-\frac{n}{2}}\exp\Big(-\dfrac{1}{2\widetilde{\sigma}^2}\displaystyle\sum_{i=1}^n (Y_i - \widetilde{\beta}_1 X_{1i} - \cdots - \widetilde{\beta}_k X_{ki})^2\Big)}{(2\pi\widehat{\sigma}^2)^{-\frac{n}{2}}\exp\Big(-\dfrac{1}{2\widehat{\sigma}^2}\displaystyle\sum_{i=1}^n (Y_i - \widehat{\beta}_1 X_{1i} - \cdots - \widehat{\beta}_k X_{ki})^2\Big)}$$

$$
\begin{aligned}
&= \frac{(\widetilde{\sigma}^2)^{-\frac{n}{2}} \exp\Big(-\dfrac{n-G}{2}\Big)}{(\widehat{\sigma}^2)^{-\frac{n}{2}} \exp\Big(-\dfrac{n-k}{2}\Big)} \\
&= \left(\frac{\dfrac{1}{n-G}\displaystyle\sum_{i=1}^n \widetilde{u}_i^2}{\dfrac{1}{n-k}\displaystyle\sum_{i=1}^n \widehat{u}_i^2}\right)^{-n/2} \exp\Big(-\frac{k-G}{2}\Big) \\
&= \exp\Big(-\frac{k-G}{2}\Big)\left(\frac{n-k}{n-G}\right)^{-n/2}\left(\frac{\sum_{i=1}^n \widetilde{u}_i^2}{\sum_{i=1}^n \widehat{u}_i^2}\right)^{-n/2} \\
&= \exp\Big(-\frac{k-G}{2}\Big)\left(\frac{n-k}{n-G}\right)^{-n/2} \\
&\quad \times \left(1 + \frac{\sum_{i=1}^n \widetilde{u}_i^2 - \sum_{i=1}^n \widehat{u}_i^2}{\sum_{i=1}^n \widehat{u}_i^2}\right)^{-n/2} \\
&= \exp\Big(-\frac{k-G}{2}\Big)\left(\frac{n-k}{n-G}\right)^{-n/2} \\
&\quad \times \left(1 + \frac{G}{n-k}\frac{(\sum_{i=1}^n \widetilde{u}_i^2 - \sum_{i=1}^n \widehat{u}_i^2)/G}{\sum_{i=1}^n \widehat{u}_i^2/(n-k)}\right)^{-n/2} \\
&< c
\end{aligned}
$$

のとき仮説を棄却する。

$$
\frac{(\sum_{i=1}^n \widetilde{u}_i^2 - \sum_{i=1}^n \widehat{u}_i^2)/G}{\sum_{i=1}^n \widehat{u}_i^2/(n-k)} \sim F(G, n-k)
$$

を利用すると c が求まる。

ただし，途中で以下を利用

$$
\begin{aligned}
\widetilde{\sigma}^2 &= \frac{1}{n-G}\sum_{i=1}^n (Y_i - \widetilde{\beta}_1 X_{1i} - \cdots - \widetilde{\beta}_k X_{ki})^2 \\
&= \frac{1}{n-G}\sum_{i=1}^n \widetilde{u}_i^2
\end{aligned}
$$

$$
\begin{aligned}
\widehat{\sigma}^2 &= \frac{1}{n-k}\sum_{i=1}^n (Y_i - \widehat{\beta}_1 X_{1i} - \cdots - \widehat{\beta}_k X_{ki})^2 \\
&= \frac{1}{n-k}\sum_{i=1}^n \widehat{u}_i^2
\end{aligned}
$$

近似的には，

$$
\begin{aligned}
-2\log\frac{l(\widetilde{\theta})}{l(\widehat{\theta})} &= -2\log\frac{(\widetilde{\sigma}^2)^{-\frac{n}{2}} \exp\Big(-\dfrac{n-G}{2}\Big)}{(\widehat{\sigma}^2)^{-\frac{n}{2}} \exp\Big(-\dfrac{n-k}{2}\Big)} \\
&= n\log\Big(\frac{\widetilde{\sigma}^2}{\widehat{\sigma}^2}\Big) + (k-G) \\
&\longrightarrow \chi^2(G)
\end{aligned}
$$

