

エコノメトリックス

(2014年度後期 講義ノート)

平成 26 年 10 月 1 日 (水) 版

教科書『計量経済学』
(山本拓著, 新世社, 1995 年)

谷崎 久志
大阪大学・経済学部

目 次

1	計量経済学について	1	4	統計学の回帰分析への応用	9
1.1	例 1: マクロの消費関数	1	4.1	回帰モデルの仮定	10
1.2	例 2: 日本酒の需要関数	2	4.2	誤差項 (攪乱項) の経済学的意味	11
2	行列について	2	4.3	$\hat{\alpha}$, $\hat{\beta}$ の統計的性質	11
3	最小二乗法について	6	4.3.1	$\hat{\beta}$ について	11
3.1	最小二乗法と回帰直線	6	4.3.2	$\hat{\alpha}$ について	11
3.2	切片 α と傾き β の推定	6	4.3.3	$\hat{\alpha}$, $\hat{\beta}$ の平均	11
3.3	残差 $\hat{u}_i$ の性質について	7	4.3.4	$\hat{\alpha}$, $\hat{\beta}$ の分散	12
3.4	決定係数 R^2 について	8	4.3.5	$\hat{\alpha}$, $\hat{\beta}$ の分布 (σ^2 が既知の場合)	14
3.5	まとめ	9	4.3.6	$\hat{\alpha}$, $\hat{\beta}$ の性質: 最良線型不偏性と一 致性	15
			4.4	誤差項 (または, 攪乱項) u_i の分散 σ^2 に ついて	17
			4.4.1	$\hat{\alpha}$, $\hat{\beta}$ の分散の不偏推定量	19

4.5	$\hat{\alpha}$, $\hat{\beta}$ の分布	20	12.2.3 誤差項に系列相関がある場合	51
4.5.1	統計学の復習 (t 分布)	20	12.3 尤度比検定	52
4.5.2	$\hat{\beta}$ について:	20	13 時系列分析と季節調整	55
4.5.3	$\hat{\alpha}$ について:	21	13.1 季節変動	55
4.5.4	まとめ:	21	13.2 トレンド	55
4.6	α , β の区間推定 (信頼区間)	21	13.3 循環変動	56
4.6.1	統計学の復習: 区間推定 (信頼区間)	21	14 時系列分析と定常性	56
4.6.2	α , β の区間推定 (信頼区間)	22	14.1 時系列モデルの特定化	57
4.7	α , β の仮説検定	22	14.1.1 自己回帰 (AR) モデル	57
4.7.1	統計学の復習: 仮説検定	22	14.1.2 移動平均 (MA) モデル	57
4.7.2	α , β の仮説検定	23	14.1.3 より複雑なモデル	57
4.7.3	t 値について	24	14.2 時系列モデルの作成手順と予測	58
5	多重回帰	25	14.3 非定常時系列	58
5.1	推定量の性質	26	14.3.1 単位根	58
5.2	ダミー変数について	27	14.3.2 見せかけ回帰	59
5.2.1	異常値	27	14.3.3 共和分	60
5.2.2	構造変化	28		
6	関数型について	29		
7	系列相関: DW について	30		
7.1	DW について	30		
7.2	系列相関のもとで回帰式の推定	32		
8	不均一分散 (不等分散)	33		
9	多重共線性について	34		
10	F 検定について	35		
10.1	いくつかの例	35		
10.2	統計学の復習	36		
10.3	検定の方法	36		
11	応用例	37		
11.1	マクロの消費関数	37		
11.2	ミクロの消費関数 (需要関数)	40		
11.3	株価, 金利, 為替レート	43		
12	推定量の求め方	44		
12.1	最小二乗法	44		
12.2	最尤法	44		
12.2.1	変数変換	49		
12.2.2	回帰分析への応用	49		

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, ...)
- データ(GNP, 消費, 投資, 金利, 為替レート, ...)

計量経済学 $\Rightarrow$ 経済理論が現実にな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。

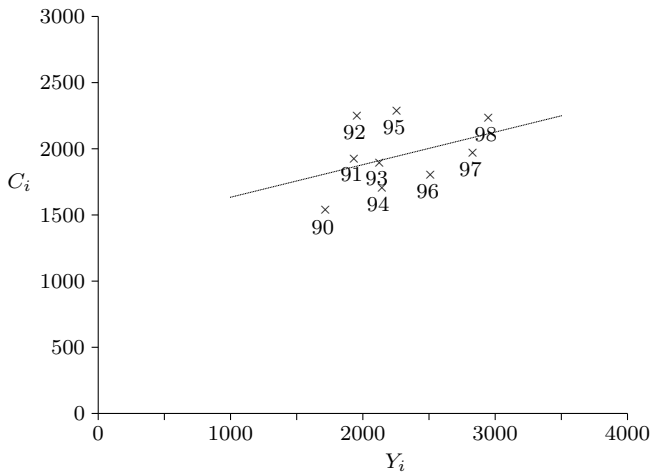
1. $Y \nearrow \Rightarrow C \nearrow$
2. $\frac{dC}{dY}$ = 限界消費性向 = 所得1円増加で消費が何円増加するか
3. すなわち, $\frac{dC}{dY} > 0$

モデルの定式化

1. $C = a + bY$
2. $b = \frac{dC}{dY}$ = 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/2014>
からダウンロード可。

図 1：消費 (C_i) と所得 (Y_i)



1. $\times \rightarrow$ 実際のデータ
2. $(Y_i, C_i) \Rightarrow t$ 期のデータ, i.e., $i = 1, 2, \dots, 9$
3. $i = 1 \Rightarrow$ 1990 年,
 $i = 2 \Rightarrow$ 1991 年,
 $\dots$,
 $i = 9 \Rightarrow$ 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. $Y \nearrow \Rightarrow Q \nearrow$,
 $P_1 \nearrow \Rightarrow Q \searrow$,
 $P_2 \nearrow \Rightarrow Q \nearrow$
2. $\frac{\partial Q}{\partial Y} > 0, \frac{\partial Q}{\partial P_1} < 0, \frac{\partial Q}{\partial P_2} > 0$
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., $i = 1, 2, \dots, n$
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)

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

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,

$$a = (a_1 \quad a_2)$$

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,

$$a = (a_1 \quad \cdots \quad a_k)$$

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 $i = 1, \dots, n, j = 1, \dots, k$ 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。

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \quad \text{または} \quad (x \ y) = (3 \ 2)$$

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。

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

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

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

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。

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

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

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

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

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

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

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

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

このように, AB と BA の次元は異なる。

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

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

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

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。

$$I_n A = A I_n = A \quad I_n x = x$$

$$\begin{aligned}
 &\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \\
 &= \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \\
 &= \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \\
 &\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}
 \end{aligned}$$

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。

$$(A')' = A$$

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad x' = (x_1 \quad x_2 \quad \cdots \quad x_n)$$

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 $\Rightarrow$ 最小二乗法

3.1 最小二乗法と回帰直線

$(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ 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 $\{(X_i, Y_i), i = 1, 2, \dots, n\}$ 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 $\hat{\alpha}, \hat{\beta}$ 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 $\hat{\alpha}, \hat{\beta}$ となる。

すなわち, $\hat{\alpha}, \hat{\beta}$ は,

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

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

を満たす。

さらに,

$$\begin{aligned} \sum_{i=1}^n Y_i &= n\hat{\alpha} + \hat{\beta} \sum_{i=1}^n X_i, \\ \sum_{i=1}^n X_i Y_i &= \hat{\alpha} \sum_{i=1}^n X_i + \hat{\beta} \sum_{i=1}^n X_i^2, \end{aligned} \quad (3)$$

行列表示によって,

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

逆行列の公式:

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

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

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

さらに, $\hat{\beta}$ について解くと,

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

連立方程式の (3) 式から,

$$\hat{\alpha} = \bar{Y} - \hat{\beta} \bar{X}$$

となる。ただし,

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

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 $\hat{\alpha}$, $\hat{\beta}$ を求める。

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

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

$$\hat{\beta} = \frac{\sum_{i=1}^n X_i Y_i - n \bar{X} \bar{Y}}{\sum_{i=1}^n X_i^2 - n \bar{X}^2}$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta} \bar{X}$$

なので、必要なものは $\bar{X}$, $\bar{Y}$, $\sum_{i=1}^n X_i^2$, $\sum_{i=1}^n X_i Y_i$ 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	$\sum Y_i$ 35	$\sum X_i$ 52	$\sum X_i Y_i$ 468	$\sum X_i^2$ 696
平均	$\bar{Y}$ 8.75	$\bar{X}$ 13		

よって,

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

$$\hat{\alpha} = 8.75 - 0.65 \times 13 = 0.3$$

となる。

注意事項：

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

回帰直線は

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

として与えられる。

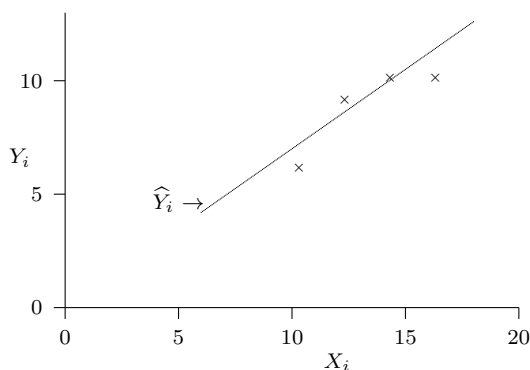
上の数値例では,

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

となる。

i	Y_i	X_i	$X_i Y_i$	X_i^2	$\hat{Y}_i$
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	$\sum Y_i$ 35	$\sum X_i$ 52	$\sum X_i Y_i$ 468	$\sum X_i^2$ 696	$\sum \hat{Y}_i$ 35.0
平均	$\bar{Y}$ 8.75	$\bar{X}$ 13			