例 4： 回帰モデル

$$
\begin{aligned}
&Y_i = \alpha + \beta X_i + u_i \\
&u_i = \rho u_{i-1} + \epsilon_i \\
&\epsilon_i \sim N(0, \sigma^2) \qquad i = 2, 3, \cdots, n
\end{aligned}
$$

について，H_0： $\rho = 0$，H_1： $\rho \neq 0$ の尤度比検定を行う。

$\theta = (\alpha, \beta, \sigma^2, \rho)$ とする。対数尤度関数は，

$$
\begin{aligned}
\log l(\theta) &= \sum_{i=2}^n \log f(Y_i;\theta) = -\frac{n-1}{2}\log(2\pi) - \frac{n-1}{2}\log(\sigma^2) \\
&\quad -\frac{1}{2\sigma^2}\sum_{i=2}^n \Big((Y_i - \rho Y_{i-1}) - \alpha(1-\rho) - \beta(X_i - \rho X_{i-1})\Big)^2
\end{aligned}
$$

となる。
対数尤度関数をそれぞれ α，β，σ^2，ρ について微分し，ゼロとおく。4 本の連立方程式を解いて，制約なし最尤推定量 $\widehat{\theta} = (\widehat{\alpha}$，$\widehat{\beta}$，$\widehat{\sigma}^2$，$\widehat{\rho})$ が得られる。

$\rho = 0$ と制約をおく。$\theta = (\alpha, \beta, \sigma^2, 0)$ とする。対数尤度関数は，

$$
\begin{aligned}
\log l(\theta) &= \sum_{i=2}^n \log f(Y_i;\theta) = -\frac{n-1}{2}\log(2\pi) - \frac{n-1}{2}\log(\sigma^2) \\
&\quad -\frac{1}{2\sigma^2}\sum_{i=2}^n (Y_i - \alpha - \beta X_i)^2
\end{aligned}
$$

となる。
上記の対数尤度関数をそれぞれ α，β，σ^2 について微分し，ゼロとおく。3 本の連立方程式を解いて，$\rho = 0$ の制約付き最尤推定量 $\widetilde{\theta} = (\widetilde{\alpha}$，$\widetilde{\beta}$，$\widetilde{\sigma}^2$，$0)$ が得られる。

すなわち，

$$
\frac{\displaystyle\max_{\alpha,\beta,\sigma^2} l(\alpha,\beta,\sigma^2,0)}{\displaystyle\max_{\alpha,\beta,\sigma^2,\rho} l(\alpha,\beta,\sigma^2,\rho)} = \frac{l(\widetilde{\alpha},\widetilde{\beta},\widetilde{\sigma}^2,0)}{l(\widehat{\alpha},\widehat{\beta},\widehat{\sigma}^2,\widehat{\rho})} = \frac{l(\widetilde{\theta})}{l(\widehat{\theta})}
$$

$\log l(\widehat{\theta})$ は，$\widehat{\sigma}^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=2}^n \Big((Y_i - \widehat{\rho}Y_{i-1}) - \widehat{\alpha}(1-\widehat{\rho}) - \widehat{\beta}(X_i - \widehat{\rho}X_{i-1})\Big)^2$ に注意して，

$$
\begin{aligned}
\log l(\widehat{\theta}) &= -\frac{n-1}{2}\log(2\pi) - \frac{n-1}{2}\log(\widehat{\sigma}^2) \\
&\quad -\frac{1}{2\widehat{\sigma}^2}\sum_{i=2}^n \Big((Y_i - \widehat{\rho}Y_{i-1}) - \widehat{\alpha}(1-\widehat{\rho}) - \widehat{\beta}(X_i - \widehat{\rho}X_{i-1})\Big)^2 \\
&= -\frac{n-1}{2}\log(2\pi) - \frac{n-1}{2}\log(\widehat{\sigma}^2) - \frac{n-1}{2}
\end{aligned}
$$