図 2: Y_i , X_i , $\hat{Y}_i$



$\hat{Y}_i$ を実績値 Y_i の予測値または理論値と呼ぶ。

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

$\hat{u}_i$ を残差と呼ぶ。

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

さらに, $\bar{Y}$ を両辺から引いて,

$$(Y_i - \bar{Y}) = (\hat{Y}_i - \bar{Y}) + \hat{u}_i,$$

3.3 残差 $\hat{u}_i$ の性質について

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

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

を得る。

(2) 式から,

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

を得る。

$\hat{Y}_i = \hat{\alpha} + \hat{\beta}X_i$ から、

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

を得る。なぜなら、

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

である。

i	Y_i	X_i	$\hat{Y}_i$	$\hat{u}_i$	$X_i \hat{u}_i$	$\hat{Y}_i \hat{u}_i$
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	$\sum Y_i$ 35	$\sum X_i$ 52	$\sum \hat{Y}_i$ 35.0	$\sum \hat{u}_i$ 0.0	$\sum X_i \hat{u}_i$ 0.0	$\sum \hat{Y}_i \hat{u}_i$ 0.00

3.4 決定係数 R^2 について

次の式

$$(Y_i - \bar{Y}) = (\hat{Y}_i - \bar{Y}) + \hat{u}_i,$$

の両辺を二乗して、総和すると、

$$\begin{aligned} &\sum_{i=1}^n (Y_i - \bar{Y})^2 \\ &= \sum_{i=1}^n \left((\hat{Y}_i - \bar{Y}) + \hat{u}_i \right)^2 \\ &= \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2 + 2 \sum_{i=1}^n (\hat{Y}_i - \bar{Y}) \hat{u}_i + \sum_{i=1}^n \hat{u}_i^2 \\ &= \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2 + \sum_{i=1}^n \hat{u}_i^2 \end{aligned}$$

となる。まとめると、

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

を得る。さらに、

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

それぞれの項は、

$$1. \sum_{i=1}^n (Y_i - \bar{Y})^2 \implies y \text{ の全変動}$$

$$2. \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2 \implies \hat{Y}_i \text{ (回帰直線) で説明される部分}$$

$$3. \sum_{i=1}^n \hat{u}_i^2 \implies \hat{Y}_i \text{ (回帰直線) で説明されない部分}$$

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。

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

または、

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

として書き換えられる。

または、 $Y_i = \hat{Y}_i + \hat{u}_i$ と

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

を用いて、

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

と書き換えられる。すなわち、 R^2 は Y_i と $\hat{Y}_i$ の相関係数の二乗と解釈される。

$$\sum_{i=1}^n (Y_i - \bar{Y})^2 = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2 + \sum_{i=1}^n \hat{u}_i^2 \text{ から、明らかに、}$$

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

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。

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

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

i	Y_i	X_i	$\hat{Y}_i$	$\hat{u}_i$	$\hat{u}_i^2$	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	$\sum Y_i$ 35	$\sum X_i$ 52	$\sum \hat{Y}_i$ 35.0	$\sum \hat{u}_i$ 0.0	$\sum \hat{u}_i^2$ 2.30	$\sum Y_i^2$ 317

$\sum \hat{u}_i^2 = 2.30$, $\bar{X} = 13$, $\bar{Y} = 8.75$, $\sum_{i=1}^n Y_i^2 = 317$ 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

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

$$\hat{\beta} = \frac{\sum_{i=1}^n X_i Y_i - n\bar{X}\bar{Y}}{\sum_{i=1}^n X_i^2 - n\bar{X}^2}$$

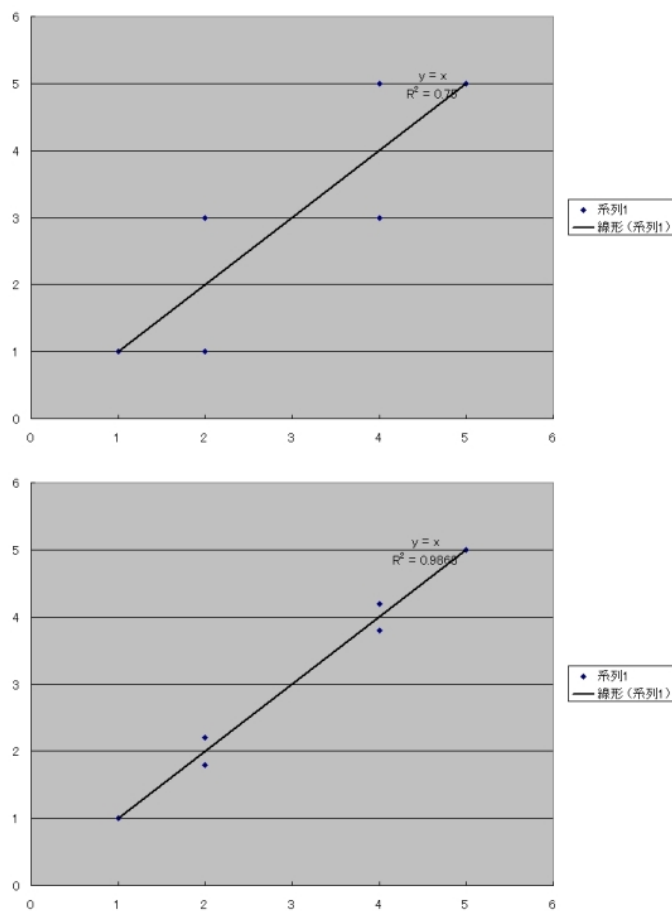
$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X}$$

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

決定係数の計算には以下の公式を用いる。

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

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



4 統計学の回帰分析への応用

$(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ 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。

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

$$\hat{\beta} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

$$= \frac{\sum_{i=1}^n (X_i - \bar{X})Y_i}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X},$$

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

である。

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

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

残差 $\hat{u}_i$ 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, 攪乱項) u_i を含め, それを確率変数として考える。

⇒ 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)

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

1. 残差 $\hat{u}_i$ は u_i の実現値としてみなすことができる。
2. $\hat{\alpha}, \hat{\beta}$ 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

$$\Rightarrow X_1, X_2, \dots, X_n$$

⇒ 確率変数

2. 実現された標本, 実現された観測値, 実現値

$$\Rightarrow x_1, x_2, \dots, x_n$$

⇒ 数値

1. 理論観測値 $X_1, X_2, \dots, X_n$ の関数 ⇒ 統計量
2. すべての i について, $\mu = E(X_i)$ と仮定する。
3. 母平均 μ の推定に使われる統計量 ⇒ μ の推定量

$$(a) \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \text{ は } \mu \text{ の推定量}$$

$$(b) S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \text{ は } \sigma^2 \text{ の推定量}$$

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
⇒ 推定値

$$(a) \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \text{ は } \mu \text{ の推定値}$$

$$(b) s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \text{ は } \sigma^2 \text{ の推定値}$$

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. $\hat{\alpha}, \hat{\beta}$ 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, $E(u_i) = 0$ とする。
3. すべての i について, $V(u_i) = \sigma^2$ とする。 ($V(u_i) = E(u_i^2) = \sigma^2$ に注意)
4. すべての $i \neq j$ について, $\text{Cov}(u_i, u_j) = 0$ とする。 ($\text{Cov}(u_i, u_j) = E(u_i u_j) = 0$ に注意)
5. すべての i について, $u_i \sim N(0, \sigma^2)$ とする。
6. $n \rightarrow \infty$ のとき, $\sum_{i=1}^n (X_i - \bar{X})^2 \rightarrow \infty$ とする。

攪乱項 $u_1, u_2, \dots, u_n$ 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)

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

特に, 回帰直線は,

$$E(Y_i) = \alpha + \beta X_i$$

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…、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…
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 $\hat{\alpha}$, $\hat{\beta}$ の統計的性質

準備：

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

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

ただし、

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i,$$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i,$$

$$\bar{u} = \frac{1}{n} \sum_{i=1}^n u_i,$$

とする。辺々を引いて、

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

を得る。

4.3.1 $\hat{\beta}$ について

β の最小二乗推定量 $\hat{\beta}$ に代入すると、

$$\hat{\beta} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

$$\begin{aligned} &= \frac{\sum_{i=1}^n (X_i - \bar{X}) \left(\beta(X_i - \bar{X}) + (u_i - \bar{u}) \right)}{\sum_{i=1}^n (X_i - \bar{X})^2} \\ &= \frac{\sum_{i=1}^n (X_i - \bar{X}) \left(\beta(X_i - \bar{X}) \right)}{\sum_{i=1}^n (X_i - \bar{X})^2} \\ &\quad + \frac{\sum_{i=1}^n (X_i - \bar{X})(u_i - \bar{u})}{\sum_{i=1}^n (X_i - \bar{X})^2} \\ &= \beta + \frac{\sum_{i=1}^n (X_i - \bar{X})u_i}{\sum_{i=1}^n (X_i - \bar{X})^2} \end{aligned}$$

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

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

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

4.3.2 $\hat{\alpha}$ について

α の最小二乗推定量 $\hat{\alpha}$ については、

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

ただし、 $\bar{u} = \frac{1}{n} \sum_{i=1}^n u_i$ である。 $\bar{Y} = \alpha + \beta\bar{X} + \bar{u}$ を途中で使う。

4.3.3 $\hat{\alpha}$, $\hat{\beta}$ 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。

$$E(a + bX) = a + bE(X),$$

となる。ただし、 a, b は定数とする。

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

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

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

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

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

の両辺に期待値をとると、

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

となり、 $\hat{\beta}$ は β の不偏推定量であると言える。

$\hat{\alpha}$ については、

$$\hat{\alpha} = \alpha - (\hat{\beta} - \beta)\bar{X} + \bar{u}$$

を利用して、辺々に期待値をとると、

$$\begin{aligned} E(\hat{\alpha}) &= \alpha - E(\hat{\beta} - \beta)\bar{X} + E(\bar{u}) \\ &= \alpha \end{aligned}$$

となる。 $E(\hat{\beta} - \beta) = 0$ に注意。また、 $E(\bar{u})$ の計算は以下のとおり。

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

$\hat{\alpha}$ は α の不偏推定量であると言える。

4.3.4 $\hat{\alpha}, \hat{\beta}$ の分散

統計学の復習 (分散の公式) :

1. X を確率変数とする。

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

となる。ただし、 $\mu = E(X)$ とする。

2. X を確率変数とする。

$$V(a + bX) = V(bX) = b^2 V(X),$$

となる。ただし、 a, b は定数とする。

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

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

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

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

$$\begin{aligned} V(\hat{\beta}) &= V\left(\beta + \sum_{i=1}^n \omega_i u_i\right) \\ &= V\left(\sum_{i=1}^n \omega_i u_i\right) \\ &= \sum_{i=1}^n \omega_i^2 E(u_i^2) \\ &= \sigma^2 \sum_{i=1}^n \omega_i^2 \\ &= \frac{\sigma^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \end{aligned}$$

誤差項 (または、攪乱項) の仮定より、

$$V(u_i) = \sigma^2,$$

を用いる。

最後の行は、 $\omega_i = \frac{X_i - \bar{X}}{\sum_{i=1}^n (X_i - \bar{X})^2}$ に注意して、

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

を用いる。

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

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

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

途中で、以下の計算が使われる。

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

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

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

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

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

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

となる。

数値例：

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	$\sum Y_i$ 35	$\sum X_i$ 52	$\sum X_i Y_i$ 468	$\sum X_i^2$ 696
平均	$\bar{Y}$ 8.75	$\bar{X}$ 13		

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

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

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

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 $\hat{\alpha}, \hat{\beta}$ となる。

すなわち、 $\hat{\alpha}, \hat{\beta}$ は、

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

を満たす。

さらに、

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

行列表示によって、

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

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

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

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。

$$\begin{aligned}
&\begin{pmatrix} V(\hat{\alpha}) & \text{Cov}(\hat{\alpha}, \hat{\beta}) \\ \text{Cov}(\hat{\alpha}, \hat{\beta}) & V(\hat{\beta}) \end{pmatrix} \\
&= \sigma^2 \begin{pmatrix} n & \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i^2 \end{pmatrix}^{-1} \\
&= \frac{\sigma^2}{n \sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2} \\
&\quad \times \begin{pmatrix} \sum_{i=1}^n X_i^2 & -\sum_{i=1}^n X_i \\ -\sum_{i=1}^n X_i & n \end{pmatrix} \\
&= \begin{pmatrix} \frac{\sigma^2 \sum_{i=1}^n X_i^2}{n \sum_{i=1}^n (X_i - \bar{X})^2} & -\frac{\sigma^2 \bar{X}}{\sum_{i=1}^n (X_i - \bar{X})^2} \\ -\frac{\sigma^2 \bar{X}}{\sum_{i=1}^n (X_i - \bar{X})^2} & \frac{\sigma^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \end{pmatrix}
\end{aligned}$$

4.3.5 $\hat{\alpha}, \hat{\beta}$ 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 $X_1, X_2, \dots, X_n$ が同一の分布に従うものとする。このとき、

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

となる。

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

$$\sum_{i=1}^n c_i X_i \sim N\left(E\left(\sum_{i=1}^n c_i X_i\right), V\left(\sum_{i=1}^n c_i X_i\right)\right)$$

となる。すなわち、

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

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

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

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

$\hat{\alpha}, \hat{\beta}$ の分布:

1. $\hat{\beta} = \beta + \sum_{i=1}^n \omega_i u_i$
2. $E(\hat{\beta}) = \beta$
3. $V(\hat{\beta}) = \frac{\sigma^2}{\sum_{i=1}^n (X_i - \bar{X})^2}$

よって,

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

となる。

1. $\hat{\alpha} = \alpha - (\hat{\beta} - \beta)\bar{X} + \bar{u}$
2. $E(\hat{\alpha}) = \alpha$
3. $V(\hat{\alpha}) = \sigma^2 \left(\frac{\bar{X}^2}{\sum_{i=1}^n (X_i - \bar{X})^2} + \frac{1}{n} \right)$

よって,

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

となる。

4.3.6 $\hat{\alpha}, \hat{\beta}$ 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): $\hat{\alpha}, \hat{\beta}$ 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 $\hat{\theta}$ を考える。

このとき,

$$E(\hat{\theta}) = \theta$$

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

$\hat{\theta}$ は不偏性を持つと言う。

$E(\hat{\theta}) - \theta$ は偏りと定義される。

(a) 標本平均 $\bar{X}$ は μ の不偏推定量である。

証明:

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

このように, $E(\bar{X}) = \mu$ なので, 標本平均 $\bar{X}$ 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, $\hat{\theta}_1$ と $\hat{\theta}_2$ の 2 つの不偏推定量を考える。

このとき, $V(\hat{\theta}_1) \leq V(\hat{\theta}_2)$ が成り立つとき, $\hat{\theta}_1$ は $\hat{\theta}_2$ 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, $\hat{\theta}$ 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, $\hat{\theta}$ 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, $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ が最も小さな分散を持つ推定量となる。

$\Rightarrow$ 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 $\hat{\theta}$ を考える。 n 個の標本から構成された推定量を $\hat{\theta}^{(n)}$ と定義する。

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

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

$$\text{plim}_{n \rightarrow \infty} \hat{\theta} = \theta$$

と表現する。

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

(b) μ の推定量 $\bar{X}$ を調べる。

$$E(\bar{X}) = \mu$$

である。

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

となる。 $n \rightarrow \infty$ のとき、

$$V(\bar{X}) = \frac{\sigma^2}{n} \rightarrow 0$$

となるので、 $\bar{X}$ は μ の一致推定量であると言える。

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

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

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

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

ただし、 $\omega_i = \frac{(X_i - \bar{X})}{\sum_{i=1}^n (X_i - \bar{X})^2}$ とする。このように、 $\hat{\beta}$ 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。

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

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

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

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

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、

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

となる。 $\tilde{\beta}$ が不偏であるためには、

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

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、

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

を利用して、

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

$\tilde{\beta}$ 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、

$$\begin{aligned}\sum_{i=1}^n \omega_i d_i &= \frac{\sum_{i=1}^n (X_i - \bar{X}) d_i}{\sum_{i=1}^n (X_i - \bar{X})^2} \\ &= \frac{\sum_{i=1}^n X_i d_i - \bar{X} \sum_{i=1}^n d_i}{\sum_{i=1}^n (X_i - \bar{X})^2} = 0\end{aligned}$$

を得る。

まとめると、 $\tilde{\beta}$ の分散は、

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

となる。 $\hat{\beta}$ の分散は、

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

なので、

$$V(\tilde{\beta}) \geq V(\hat{\beta})$$

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、 $d_1 = d_2 = \dots = d_n = 0$ のときとなり、これは $\hat{\beta}$ に一致する。

よって、 $\hat{\beta}$ 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。

⇒ ガウス＝マルコフの定理

$\hat{\alpha}$ についても、同様に、 α の最小分散線型不偏推定量となる。

証明は、

$$\hat{\alpha} - \alpha = -(\hat{\beta} - \beta)\bar{X} + \bar{u}$$

を利用すればよい。

推定量の関係 ⇒

最小分散 (最良) 線型不偏推定量 ⊂ 線型不偏推定量 ⊂ 線型推定量 ⊂ 全推定量

一貫性: $E(\hat{\beta}) = \beta$ となることが分かった。

n が大きくなると、 $\hat{\beta}$ は β に近づくかどうかを調べる。

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

$n \rightarrow \infty$ のとき、 $V(\hat{\beta}) \rightarrow 0$ となれば、 $\hat{\beta}$ は β の一致推定量となる。

最小二乗法の仮定の一つに、「 $n \rightarrow \infty$ のとき、 $\sum_{i=1}^n (X_i - \bar{X})^2 \rightarrow \infty$ 」というものがあつた。この仮定は、「 $n \rightarrow \infty$ のとき、 $V(\hat{\beta}) \rightarrow 0$ 」を保証する。よって、 $\hat{\beta}$ は β の一致推定量である。

$\hat{\alpha}$ についても、同様に、 $E(\hat{\alpha}) = \alpha$ であることは分かっている。

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

となり、「 $n \rightarrow \infty$ のとき、 $V(\hat{\alpha}) \rightarrow 0$ 」となるので、 $\hat{\alpha}$ 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: $u_i \sim N(0, \sigma^2)$

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

u_i の分散 σ^2 の不偏推定量:

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

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

誤差項 (または、攪乱項) の母分散 σ^2 の不偏推定量 s^2 は、

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

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。

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

両辺を二乗する。

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

総和をとる。

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

期待値をとる。

$$\begin{aligned}
E\left(\sum_{i=1}^n u_i^2\right) &= nE(\hat{\alpha} - \alpha)^2 + E(\hat{\beta} - \beta)^2 \sum_{i=1}^n X_i^2 \\
&\quad + E\left(\sum_{i=1}^n \hat{u}_i^2\right) + 2nE((\hat{\alpha} - \alpha)(\hat{\beta} - \beta))\bar{X} \\
n\sigma^2 &= \frac{\sigma^2 \sum_{i=1}^n X_i^2}{\sum_{i=1}^n (X_i - \bar{X})^2} + \frac{\sigma^2 \sum_{i=1}^n X_i^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \\
&\quad + E\left(\sum_{i=1}^n \hat{u}_i^2\right) - \frac{2n\sigma^2 \bar{X}^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \\
&= 2\sigma^2 \left(\frac{\sum_{i=1}^n X_i^2 - n\bar{X}^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \right) + E\left(\sum_{i=1}^n \hat{u}_i^2\right) \\
&= 2\sigma^2 + E\left(\sum_{i=1}^n \hat{u}_i^2\right)
\end{aligned}$$

途中の計算には以下が使われる。

$$\begin{aligned}
E\left(\sum_{i=1}^n u_i^2\right) &= n\sigma^2 \\
E(\hat{\alpha} - \alpha)^2 &= \frac{\sigma^2 \sum_{i=1}^n X_i^2}{n \sum_{i=1}^n (X_i - \bar{X})^2} \\
E(\hat{\beta} - \beta)^2 &= \frac{\sigma^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \\
E((\hat{\alpha} - \alpha)(\hat{\beta} - \beta)) &= -\frac{\sigma^2 \bar{X}}{\sum_{i=1}^n (X_i - \bar{X})^2}
\end{aligned}$$

よって、

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

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 $Z_1, Z_2, \dots, Z_m$ 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、 $E(Y) = m$, $V(Y) = 2m$ となる。(証明略)

- 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)

- n 個の独立な確率変数 $X_1, X_2, \dots, X_n$ が同一の正規分布 $N(\mu, \sigma^2)$ に従うものとする。

- $\frac{X_i - \mu}{\sigma} \sim N(0, 1)$ なので、 $\left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi^2(1)$ となる。
 $\frac{X_1 - \mu}{\sigma}, \frac{X_2 - \mu}{\sigma}, \dots, \frac{X_n - \mu}{\sigma}$ 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。