となる。

同様に，$\log l(\widetilde{\theta})$ は，$\widetilde{\sigma}^2 = \frac{1}{n-1}\sum_{i=2}^{n}(Y_i - \widetilde{\alpha} - \widetilde{\beta}X_i)^2$ に注意して，

$$\begin{aligned}\log l(\widetilde{\theta}) &= -\frac{n-1}{2}\log(2\pi) - \frac{n-1}{2}\log(\widetilde{\sigma}^2) \\ &\quad -\frac{1}{2\widetilde{\sigma}^2}\sum_{i=2}^{n}(Y_i - \widetilde{\alpha} - \widetilde{\beta}X_i)^2 \\ &= -\frac{n-1}{2}\log(2\pi) - \frac{n-1}{2}\log(\widetilde{\sigma}^2) - \frac{n-1}{2}\end{aligned}$$

となる。

したがって，尤度比検定統計量

$$-2\log\frac{l(\widetilde{\theta})}{l(\widehat{\theta})} = (n-1)\log\frac{\widetilde{\sigma}^2}{\widehat{\sigma}^2}$$

は，n が大きくなると，$\chi^2(1)$ 分布に近づく。

13 時系列分析と季節調整

時系列データを分析する主要な目的は，そのデータがどのようなメカニズムで生成されたかを明らかにし，その構造に基づいてさまざまな状況での予測を行うことである。

時系列データ（X_t）が以下の 4 つの要素から構成されていると仮定されることが多い。

(1) 時間の経過とともに傾向的に変動する部分。トレンド（T_t，trend）という。

(2) ほぼ一定の周期をもって循環する部分。循環変動（C_t，cycle）という。

(3) 毎年同じ時期（季節）に規則的に観察される変動部分。季節変動（S_t，seasonality）という。

(4) (1) ～ (3) によって説明されない部分。不規則変動（I_t，irregularity）という。

すなわち，X_t は，

加法型モデル： $X_t = T_t + C_t + S_t + I_t$

または，

乗法型モデル： $X_t = T_t \times C_t \times S_t \times I_t$

と表される。

この接近方法を時系列の分解アプローチと呼ぶ。このアプローチではそれぞれの変動を特定化し，それらを総合することにより原系列の変動の特徴を探り，予測に利用する。このアプローチでは，生成メカニズムの理論的背景（例えば，経済学の理論的背景）についての知識は必要としない。

また，時系列データの生成過程を確率的に描写し，データの生成された確率的構造を明らかにし，その構造に基づいて予測などを行う場合も考えられる。このような接近方法を時系列分析（time series analysis）という。この分析方法も生成メカニズムの理論的背景についての知識は必要としない。

13.1 季節変動

季節調整済みデータ（seasonally adjusted data）: 原系列（当初に与えられたデータの系列）から季節変動をとり除いたデータ

季節調整済みデータを作成する簡単な方法 $\longrightarrow$ 移動平均法 (moving average) : 1 年分のデータを平均することにより季節変動を除去する。

移動平均値の中心化による t 期の季節調整値を Y_t で表す。

月次データの季節調整：

$$\begin{aligned}Y_t &= \underbrace{\frac{1}{2}\Big(\underbrace{\frac{1}{12}(X_{t-6} + \cdots + X_{t+5})}_{(t-0.5)\text{ 期の季節調整値}} + \underbrace{\frac{1}{12}(X_{t-5} + \cdots + X_{t+6})}_{(t+0.5)\text{ 期の季節調整値}}\Big)}_{t\text{ 期の季節調整値}} \\ &= \frac{1}{24}\Big(X_{x-6} + 2(X_{t-5} + \cdots + X_{t+5}) + X_{t+6}\Big)\end{aligned}$$

四半期データの季節調整：

$$\begin{aligned}Y_t &= \underbrace{\frac{1}{2}\Big(\underbrace{\frac{1}{4}(X_{t-2} + \cdots + X_{t+1})}_{(t-0.5)\text{ 期の季節調整値}} + \underbrace{\frac{1}{4}(X_{t-1} + \cdots + X_{t+2})}_{(t+0.5)\text{ 期の季節調整値}}\Big)}_{t\text{ 期の季節調整値}} \\ &= \frac{1}{8}\Big(X_{x-2} + 2(X_{t-1} + X_t + X_{t+1}) + X_{t+2}\Big)\end{aligned}$$

この移動平均の中心化法による季節調整の欠点：

(1) 対象時系列データの前後 6 カ月の季節調整ができない。
(2) 季節調整値を得るためには移動平均とり，さらにその平均をとるといった作業を行っている。時系列データの季節変動だけでなく，不規則変動もならしてしまい，過度にスムーズな系列ができる危険性がある。