- μ を $\bar{X}$ に置き換えると、

$$\sum_{i=1}^n \left(\frac{X_i - \bar{X}}{\sigma}\right)^2 \sim \chi^2(n-1)$$

となる。(証明は後述)

さらに、

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

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)。

- すなわち、

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

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

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、

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

となる。さらに、

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

なので,

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

を得る。

s^2 の一致性の証明: s^2 は

$$\begin{aligned} s^2 &= \frac{1}{n-2} \sum_{i=1}^n \hat{u}_i^2 \\ &= \frac{1}{n-2} \sum_{i=1}^n (Y_i - \hat{\alpha} - \hat{\beta}X_i)^2 \end{aligned}$$

と定義される。

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

$$\begin{aligned} E\left(\frac{(n-2)s^2}{\sigma^2}\right) &= n-2, \\ V\left(\frac{(n-2)s^2}{\sigma^2}\right) &= 2(n-2), \end{aligned}$$

となる。さらに, 書き直すと,

$$\begin{aligned} \frac{(n-2)^2}{\sigma^4} V(s^2) &= 2(n-2), \\ V(s^2) &= \frac{2\sigma^4}{n-2}, \end{aligned}$$

を得る。「 $E(s^2) = \sigma^2$ で, しかも, $n \rightarrow \infty$ のとき $V(s^2) \rightarrow 0$ 」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

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

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

i	Y_i	X_i	$X_i Y_i$	X_i^2	$\hat{Y}_i$	$\hat{u}_i$
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	$\sum Y_i$ 35	$\sum X_i$ 52	$\sum X_i Y_i$ 468	$\sum X_i^2$ 696	$\sum \hat{Y}_i$ 35	$\sum \hat{u}_i$ 0
平均	$\bar{Y}$ 8.75	$\bar{X}$ 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,

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

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

$\hat{\alpha}$ と $\hat{\beta}$ の分散は,

$$\begin{aligned} V(\hat{\alpha}) &= \sigma_{\hat{\alpha}}^2 = \frac{\sigma^2 \sum_{i=1}^n X_i^2}{n \sum_{i=1}^n (X_i - \bar{X})^2} \\ V(\hat{\beta}) &= \sigma_{\hat{\beta}}^2 = \frac{\sigma^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \end{aligned}$$

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, $\hat{\alpha}, \hat{\beta}$ の分散の不偏推定量を次のように得ることができる。

$$\begin{aligned} s_{\hat{\alpha}}^2 &= \frac{s^2 \sum_{i=1}^n X_i^2}{n \sum_{i=1}^n (X_i - \bar{X})^2} \\ s_{\hat{\beta}}^2 &= \frac{s^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \end{aligned}$$

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

$$\begin{aligned} s_{\hat{\alpha}} &= s \sqrt{\frac{\sum_{i=1}^n X_i^2}{n \sum_{i=1}^n (X_i - \bar{X})^2}} \\ s_{\hat{\beta}} &= \frac{s}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2}} \end{aligned}$$

として与えられる。

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

i	Y_i	X_i	$X_i Y_i$	X_i^2	$\hat{Y}_i$	$\hat{u}_i$
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	$\sum Y_i$ 35	$\sum X_i$ 52	$\sum X_i Y_i$ 468	$\sum X_i^2$ 696	$\sum \hat{Y}_i$ 35	$\sum \hat{u}_i$ 0
平均	$\bar{Y}$ 8.75	$\bar{X}$ 13				

$s^2 = 1.15$ なので,

$$\begin{aligned} s_{\hat{\beta}}^2 &= \frac{s^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \\ &= \frac{s^2}{\sum_{i=1}^n X_i^2 - n\bar{X}^2} \\ &= 0.05 \times 1.15 \\ &= 0.0575 \\ s_{\hat{\alpha}}^2 &= \frac{s^2 \sum_{i=1}^n X_i^2}{n \sum_{i=1}^n (X_i - \bar{X})^2} \\ &= \frac{s^2 \sum_{i=1}^n X_i^2}{n(\sum_{i=1}^n X_i^2 - n\bar{X}^2)} \\ &= 8.7 \times 1.15 \\ &= 10.005 \end{aligned}$$

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

$$\begin{aligned} s_{\hat{\beta}} &= \sqrt{0.0575} = 0.240 \\ s_{\hat{\alpha}} &= \sqrt{10.005} = 3.163 \end{aligned}$$

となる。

4.5 $\hat{\alpha}$, $\hat{\beta}$ 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 $X_1, X_2, \dots, X_n$ が同一の正規分布 $N(\mu, \sigma^2)$ に従うものとする。このとき,

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

となる。ただし, $c_1, c_2, \dots, c_n$ 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, $U = \frac{Z}{\sqrt{Y/m}}$ 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 $E(U) = 0$, $m > 2$ について $V(U) = \frac{m}{m-2}$ 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。($E(U) = 0$)

2. t 分布は, 標準正規分布より裾野の広い分布 (なぜなら, $V(U) = \frac{m}{m-2} > 1$)

3. $m \rightarrow \infty$ のとき, $t(m) \rightarrow N(0, 1)$ となる。(期待値は $m > 1$ について $E(U) = 0$, 分散は $V(U) = \frac{m}{m-2} \rightarrow 1$)

標本平均 $\bar{X}$ の分布: $X_1, X_2, \dots, X_n$ 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. $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$ なので, $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$ となる。
2. $\frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2} \sim \chi^2(n-1)$ である。(証明は略)
3. $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$ と $\frac{(n-1)S^2}{\sigma^2}$ は独立。(証明は略)
すなわち, $\bar{X}$ と S^2 は独立。
4. したがって,

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

を得る。

重要な結果は,

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

ただし, $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$, $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ 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。

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

4.5.2 $\hat{\beta}$ について:

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

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

$$E(\hat{\beta}) = \beta, \\ V(\hat{\beta}) = \frac{\sigma^2}{\sum_{i=1}^n (X_i - \bar{X})^2},$$

となるので、

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

を得る。変形すると、

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

となる。

さらに、

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

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

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

4.5.3 $\hat{\alpha}$ について :

また、 $\hat{\alpha}$ の平均、分散はそれぞれ、

$$E(\hat{\alpha}) = \alpha, \\ V(\hat{\alpha}) = \frac{\sigma^2 \sum_{i=1}^n X_i^2}{n \sum_{i=1}^n (X_i - \bar{X})^2},$$

となるので、

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

を得る。変形すると、

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

となる。

さらに、 σ を s で置き換えると、

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

となる。

4.5.4 まとめ :

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

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

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)

$\bar{X}$ の分布を利用して、 μ の信頼区間を求める。

1. $\bar{X}$ の分布は以下の通り。

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

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 $100 \times \frac{\alpha}{2} \%$ 点, $100 \times (1 - \frac{\alpha}{2}) \%$ 点の値とする。このとき、

$$\text{Prob}\left(t_{1-\alpha/2}(n-1) < \frac{\bar{X} - \mu}{S / \sqrt{n}} < t_{\alpha/2}(n-1)\right) = 1 - \alpha$$

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、

$$\begin{aligned} & \text{Prob}\left(\bar{X} - t_{\alpha/2}(n-1) \frac{S}{\sqrt{n}} \right. \\ & \quad \left. < \mu < \bar{X} + t_{\alpha/2}(n-1) \frac{S}{\sqrt{n}}\right) = 1 - \alpha \end{aligned}$$

となる。

5. μ が区間 $(\bar{X} - t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}}, \bar{X} + t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}})$ にある確率は $1 - \alpha$ である。
6. 推定量 $\bar{X}$, S^2 をその推定値 $\bar{x}$, s^2 で置き換える。ただし, $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$, $s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$ とする。
7. 区間 $(\bar{x} - t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}}, \bar{x} + t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}})$ を信頼係数 $1 - \alpha$ の信頼区間といい, $\bar{x} - t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}}$ を信頼下限, $\bar{x} + t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}}$ を信頼上限と呼ぶ。

4.6.2 α , β の区間推定 (信頼区間)

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

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

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

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

$$\text{Prob}\left(t_{1-\alpha/2}(n-2) < \frac{\hat{\beta} - \beta}{s_{\hat{\beta}}} < t_{\alpha/2}(n-2)\right) = 1 - \alpha,$$

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

$$\text{Prob}\left(-t_{\alpha/2}(n-2) < \frac{\hat{\beta} - \beta}{s_{\hat{\beta}}} < t_{\alpha/2}(n-2)\right) = 1 - \alpha,$$

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,

$$\begin{aligned} \text{Prob}\left(\hat{\beta} - t_{\alpha/2}(n-2)s_{\hat{\beta}} \right. \\ \left. < \beta < \right. \\ \left. \hat{\beta} + t_{\alpha/2}(n-2)s_{\hat{\beta}}\right) = 1 - \alpha, \end{aligned}$$

と表される。

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

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

となる。

同様に, 信頼係数 $1 - \alpha$ の α の信頼区間は,

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

となる。

数値例: 今までと同様に, 以下の数値例を取りあげる。

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。

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

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

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

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

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

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

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,

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

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,

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

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,

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

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

$\bar{X}$ 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 $H_0: \mu = \mu_0$ が正しいもとの分布は、

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

となる。

3. $\text{Prob}\left(t_{1-\alpha/2}(n-1) < \frac{\bar{X} - \mu_0}{S/\sqrt{n}} < t_{\alpha/2}(n-1)\right) = 1 - \alpha$
 $t_{\alpha/2}(n-1)$, $t_{1-\alpha/2}(n-1)$ をそれぞれ自由度 $n-1$ の t 分布の上から $100 \times \frac{\alpha}{2} \%$ 点, $100 \times \frac{1-\alpha}{2} \%$ 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. $-t_{\alpha/2}(n-1) > \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$, または,

$\frac{\bar{X} - \mu_0}{S/\sqrt{n}} > t_{\alpha/2}(n-1)$ 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。

$\Rightarrow$ 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) $\bar{X}$, S^2 を実績値で置き換えて,

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

を得る。

$$\text{ただし, } \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

とする。

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

$\frac{\bar{x} - \mu_0}{s/\sqrt{n}} > t_{\alpha/2}(n-1)$ 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,

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

の分布に従う。

よって, 検定の手順は,

1. 検定統計値

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

を計算する。

ただし,

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

$$\hat{\beta} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2},$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X},$$

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,

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

または,

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

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, 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 $\hat{\alpha}$ 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，

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

の分布に従う。

このときの t 統計量の値 (すなわち， $\frac{\hat{\beta}}{s/\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2}}$)

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)，

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

または，

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

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， $\hat{\beta} > 0$ が得られた場合

(a) t 値が，

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

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，

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

となった場合，帰無仮説 $H_0 : \beta = 0$ を棄却できず， $\hat{\beta} > 0$ 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， $\hat{\beta} < 0$ が得られた場合

(a) t 値が，

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

となった場合，帰無仮説 $H_0 : \beta = 0$ を棄却できず， $\hat{\beta} < 0$ 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，

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

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，以下の数値例を取りあげる。

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。

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

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

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 ($\alpha = 0.05$, $n = 4$)，

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

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。

$$\hat{\alpha} = 0.3, \\ s_{\hat{\alpha}} = \sqrt{10.005} = 3.163,$$

なので、 t 値は、

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

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、 $\hat{\alpha} = 0.3$, $\hat{\beta} = 0.65$, $s_{\hat{\alpha}} = \sqrt{10.0005} = 3.163$, $s_{\hat{\beta}} = \sqrt{0.0575} = 0.240$, $\frac{\hat{\alpha}}{s_{\hat{\alpha}}} = 0.095$, $\frac{\hat{\beta}}{s_{\hat{\beta}}} = 2.708$, $s^2 = 1.15$ (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、係数の推定値の下に括弧内は 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、係数の推定値の下に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 $(Y_i, X_{1i}, X_{2i}, \dots, X_{ki})$, $i = 1, 2, \dots, n$ を用いて、 k 変数の多重回帰モデルを考える。

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

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、 $u_1, u_2, \dots, u_n$ は互いに独立に、平均ゼロ、分散 σ^2 の正規分布に従う)。

$\beta_1, \beta_2, \dots, \beta_k$ 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 $S(\beta_1, \beta_2, \dots, \beta_k)$ を定義する。

$$S(\beta_1, \beta_2, \dots, \beta_k) \\ = \sum_{i=1}^n u_i^2 \\ = \sum_{i=1}^n (Y_i - \beta_1 X_{1i} - \beta_2 X_{2i} - \dots - \beta_k X_{ki})^2$$

このとき、

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

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

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

最小化のためには、

$$\frac{\partial S(\beta_1, \beta_2, \dots, \beta_k)}{\partial \beta_1} = 0 \\ \frac{\partial S(\beta_1, \beta_2, \dots, \beta_k)}{\partial \beta_2} = 0 \\ \vdots \\ \frac{\partial S(\beta_1, \beta_2, \dots, \beta_k)}{\partial \beta_k} = 0$$

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

すなわち、 $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k$ は、

$$\sum_{i=1}^n (Y_i - \hat{\beta}_1 X_{1i} - \hat{\beta}_2 X_{2i} - \dots - \hat{\beta}_k X_{ki}) X_{1i} = 0, \\ \sum_{i=1}^n (Y_i - \hat{\beta}_1 X_{1i} - \hat{\beta}_2 X_{2i} - \dots - \hat{\beta}_k X_{ki}) X_{2i} = 0, \\ \vdots \\ \sum_{i=1}^n (Y_i - \hat{\beta}_1 X_{1i} - \hat{\beta}_2 X_{2i} - \dots - \hat{\beta}_k X_{ki}) X_{ki} = 0,$$

を満たす。

さらに,

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

行列表示によって,

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

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

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

を解くことになる。⇒ コンピュータによって計算

5.1 推定量の性質

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

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

として表される。

このとき,

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

を証明することが出来る。(証明略)

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

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

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

$$V(\hat{\beta}_j) = \sigma_{\beta_j}^2,$$

として, その推定量を $s_{\beta_j}^2$ とする。

このとき,

$$\hat{\beta}_j \sim N(\beta_j, \sigma_{\beta_j}^2),$$

となり、標準化すると、

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

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、 $\hat{\beta}_j$ と s^2 の独立性から (証明略)、

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

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。

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

ただし、 $\hat{Y}_i = \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \dots + \hat{\beta}_k X_{ki}$, $Y_i = \hat{Y}_i + \hat{u}_i$ 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、 $\sum_{i=1}^n \hat{u}_i^2$ 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 $\bar{R}^2$ を用いる。

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

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

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

となる。さらに、

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

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

数値例： 今までと同じ数値例で、 $\bar{R}^2$ を計算する。

i	Y_i	X_i	$X_i Y_i$	X_i^2	$\hat{Y}_i$	$\hat{u}_i$
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	$\sum Y_i$ 35	$\sum X_i$ 52	$\sum X_i Y_i$ 468	$\sum X_i^2$ 696	$\sum \hat{Y}_i$ 35	$\sum \hat{u}_i$ 0
平均	$\bar{Y}$ 8.75	$\bar{X}$ 13				

まず R^2 は、

$$\begin{aligned} R^2 &= 1 - \frac{\sum \hat{u}_i^2}{\sum Y_i^2 - n\bar{Y}^2} \\ &= 1 - \frac{(-0.8)^2 + 0.9^2 + 0.6^2 + (-0.7)^2}{35 - 4 \times 8.75^2} \\ &= 1 - \frac{2.30}{10.75} \\ &= 0.786 \end{aligned}$$

となり、 $\bar{R}^2$ は、

$$\begin{aligned} \bar{R}^2 &= 1 - \frac{\sum \hat{u}_i^2 / (n-k)}{(\sum Y_i^2 - n\bar{Y}^2) / (n-1)} \\ &= 1 - \frac{2.30 / (4-2)}{10.75 / (4-1)} \\ &= 0.679 \end{aligned}$$

となる。

注意： R^2 や $\bar{R}^2$ 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)。

⇒ 関数型の選択

5.2 ダミー変数について

5.2.1 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,

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

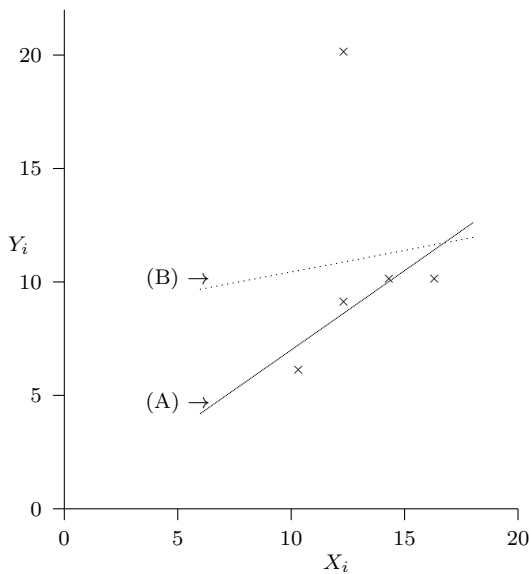
を推定する。 δ の推定値 $\hat{\delta}$ の有意性を調べることによって, 異常値かどうかの検定ができる。

数値例: 今までと同様に, 以下の数値例を取りあげる。

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。

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

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

ただし, 係数の推定値の下括弧内は 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,

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

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

5.2.2 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。

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

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	$D_i X_i$
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、

$$\beta = \frac{dY_i}{dX_i}$$

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。 $100 \times \frac{Y_i - Y_{i-1}}{Y_{i-1}}$ 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 $100 \times \frac{Y_i - Y_{i-1}}{Y_{i-1}} = \alpha + \beta X_i + u_i$ 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 $\bar{R}^2$ による比較はこの場合出来ない。

$\Rightarrow s^2$ 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、

$$\beta = \frac{d \log(Y_i)}{d \log(X_i)} = \frac{\frac{dY_i}{Y_i}}{\frac{dX_i}{X_i}} = \frac{100 \frac{dY_i}{Y_i}}{100 \frac{dX_i}{X_i}}$$

となる。

2 つ目の等号では、 $\frac{d \log(Y_i)}{dY_i} = \frac{1}{Y_i}$ が利用される。

3 つ目の等号の分子 $100 \frac{dY_i}{Y_i}$ や分母 $100 \frac{dX_i}{X_i}$ 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、

$$\begin{aligned} \log(Q_i) &= \beta'_1 + \beta_2 \log(K_i) + \beta_3 \log(L_i) \\ &= \beta'_1 + \beta_2 \log(K_i) + (1 - \beta_2) \log(L_i) + u_i \\ &= \beta'_1 + \beta_2 (\log(K_i) - \log(L_i)) + \log(L_i) + u_i, \end{aligned}$$

となるので、

$$\log(Q_i) - \log(L_i) = \beta'_1 + \beta_2 (\log(K_i) - \log(L_i)) + u_i,$$

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：

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

$\Rightarrow$ 平均費用と生産量との関係等

逆数：

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

$\Rightarrow$ 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、 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、 $\frac{\beta}{1-\gamma}$ 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。
⇒ 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 $u_1, u_2, \dots, u_n$ はそれぞれ独立に分布する」というものがあつた。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。
⇒ 時系列データのときのみ有効

$u_1, u_2, \dots, u_n$ 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、 $u_1, u_2, \dots, u_n$ 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、 $u_1, u_2, \dots, u_n$ 負の系列相関があると言う。

特徴： $u_1, u_2, \dots, u_i$ から u_{i+1} の符号が予想できる。⇒ 「 $u_1, u_2, \dots, u_n$ 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、 $\epsilon_1, \epsilon_2, \dots, \epsilon_n$ 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。

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

DW は近似的に、次のように表される。

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

以下の2つの近似が用いられる。

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

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

すなわち、 $\hat{\rho}$ は $\hat{u}_i$ と $\hat{u}_{i-1}$ 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 $\hat{u}_i, \hat{u}_{i-1}$ に置き換えて、 ρ の推定値 $\hat{\rho}$ を求める。

1. DW の値が2前後のとき、系列相関なし ($\hat{\rho} = 0$ 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	$X_i Y_i$	X_i^2	$\hat{Y}_i$	$\hat{u}_i$
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	$\sum Y_i$	$\sum X_i$	$\sum X_i Y_i$	$\sum X_i^2$	$\sum \hat{Y}_i$	$\sum \hat{u}_i$
	35	52	468	696	35	0
平均	$\bar{Y}$	$\bar{X}$				
	8.75	13				