現在よく利用されている米国のセンサス局で開発されたセンサス局法（X11，X12-ARIMA）は，このような点を考慮して開発された季節調整法

13.2　トレンド

時系列データの傾向的な動きを特定化する方法として，トレンド部分を時間の具体的な関数で近似し，その係数を回帰分析によって求める。
一次式：

$$T_t = \alpha + \beta\, t$$

二次式：

$$T_t = \alpha + \beta\, t + \gamma\, t^2$$

成長率が一定の場合には，次のような指数関数によってトレンド部分を描写する（b は成長率を表す）。

$$T_t = Ae^{bt}$$

対数変換して，

$$\log T_t = A' + b\, t$$

とする。

T_t には季節調整済みデータを用いる。

13.3　循環変動

季節調整系列からトレンド部分を差し引いた系列を求める。

$$C_t = T_t - \widehat{\alpha} - \widehat{\beta}\, t$$

時系列データに内蔵されている周期性を計測し，それがどのような要因によって引き起こされているのかを探ることにより景気変動のメカニズムを理解しようという試みが数多くなされてきた。

14　説明変数と誤差項に相関がある場合

$$Y_i = \alpha + \beta X_i + u_i, \qquad i = 1, 2, \cdots, n$$

$u_1, u_2, \cdots, u_n$ は互いに独立で，すべての i について $u_i \sim N(0, \sigma^2)$ と仮定する。

(1) X_i は非確率変数 $\longrightarrow$ 今までの最小二乗推定量

(2) X_i は確率変数

(2a) X_i と u_i は相関なし $\longrightarrow$ $\mathrm{E}(X_i u_i) = 0$

(2b) X_i は u_i は相関あり $\longrightarrow$ $\mathrm{E}(X_i u_i) \neq 0$

最小二乗推定量

$$\begin{aligned}\hat{\beta} &= \frac{\sum_{i=1}^n (X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^n (X_i - \overline{X})^2} \\ &= \beta + \sum_{i=1}^n \omega_i u_i\end{aligned}$$

ただし，$\omega_i = \dfrac{X_i - \overline{X}}{\sum_{j=1}^n (X_j - \overline{X})^2}$ である。

(2a) の場合，

$$\begin{aligned}\mathrm{E}(\hat{\beta}) &= \beta + \sum_{i=1}^n \mathrm{E}(\omega_i u_i) \\ &= \beta + \sum_{i=1}^n \mathrm{E}(\omega_i)\mathrm{E}(u_i) = \beta\end{aligned}$$

$\longrightarrow$ 不偏推定量

(2b) の場合，

$$\begin{aligned}\mathrm{E}(\hat{\beta}) &= \beta + \sum_{i=1}^n \mathrm{E}(\omega_i u_i) \\ &= \beta + \sum_{i=1}^n \mathrm{E}(\omega_i u_i) \neq \beta\end{aligned}$$

$\longrightarrow$ 不偏推定量でない

復習（?）: 統計学 — 大数の法則 1 :　$X_1, X_2, \cdots, X_n$ は互いに独立で同じ分布にしたがうものとする。$\mathrm{E}(X_i) = \mu$ とする。

$$n \longrightarrow \infty \quad \text{のとき，} \quad \overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i \longrightarrow \mu$$

または，

$$\text{plim}\ \overline{X}_n = \mu$$

plim とは probability limit の意味で，「$\overline{X}_n$ は μ に確率収束する」，または，「$\overline{X}_n$ の確率極限は μ」という。$\longrightarrow$ 大数の法則

復習（？）：統計学 — 大数の法則 2： $X_1, X_2, \cdots, X_n$ は $\text{E}(X_i) = \mu_i$，$\text{V}(X_i) = \sigma^2$ の分布にしたがうものとする。$n \longrightarrow \infty$ のとき，

$$\frac{1}{n}\sum_{i=1}^{n}\mu_i \longrightarrow \mu < \infty, \qquad \frac{1}{n}\sum_{i=1}^{n}\sigma_i^2 \longrightarrow \sigma^2 < \infty$$

とする。このとき，

$$n \longrightarrow \infty \quad \text{のとき，} \quad \overline{X}_n = \frac{1}{n}\sum_{i=1}^{n}X_i \longrightarrow \mu$$