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

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、 $\hat{\alpha} = 0.3$, $\hat{\beta} = 0.65$, $s_{\hat{\alpha}} = \sqrt{10.0005} = 3.163$, $s_{\hat{\beta}} = \sqrt{0.0575} = 0.240$, $\frac{\hat{\alpha}}{s_{\hat{\alpha}}} = 0.095$, $\frac{\hat{\beta}}{s_{\hat{\beta}}} = 2.708$, $s^2 = 1.15$ (すなわち、 $s = 1.07$), $R^2 = 0.786$, $\bar{R}^2 = 0.679$, $DW = 2.03$ を得た。これらをまとめて、

$$\begin{aligned} Y_i &= \begin{matrix} 0.3 \\ (0.095) \end{matrix} + \begin{matrix} 0.65 \\ (2.708) \end{matrix} X_i, \\ R^2 &= 0.786, \quad \bar{R}^2 = 0.679, \\ s &= 1.07, \quad DW = 2.03, \end{aligned}$$

ただし、係数の推定値の下括弧内は t 値を表すものとする。

または、

$$\begin{aligned} Y_i &= \begin{matrix} 0.3 \\ (3.163) \end{matrix} + \begin{matrix} 0.65 \\ (0.240) \end{matrix} X_i, \\ R^2 &= 0.786, \quad \bar{R}^2 = 0.679, \\ s &= 1.07, \quad DW = 2.03, \end{aligned}$$

ただし、係数の推定値の下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

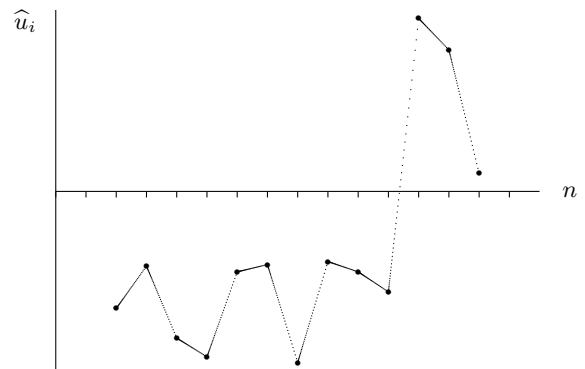
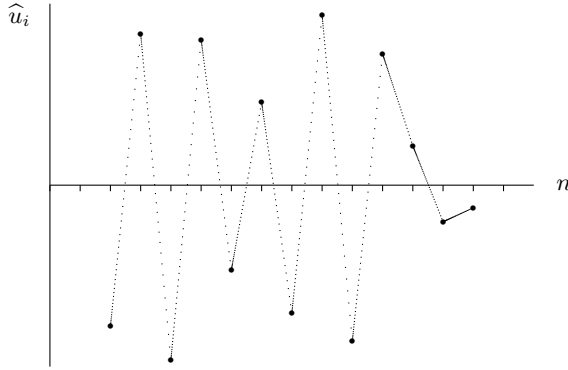


図 5： 負の系列相関



7.2 系列相関のもとで回帰式の推定

回帰式が

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

$$u_i = \rho u_{i-1} + \epsilon_i,$$

のときの推定を考える。ただし、 $\epsilon_1, \epsilon_2, \dots, \epsilon_n$ は互いに独立とする。

u_i を消去すると、

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

となり、

$$Y_i^* = (Y_i - \rho Y_{i-1}),$$

$$X_i^* = (X_i - \rho X_{i-1})$$

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。 $\epsilon_1, \epsilon_2, \dots, \epsilon_n$ 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。ただし、 $\alpha' = \alpha(1 - \rho)$ の関係が成り立つことに注意。

より一般的に、回帰式が

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

$$u_i = \rho u_{i-1} + \epsilon_i,$$

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
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
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
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
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
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

n	$k'=1$	$k'=2$	$k'=3$	$k'=4$	$k'=5$	$k'=6$	$k'=7$	$k'=8$	$k'=9$	$k'=10$	$k'=11$	$k'=12$	$k'=13$
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 1.690	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	0.147 3.266	—	—	—	—	—
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	0.200 3.111	0.127 3.360	—	—	—	—
15	1.077 1.361	0.946 1.543	0.814 1.759	0.685 1.971	0.562 2.220	0.447 2.471	0.343 2.727	0.251 2.979	0.175 3.216	0.111 3.438	—	—	—
16	1.106 1.371	0.982 1.536	0.860 1.747	0.730 1.963	0.615 2.194	0.508 2.385	0.399 2.691	0.272 2.975	0.193 3.304	0.123 3.441	—	—	—
17	1.136 1.381	0.982 1.536	0.860 1.747	0.730 1.963	0.615 2.194	0.508 2.385	0.399 2.691	0.272 2.975	0.193 3.304	0.123 3.441	—	—	—
18	1.158 1.391	1.046 1.535	0.933 1.696	0.802 1.872	0.710 2.060	0.603 2.257	0.502 2.461	0.407 2.668	0.321 2.873	0.244 3.073	0.172 3.265	0.087 3.557	—
19	1.180 1.401	1.074 1.537	0.967 1.685	0.859 1.848	0.752 2.023	0.649 2.206	0.549 2.396	0.456 2.589	0.369 2.783	0.290 2.974	0.220 3.159	0.160 3.335	0.078 3.603
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	0.593 2.339	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	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	0.548 2.461	0.461 2.633	0.386 2.806	0.307 2.976	0.240 3.141	0.182 3.300
22	1.239 1.429	1.147 1.541	1.053 1.664	0.958 1.797	0.863 1.940	0.769 2.060	0.677 2.246	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	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	0.628 2.322	0.545 2.514	0.465 2.670	0.391 2.826	0.322 2.979	0.259 3.129
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	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	1.288 1.454	1.206 1.550	1.123 1.651	1.038 1.767	0.953 1.886	0.868 2.013	0.784 2.144	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	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	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	1.316 1.468	1.241 1.556	1.161 1.653	1.084 1.751	1.000 1.881	0.925 1.987	0.844 2.108	0.763 2.235	0.687 2.367	0.604 2.501	0.530 2.636	0.459 2.769	0.400 2.890
28	1.331 1.473	1.259 1.560	1.181 1.654	1.107 1.743	1.028 1.884	0.955 1.993	0.871 2.071	0.798 2.168	0.723 2.306	0.646 2.430	0.574 2.565	0.510 2.680	0.445 2.805
29	1.345 1.483	1.270 1.563	1.198 1.650	1.124 1.743	1.050 1.881	0.975 1.994	0.900 2.052	0.826 2.164	0.753 2.278	0.681 2.396	0.612 2.513	0.544 2.630	0.479 2.754
30	1.352 1.489	1.284 1.567	1.214 1.650	1.143 1.739	1.071 1.833	0.998 1.920	0.928 2.018	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	1.363 1.496	1.297 1.570	1.229 1.650	1.160 1.735	1.090 1.825	1.021 1.909	0.972 2.004	0.904 2.102	0.836 2.203	0.769 2.306	0.703 2.411	0.638 2.518	0.576 2.625
32	1.373 1.502	1.321 1.577	1.258 1.651	1.193 1.730	1.127 1.813	1.061 1.900	0.994 1.991	0.927 2.085	0.861 2.181	0.796 2.281	0.731 2.382	0.667 2.484	0.606 2.588
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	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	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	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	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	0.991 2.041	0.908 2.144	0.845 2.236	0.783 2.330	0.722 2.425	0.662 2.521
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	1.009 2.041	0.930 2.127	0.868 2.216	0.808 2.306	0.748 2.398	0.689 2.492
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	1.026 2.026	0.947 2.112	0.885 2.203	0.825 2.295	0.765 2.387	0.714 2.484
38	1.427 1.535	1.374 1.593	1.319 1.656	1.261 1.721	1.203 1.790	1.143 1.864	1.084 1.961	1.039 2.029	0.959 2.115	0.897 2.206	0.837 2.298	0.777 2.379	0.724 2.465
39	1.435 1.540	1.382 1.597	1.338 1.658	1.273 1.722	1.218 1.789	1.166 1.859	1.104 1.932	1.047 2.007	0.969 2.085	0.932 2.164	0.874 2.256	0.819 2.350	0.763 2.413
40	1.442 1.544	1.391 1.600	1.358 1.659	1.285 1.721	1.230 1.786	1.175 1.854	1.120 1.924	1.064 1.997	0.990 2.072	0.952 2.150	0.896 2.228	0.840 2.320	0.785 2.391
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	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	1.503 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	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	1.528 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	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	1.549 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	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	1.567 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	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	1.583 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	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	1.598 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	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	1.611 1.662	1.586 1.688	1.560 1.715	1.534 1.743	1.507 1.772	1.480 1.801	1.451 1.830	1.423 1.861	1.393 1.893	1.369 1.925	1.340 1.957	1.312 1.977	1.285 2.024
85	1.623 1.670	1.601 1.693	1.576 1.718	1.551 1.747	1.524 1.776	1.497 1.804	1.469 1.834	1.445 1.864	1.415 1.891	1.390 1.921	1.367 1.951	1.343 1.977	1.320 2.004
90	1.635 1.679	1.612 1.703	1.589 1.726	1.566 1.751	1.542 1.776	1.518 1.802	1.494 1.827	1.469 1.854	1.445 1.881	1.420 1.909	1.399 1.937	1.369 1.966	1.344 1.995
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	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	1.654 1.694	1.634 1.715	1.613 1.736	1.592 1.758	1.571 1.780	1.550 1.826	1.528 1.826	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	1.720 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	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	1.758 1.779	1.748 1.780	1.738 1.799	1.728 1.809	1.718 1.820	1.707 1.831	1.697 1.841	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>

のときの推定を考える。ただし, $\epsilon_1, \epsilon_2, \dots, \epsilon_n$ は互いに独立とする。

u_i を消去すると,

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

となり,

$$\begin{aligned} Y_i^* &= (Y_i - \rho Y_{i-1}), \\ X_{1i}^* &= (X_{1i} - \rho X_{1,i-1}), \\ X_{2i}^* &= (X_{2i} - \rho X_{2,i-1}), \\ &\dots, \\ X_{ki}^* &= (X_{ki} - \rho X_{k,i-1}) \end{aligned}$$

を新たな変数として,

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

最小二乗法を適用する。 $\epsilon_1, \epsilon_2, \dots, \epsilon_n$ 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。