または，

$$\text{plim}\ \overline{X}_n = \mu$$

$\longrightarrow$ 大数の法則

例： X_i をサイコロを投げて出た目とする。$i = 1, 2, \cdots, n$ として，n 回の実験で出た目を $x_1, x_2, \cdots, x_n$ とする。X_i の取り得る値は 1, 2, 3, 4, 5, 6 で，その確率はすべて 1/6 であるので，

$$\text{E}(X_i) = \frac{1}{6}(1 + 2 + \cdots + 6) = 3.5$$

大数の法則によると，

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n}x_i = \text{E}(X_i) = 3.5$$

となる。
$(1/n)\sum_{i=1}^{n}x_i$ は実験値，$\text{E}(X_i)$ は理論値

いくつか公式： $\text{plim}\ a_n = c_1$，$\text{plim}\ b_n = c_2$ とする。このとき，

$$\text{plim}\ a_n b_n = \text{plim}\ a_n \times \text{plim}\ b_n = c_1 c_2,$$

$$\text{plim}\ \frac{a_n}{b_n} = \frac{\text{plim}\ a_n}{\text{plim}\ b_n} = \frac{c_1}{c_2}, \quad \text{ただし，} c_2 \neq 0 \text{ とする}$$

最小二乗推定量の一致性について： 最小二乗推定量

$$\begin{aligned}\hat{\beta} &= \beta + \sum_{i=1}^{n}\omega_i u_i \\ &= \beta + \frac{(1/n)\sum_{i=1}^{n}(X_i - \overline{X})u_i}{(1/n)\sum_{j=1}^{n}(X_j - \overline{X})^2}\end{aligned}$$

下記を仮定する。

$$\text{plim}\ \Big(\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2\Big) = M_{xx} > 0$$

$\longrightarrow$ X の分散に対応

$$\text{plim}\ \Big(\frac{1}{n}\sum_{i=1}^{n}X_i u_i\Big) = M_{xu}$$

$\longrightarrow$ X と u の共分散に対応

$$\text{plim}\ \Big(\frac{1}{n}\sum_{i=1}^{n}u_i\Big) = 0$$

$\longrightarrow$ u の平均に対応（誤差項の仮定より）

したがって，$\hat{\beta}$ の確率極限は下記の通りとなる。

$$\begin{aligned}\text{plim}\ \hat{\beta} &= \text{plim}\ \Big(\beta + \frac{(1/n)\sum_{i=1}^{n}(X_i - \overline{X})u_i}{(1/n)\sum_{j=1}^{n}(X_j - \overline{X})^2}\Big) \\ &= \beta + \frac{\text{plim}\ (1/n)\sum_{i=1}^{n}(X_i - \overline{X})u_i}{\text{plim}\ (1/n)\sum_{j=1}^{n}(X_j - \overline{X})^2} \\ &= \beta + \frac{M_{xu}}{M_{xx}}\end{aligned}$$

ただし，右辺第二項の分子は，

$$\begin{aligned}&\text{plim}\ \Big(\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})u_i\Big) \\ &= \text{plim}\ \Big(\frac{1}{n}\sum_{i=1}^{n}X_i u_i\Big) - \text{plim}\ \overline{X}\Big(\frac{1}{n}\sum_{i=1}^{n}u_i\Big) \\ &= \lim\ \Big(\frac{1}{n}\sum_{i=1}^{n}\text{E}(X_i u_i)\Big) - \mu\ \lim\ \Big(\frac{1}{n}\sum_{i=1}^{n}\text{E}(u_i)\Big) \\ &= M_{xu}\end{aligned}$$

となることに注意。
結果をもう一度書くと，

$$\text{plim}\ \hat{\beta} = \beta + \frac{M_{xu}}{M_{xx}}$$

となる。

- $M_{xu}=0$ のとき，$\text{plim}\ \hat\beta=\beta$，すなわち，$\hat\beta \longrightarrow \beta$ $\Longrightarrow$ $\hat\beta$ は β の一致推定量

- $M_{xu}\neq 0$ のとき，$\text{plim}\ \hat\beta\neq\beta$，すなわち，$\hat\beta \not\longrightarrow \beta$ $\Longrightarrow$ $\hat\beta$ は一致性なし

説明変数 X_i と誤差項 u_i との間に相関がある場合，最小二乗推定量 $\hat\beta$ に不偏性も一致性もない。

対策：操作変数法： 説明変数 X_i と誤差項 u_i との間に相関がある場合，

$$\text{plim}\ \Big(\frac{1}{n}\sum_{i=1}^{n}X_iu_i\Big)=M_{xu}\neq 0$$

となることが問題。

したがって，X_i の代わりに，

$$\text{plim}\ \Big(\frac{1}{n}\sum_{i=1}^{n}Z_iu_i\Big)=0$$

となるような別の変数 Z_i を考える。しかも，Z_i は X_i と同じ動き（すなわち，強い相関）になることが必要。

(*) $\mathrm{E}(u_i)=0$ から，下記の条件は満たしているものとする。

$$\text{plim}\ \Big(\frac{1}{n}\sum_{i=1}^{n}u_i\Big)=0$$

次の二つの連立方程式を満たすような α，β の解が一致推定量となる。

$$\frac{1}{n}\sum_{i=1}^{n}u_i=\frac{1}{n}\sum_{i=1}^{n}(Y_i-\alpha-\beta X_i)=0$$
$$\frac{1}{n}\sum_{i=1}^{n}Z_iu_i=\frac{1}{n}\sum_{i=1}^{n}Z_i(Y_i-\alpha-\beta X_i)=0$$

この解を $(\alpha_{IV}$，$\beta_{IV})$ とする。
すなわち，

$$\sum_{i=1}^{n}(Y_i-\alpha_{IV}-\beta_{IV}X_i)=0$$
$$\sum_{i=1}^{n}Z_i(Y_i-\alpha_{IV}-\beta_{IV}X_i)=0$$

を $(\alpha_{IV}$，$\beta_{IV})$ について解く。

一つ目の式から得られる $\alpha_{IV}=\overline{Y}-\beta_{IV}\overline{X}$ を用いて，α_{IV} を消去して，β_{IV} について解くと，

$$\begin{aligned}\beta_{IV}&=\frac{\sum_{i=1}^{n}Z_iY_i-\overline{Z}\ \overline{Y}}{\sum_{i=1}^{n}Z_iX_i-\overline{Z}\ \overline{X}}\\&=\frac{\sum_{i=1}^{n}(Z_i-\overline{Z})(Y_i-\overline{Y})}{\sum_{i=1}^{n}(Z_i-\overline{Z})(X_i-\overline{X})}\end{aligned}$$

となる。Z_i の代わりに X_i を使えば，β_{IV} は $\hat\beta$ に一致する。

Z_i を操作変数（instrumental variable），
$(\alpha_{IV}$，$\beta_{IV})$ を操作変数推定量（instrumental variable estimator），または，操作変数法（instrumental variable method）と呼ぶ。

操作変数法の一致性の証明： β_{IV} に注目して，下記のように書き直す。

$$\begin{aligned}\beta_{IV}&=\frac{\sum_{i=1}^{n}(Z_i-\overline{Z})(Y_i-\overline{Y})}{\sum_{j=1}^{n}(Z_j-\overline{Z})(X_j-\overline{X})}=\sum_{i=1}^{n}\omega_i^*(Y_i-\overline{Y})\\&=\sum_{i=1}^{n}\omega_i^*Y_i\end{aligned}$$

ただし，

$$\omega_i^*=\frac{Z_i-\overline{Z}}{\sum_{j=1}^{n}(Z_j-\overline{Z})(X_j-\overline{X})}$$

とする（ω_i と区別するために，右肩に $*$ を付ける)。途中で，ω_i と同様に $\sum_{i=1}^{n}\omega_i^*=0$ を使っている。
さらに，$Y_i=\alpha+\beta X_i+u_i$ に注意して，書き換えると，

$$\begin{aligned}\beta_{IV}&=\sum_{i=1}^{n}\omega_i^*Y_i=\sum_{i=1}^{n}\omega_i^*(\alpha+\beta X_i+u_i)\\&=\beta+\sum_{i=1}^{n}\omega_i^*u_i\end{aligned}$$

式の途中で，

$$\sum_{i=1}^{n}\omega_i^*=0$$
$$\sum_{i=1}^{n}\omega_i^*X_i=\sum_{i=1}^{n}\omega_i^*(X_i-\overline{X})=1$$