ρ の求め方について: DW は近似的に $DW \approx 2(1 - \hat{\rho})$ と表されるので, DW から ρ の推定値 $\hat{\rho}$ を逆算して,

$$\begin{aligned} Y_i^* &= (Y_i - \hat{\rho} Y_{i-1}), \\ X_{1i}^* &= (X_{1i} - \hat{\rho} X_{1,i-1}), \\ X_{2i}^* &= (X_{2i} - \hat{\rho} X_{2,i-1}), \\ &\dots, \\ X_{ki}^* &= (X_{ki} - \hat{\rho} X_{k,i-1}) \end{aligned}$$

を新たな変数として,

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

に最小二乗法を適用する。

8 不均一分散 (不等分散)

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 ,

$u_2, \dots, u_n$ 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。

$$\begin{aligned}\frac{Y_i}{z_i} &= \alpha \frac{1}{z_i} + \beta \frac{X_i}{z_i} + \frac{u_i}{z_i} \\ &= \alpha \frac{1}{z_i} + \beta \frac{X_i}{z_i} + u_i^*\end{aligned}$$

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)。

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

u_i の仮定 $E(u_i) = 0$ が使われている。

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

u_i の仮定 $V(u_i) = \sigma_*^2 z_i^2$ が最後に使われている。

よって、 $\frac{Y_i}{z_i}, \frac{1}{z_i}, \frac{X_i}{z_i}$ 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

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

を推定し、 γ の推定値 $\hat{\gamma}$ 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、

$$\begin{aligned}\frac{Y_i}{X_i} &= \alpha \frac{1}{X_i} + \beta + \frac{u_i}{X_i} \\ &= \alpha \frac{1}{X_i} + \beta + u_i^*\end{aligned}$$

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。

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 $\hat{\alpha}, \hat{\beta}$ の組み合わせは無数に存在する。この意味で、 $\hat{\alpha}, \hat{\beta}$ 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 $\hat{\alpha}, \hat{\beta}$ 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。

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

行列表示により、

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

$\hat{\alpha}, \hat{\beta}$ について表すと、

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

逆行列を計算して、

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

完全な多重共線性の場合 ($W_i = \gamma X_i$ の場合),

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

となる。

また、

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

となるので、完全な多重共線性の場合、推定値の分散が無大となる。推定値の分散が無大という意味は、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、 $\bar{W} = \frac{1}{n} \sum W_i = 0$, $\bar{X} = \frac{1}{n} \sum X_i = 0$ とする。 W_i と X_i との相関係数を r とすると、

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

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

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

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)、 $V(\hat{\alpha})$, $V(\hat{\beta})$ は大きくなるということを意味する。

⇒ 係数の推定値の有意性が低くなる。

⇒ 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 $\bar{R}^2$) 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 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。

$$\text{帰無仮説 } H_0: \beta_2 + \beta_3 = 1,$$

$$\text{対立仮説 } H_1: \beta_2 + \beta_3 \neq 1,$$

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、傾き共に変化すると想定した場合、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,

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

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

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

において, パラメータ $\beta_1, \beta_2, \dots, \beta_k$ 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 $\hat{u}_i$ とする。

制約を含めて推定されたときの残差を $\tilde{u}_i$ とする。

すなわち,

帰無仮説 $H_0: \beta_{k-G+1} = \dots = \beta_k = 0$,

対立仮説 $H_1: H_0$ でない。

を検定する場合,

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

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

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

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

1. H_0 が真のとき, $\frac{\sum \tilde{u}_i^2 - \sum \hat{u}_i^2}{\sigma^2} \sim \chi^2(G)$ となる。(証明略)
2. また, $\frac{\sum \hat{u}_i^2}{\sigma^2} \sim \chi^2(n-k)$ となる。(証明略)
3. さらに, $\frac{\sum \tilde{u}_i^2 - \sum \hat{u}_i^2}{\sigma^2}$ と $\frac{\sum \hat{u}_i^2}{\sigma^2}$ とは独立に分布する。(証明略)
4. したがって, この場合,

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

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:

$$\log\left(\frac{Q_i}{L_i}\right) = \beta'_1 + \beta_2 \log\left(\frac{K_i}{L_i}\right) + u_i,$$

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。

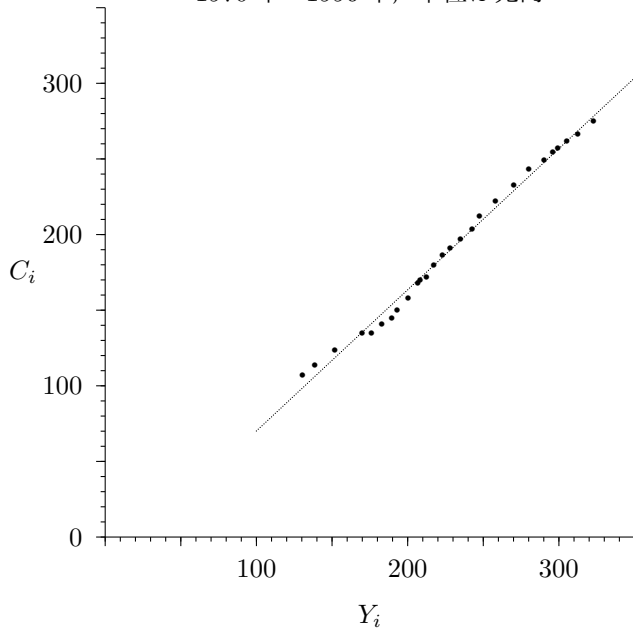
表 3: 所得と消費のデータ

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

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.
$$\text{実質値} = \frac{\text{名目値}}{\text{物価指数}}$$
 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、
$$\text{実質値} = \frac{\text{名目値}}{\text{物価指数}/100}$$
 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。

図 1: 実質消費 (縦軸) と実質所得 (横軸)

1970 年～1996 年, 単位は兆円



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ー」は 100 なので, 基準年次は 1990 年となる。
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いう。

まず最初に,

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

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。

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

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

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

ただし, 係数の推定値の下に括弧内は, 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 $\hat{\alpha}$ は -23216.7 で負, 限界消費性向の推定値 $\hat{\beta}$ 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. $\hat{\alpha}, \hat{\beta}$ 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。

⇒ 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, ……

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 $\bar{R}^2 = 0.992102$ 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。

⇒ 最小二乗法の仮定を満たしていない。

⇒ $t(25)$ 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 $\hat{\rho}$ は、最小二乗法の推定結果の DW を用いて、 $\hat{\rho} = 1 - \frac{DW}{2} = 1 - \frac{.2898375}{2} = .855081$ を得る。データを、次のように変換する。

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

そして、

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

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。

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

ただし、係数の推定値の下に括弧内は、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 $\hat{\alpha}$ は、 $\hat{\alpha} = \frac{\hat{\alpha}'}{1 - \hat{\rho}}$ によって求められ、 $\frac{-3679.78}{1 - .855081} = -25392.0$ で負、限界消費性向の推定値 $\hat{\beta}$ 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 $\hat{\alpha}'$ 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 $\bar{R}^2$ 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。

⇒ 「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は異なるため、 $\bar{R}^2$ の比較は不適當。

⇒ 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、 $1.32 \leq DW < 1.47$ であるので、系列相関の有無を判定できない。

⇒ 「明らかに誤差項に系列相関がある」とは言えない。

⇒ 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
-----+-----+-----+-----+-----+-----+-----
      Model |  6.7845e+10       1   6.7845e+10          F( 1, 25) = 3267.05
      Residual |  519164248      25  20766569.9          Prob > F      = 0.0000
-----+-----+-----+-----+-----+-----
      Total |  6.8365e+10      26  2.6294e+09          R-squared      = 0.9924
                                          Adj R-squared   = 0.9921
                                          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*1.rcons
(1 missing value generated)

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

. reg drcons dryd

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

$$DRCONS_i = RCONS_i - \hat{\rho} RCONS_{i-1}$$

$$DRY_i = RY_i - \hat{\rho} RY_{i-1}$$

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

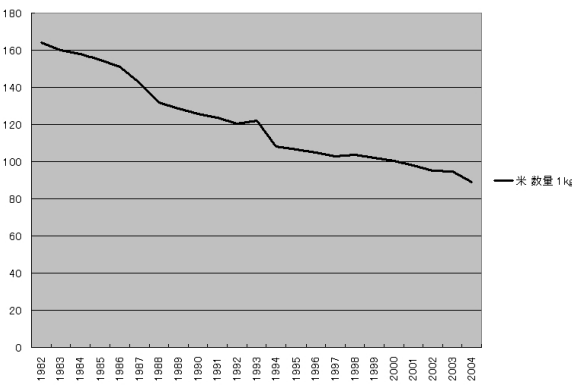
11.2 ミクロの消費関数 (需要関数)