を利用している。

$$
\begin{aligned}
\beta_{IV} &= \beta + \sum_{i=1}^{n} \omega_i^* u_i \\
&= \beta + \frac{(1/n)\sum_{i=1}^{n}(Z_i - \overline{Z})u_i}{(1/n)\sum_{j=1}^{n}(Z_j - \overline{Z})(X_j - \overline{X})} \\
&\longrightarrow\ \beta + \frac{M_{zu}}{M_{zx}} = \beta
\end{aligned}
$$

途中の式では，大数の法則を使うために，分子分母に $1/n$ を掛けている。Z と u は相関がないと仮定したので，$M_{zu} = 0$ に注意。

$$
\begin{aligned}
M_{zu} &= \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n}(Z_i - \overline{Z})u_i \\
&= \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n}\mathrm{Cov}(Z_i, u_i) = 0 \\
M_{zx} &= \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n}(Z_i - \overline{Z})(X_i - \overline{X}) \\
&= \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n}\mathrm{Cov}(Z_i, X_i) \neq 0
\end{aligned}
$$

を利用。
よって，β_{IV} は β の一致推定量となる。

代表的な Z_i の求め方： $Z_i = \hat{X}_i$ とする。ただし，$\hat{X}_i$ は X_i の最小二乗法による予測値である。
$(X_i, W_{1i}, W_{2i},\ \cdots,\ W_{mi})$, $i = 1, 2, \cdots, n$, のデータから

$$
X_i = \gamma_1 W_{1i} + \gamma_2 W_{2i} +\ \cdots\ + \gamma_m W_{mi} + v_i
$$

を最小二乗法で推定する。

$$
\hat{X}_i = \hat{\gamma}_1 W_{1i} + \hat{\gamma}_2 W_{2i} +\ \cdots\ + \hat{\gamma}_m W_{mi}
$$

を求める。ただし，$j = 1, 2, \cdots, m$ について，W_{ji} と u_i とは相関がないものとする。

すなわち，β_{TSLS} とおく。

$$
\beta_{TSLS} = \frac{\sum_{i=1}^{n}(\hat{X}_i - \overline{\hat{X}})(Y_i - \overline{Y})}{\sum_{j=1}^{n}(\hat{X}_j - \overline{\hat{X}})(X_j - \overline{X})}
$$

この推定量を 2 段階最小二乗法（two-stage least squares）という。

この推定量が一致性を持つ理由は，Z_i の代わりに $\hat{X}_i$ を用いることにしているので，

$$
\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n}(Z_i - \overline{Z})u_i &= \frac{1}{n}\sum_{i=1}^{n}(\hat{X}_i - \overline{\hat{X}})u_i \\
&= \frac{1}{n}\sum_{i=1}^{n}\hat{X}_i u_i - \overline{\hat{X}} \times \frac{1}{n}\sum_{i=1}^{n} u_i \ \longrightarrow\ \mathrm{E}(\hat{X}_i u_i) = 0
\end{aligned}
$$

を示せばよい。$(1/n)\sum_{i=1}^{n} u_i \ \longrightarrow\ \mathrm{E}(u_i) = 0$ に注意。

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\hat{X}_i u_i \\
&= \frac{1}{n}\sum_{i=1}^{n}(\hat{\gamma}_1 W_{1i} + \hat{\gamma}_2 W_{2i} +\ \cdots\ + \hat{\gamma}_m W_{mi})u_i \\
&= \hat{\gamma}_1 \frac{1}{n}\sum_{i=1}^{n} W_{1i}u_i + \hat{\gamma}_2 \frac{1}{n}\sum_{i=1}^{n} W_{2i}u_i \\
&\qquad +\ \cdots\ + \hat{\gamma}_m \frac{1}{n}\sum_{i=1}^{n} W_{mi}u_i \ \longrightarrow\ 0
\end{aligned}
$$

なぜなら，$j = 1, 2, \cdots, m$ について，$(1/n)\sum_{i=1}^{n} W_{ji}u_i \longrightarrow \mathrm{E}(W_{ji}u_i) = 0$ となる。

2 段階最小二乗法を書き換えると，

$$
\begin{aligned}
\beta_{TSLS} &= \frac{\sum_{i=1}^{n}(\hat{X}_i - \overline{\hat{X}})(Y_i - \overline{Y})}{\sum_{j=1}^{n}(\hat{X}_j - \overline{\hat{X}})(X_j - \overline{X})} \\
&= \frac{\sum_{i=1}^{n}(\hat{X}_i - \overline{X})(Y_i - \overline{Y})}{\sum_{j=1}^{n}(\hat{X}_j - \overline{X})^2}
\end{aligned}
$$

となる。最後の等号は，$\overline{X} = \overline{\hat{X}}$，$\sum_{j=1}^{n}(\hat{X}_j - \overline{\hat{X}})(X_j - \overline{X}) = \sum_{j=1}^{n}(\hat{X}_j - \overline{X})^2$ となることを利用（証明略）。

- 操作変数 W_{1i}, W_{2i}, $\cdots$, W_{ki} を用いて，X_i の予測値 $\hat{X}_i$ を最小二乗法で求める。

- $Y_i = \alpha + \beta\hat{X}_i + e_i$ を最小二乗法で推定する。$e_i = u_i - \beta\hat{v}_i$ となることに注意。

計量経済学 練習問題

$\boxed{1}$ n 組のデータ (Y_1, X_1), (Y_2, X_2), $\cdots$, (Y_n, X_n) が以下の表で与えられたとき，

$$Y_i = \alpha + \beta X_i + u_i,$$

を最小二乗法で推定する。ただし，u_1, u_2, $\cdots$, u_n は互いに独立で，$u_i \sim N(0, \sigma^2)$ とする。

i	Y_i	X_i	Y_i^2	X_i^2	Y_iX_i	$\widehat{Y}_i$	$\widehat{u}_i$	$\widehat{u}_i^2$
1	2	1						
2	1	2						
3	2	3						
4	3	2						
5	4	3						
合計								
平均								

このとき，以下の問に答えよ。表の空欄は，各自の計算のために使ってよい。

① α, β の最小二乗推定値 $\widehat{\alpha}$, $\widehat{\beta}$ をそれぞれ求めよ。

② 決定係数 R^2 を求めよ。

③ 最小二乗推定量 $\widehat{\alpha}$, $\widehat{\beta}$ の分散 $\mathrm{V}(\widehat{\alpha})$, $\mathrm{V}(\widehat{\beta})$ をそれぞれ求めよ。

④ σ^2 の不偏推定値 s^2 を求めよ。

⑤ 自由度修正済み決定係数 $\overline{R}^2$ をそれぞれ求めよ。

⑥ $\widehat{\beta}$ の標準誤差を求めよ。

⑦ β の 99% 信頼区間をそれぞれ求めよ。

⑧ 帰無仮説 H_0 : $\beta = 0$，対立仮説 H_1 : $\beta \neq 0$ を，有意水準 5% で検定せよ。

⑨ ダービン・ワトソン比を求めよ。

定期試験について (2005 年前期)

● 前ページの「計量経済学 (木 2) 練習問題」を解けるようにしておくこと。

● 以下のことを前もって考えておくこと。

為替レート：

$$Exr_i = f(r_i, r_i^*, p_i, p_i^*, B_i, s_i)$$

r_i, r_i^*, p_i, p_i^*, B_i, s_i が Exr_i に与える影響を考えておいて下さい。
ただし，各変数は次のとおり。

Exr_i	t 期の為替レート
B_i	t 期の貿易収支 (= 輸出 − 輸入)
p_i	t 期の物価上昇率
p_i^*	t 期の米国の物価上昇率
r_i	t 期の金利
r_i^*	t 期の米国の金利
s_i	t 期の株価

次の質問に答えられるようにしておいて下さい。

1. 係数の符号はどうなるべきか？
2. R^2 の意味は？
3. R^2 の問題点を含めて，$\overline{R}^2$ の意味は？
4. 真の係数の値はどのあたりにあるか？
5. 真の係数の値がゼロという仮説はどのような意味を持つか？
6. DW の意味と解釈は？
7. 1〜6 について，実際に推定した結果，どのような結果が得られたか？

● 『基本統計学』(第 3 版) の p.190 の練習問題 9.6 と p.217 の練習問題 10.1〜10.4 も参考に。