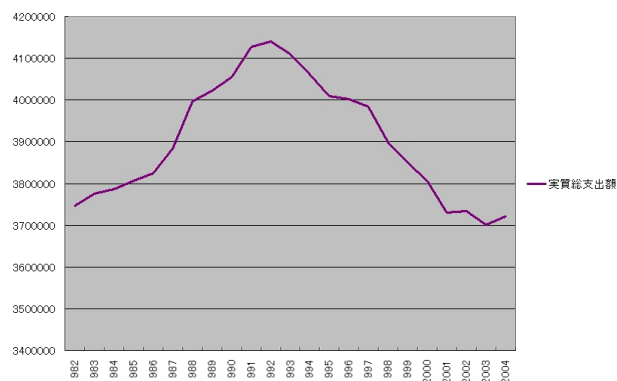
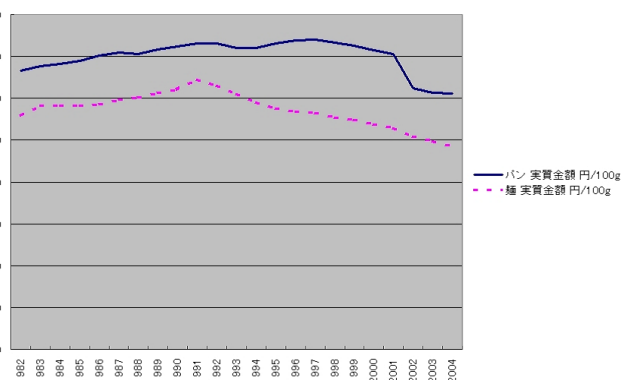
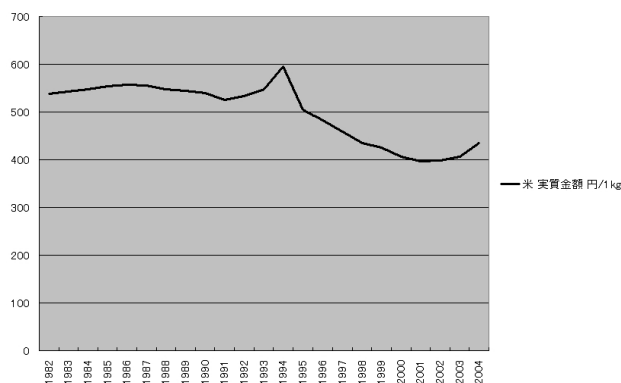
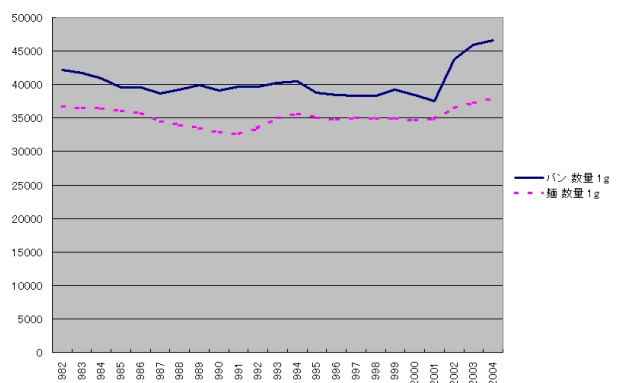
米, パン, 麵の選択を考える。

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		麵		総支出 額	消費者 物価指数 (全国総合)
	数量 1kg	金額 1kg	数量 1g	金額 100g	数量 1g	金額 100g		
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

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

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

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

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

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

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

●対数線型

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

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

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

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

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

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

●対数線型 (変数を減らす)

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

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

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

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

● 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 =
Model	11605.1437	4	2901.28592	23
Residual	837.923294	18	46.5512941	F(4, 18) = 62.32
Total	12443.067	22	565.593953	Prob > F = 0.0000
				R-squared = 0.9327
				Adj R-squared = 0.9177
				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 =
Model	113479303	4	28369825.8	23
Residual	8424108.48	18	468006.027	F(4, 18) = 60.62
Total	121903412	22	5541064.17	Prob > F = 0.0000
				R-squared = 0.9309
				Adj R-squared = 0.9155
				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 =
Model	37006848.3	4	9251712.07	23
Residual	4782689.62	18	265704.979	F(4, 18) = 34.82
Total	41789537.9	22	1899524.45	Prob > F = 0.0000
				R-squared = 0.8856
				Adj R-squared = 0.8601
				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 =
Model	.766574852	4	.191643713	23
Residual	.04086074	18	.002270041	F(4, 18) = 84.42
Total	.807435592	22	.036701618	Prob > F = 0.0000
				R-squared = 0.9494
				Adj R-squared = 0.9381
				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 =
Model	.064720875	4	.016180219	23
Residual	.004742568	18	.000263476	F(4, 18) = 61.41
Total	.069463442	22	.003157429	Prob > F = 0.0000
				R-squared = 0.9317
				Adj R-squared = 0.9166
				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 =
Model	.029851208	4	.007462802	23
Residual	.003971008	18	.000220612	F(4, 18) = 33.83
Total	.033822216	22	.001537373	Prob > F = 0.0000
				R-squared = 0.8826
				Adj R-squared = 0.8565
				Root MSE = .01485

```

      lq3 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-----+-----
      lre | -.1868154   .158742   -1.18   0.255   -.52032   .166892
      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 =
Model	.766565902	3	.255521967	23
Residual	.040869691	19	.002151036	F(3, 19) = 118.79
Total	.807435592	22	.036701618	Prob > F = 0.0000
				R-squared = 0.9494
				Adj R-squared = 0.9414
				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 =
Model	.029835392	3	.009945131	23
Residual	.003986824	19	.000209833	F(3, 19) = 47.40
Total	.033822216	22	.001537373	Prob > F = 0.0000
				R-squared = 0.8821
				Adj R-squared = 0.8635
				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

普通国債(東京店随最長期物)



$$Kabui_i = - \underset{(2.58)}{1883.1} + \underset{(8.27)}{46.3} \text{ExRate}_i + \underset{(51.7)}{6802.6} R_i$$

$$s = 2309.9, \quad R^2 = 0.570, \quad \bar{R}^2 = 0.570, \quad DW = 0.021$$

$$Kabui_i = - \underset{(0.03)}{2.36} + \underset{(0.31)}{0.164} \text{ExRate}_i + \underset{(1.96)}{36.5} R_i +$$

$$\underset{(482)}{0.995} Kabui_{i-1}$$

$$s = 214.8, \quad R^2 = 0.996, \quad \bar{R}^2 = 0.996, \quad DW = 2.100$$

$$\Delta Kabui_i = \underset{(0.119)}{8.11} - \underset{(0.173)}{0.091} \text{ExRate}_i - \underset{(0.070)}{0.862} R_i$$

$$s = 215.1, \quad R^2 = 0.000015, \quad \bar{R}^2 = -0.00097, \quad DW = 2.106$$

● 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
Total | 2.5123e+10  2027  12394036.8              Adj R-squared = 0.5695
                                           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
```

11.3 株価，金利，為替レート

東証株価・日経225種平均・終値・円



東京・終値・円/ドル



```

      _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
-----+-----+-----+-----+-----+-----
      Model | 2.4997e+10           3    8.3323e+09      F( 3, 2023) = .
      Residual | 93298782.6        2023    46119.0225      Prob > F      = 0.0000
      Total | 2.5090e+10        2026   12384078.2      R-squared      = 0.9963
                                          Adj R-squared  = 0.9963
                                          Root MSE      = 214.75

      kabu |         Coef.      Std. Err.      t    P>|t|     [95% Conf. Interval]
-----+-----+-----+-----+-----+-----
      exrate |   .1643596       .5295945      0.31   0.756    -1.8742481    1.202967
      r      |   36.51834       18.6521      1.96   0.050    -1.060989    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
-----+-----+-----+-----+-----+-----
      Model | 1443.23541           2    721.617703      F( 2, 2024) = 0.02
      Residual | 93624927.9        2024    46257.3754      Prob > F      = 0.9845
      Total | 93626371.1        2026    46212.424      R-squared      = 0.0000
                                          Adj R-squared  = -0.0010
                                          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) : $x_1, x_2, \dots, x_n$
- ・ 背後に対応する確率変数を仮定 : $X_1, X_2, \dots, X_n$
- ・ $E(X_i) = \mu$, $V(X_i) = \sigma^2$ を仮定

母数 (μ, σ^2) を推定する。

観測データ $x_1, x_2, \dots, x_n$ 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 $\hat{\mu}$ とすると,

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

となり, $\hat{\mu} \equiv \bar{x}$ を得る。

すなわち,

$$\frac{d \sum_{i=1}^n (x_i - \mu)^2}{d\mu} = 0$$

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,

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

となり, $\hat{\mu} \equiv \bar{X}$ を得る ($\hat{\mu}$ 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 $X_1, X_2, \dots, X_n$ 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。

$X_1, X_2, \dots, X_n$ の結合分布は, 互いに独立なので,

$$f(x_1, x_2, \dots, x_n; \theta) \equiv \prod_{i=1}^n f(x_i; \theta)$$

と表される。

観測データ $x_1, x_2, \dots, x_n$ 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 $\hat{\theta} = \hat{\theta}(x_1, x_2, \dots, x_n)$ と呼ぶ。

データ $x_1, x_2, \dots, x_n$ を確率変数 $X_1, X_2, \dots, X_n$ で置き換えて、 $\hat{\theta} = \hat{\theta}(X_1, X_2, \dots, X_n)$ 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、

$$\hat{\theta} \sim N(\theta, \sigma_{\theta}^2)$$

ただし、

$$\begin{aligned} \sigma_{\theta}^2 &= \frac{1}{\sum_{i=1}^n E\left[\left(\frac{d \log f(X_i; \theta)}{d\theta}\right)^2\right]} \\ &= -\frac{1}{\sum_{i=1}^n E\left[\frac{d^2 \log f(X_i; \theta)}{d\theta^2}\right]} \end{aligned}$$

θ がベクトル ($k \times 1$) の場合、 n が大きいとき、

$$\hat{\theta} \sim N(\theta, \Sigma_{\theta})$$

ただし、

$$\begin{aligned} \Sigma_{\theta} &= \left(\sum_{i=1}^n E\left[\left(\frac{\partial \log f(X_i; \theta)}{\partial \theta}\right) \left(\frac{\partial \log f(X_i; \theta)}{\partial \theta}\right)'\right] \right)^{-1} \\ &= -\left(\sum_{i=1}^n E\left[\frac{\partial^2 \log f(X_i; \theta)}{\partial \theta \partial \theta'}\right] \right)^{-1} \end{aligned}$$

例 1： 正規母集団 $N(\mu, \sigma^2)$ からの標本値 $x_1, x_2, \dots, x_n$ 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】 $N(\mu, \sigma^2)$ の密度関数は、

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x - \mu)^2\right)$$

となる。したがって、互いに独立な $X_1, X_2, \dots, X_n$ の結合分布は、

$$\begin{aligned} f(x_1, x_2, \dots, x_n; \mu, \sigma^2) &= \prod_{i=1}^n f(x_i; \mu, \sigma^2) \\ &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x_i - \mu)^2\right) \\ &= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\right) \end{aligned}$$

となる。

- (1) σ^2 が既知のとき、尤度関数 $l(\mu)$ は、

$$l(\mu) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\right)$$

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、

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

となる μ を求める。 μ の解を $\hat{\mu}$ とすると、 μ の最尤推定値は、

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i \equiv \bar{x}$$

を得る。

さらに、観測値 $x_1, x_2, \dots, x_n$ をその確率変数 $X_1, X_2, \dots, X_n$ で置き換えて、 μ の最尤推定量は、

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n X_i \equiv \bar{X}$$

となる。

$\hat{\mu}$ の分散を求めるために、

$$\log f(X_i; \mu) = -\frac{1}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} (X_i - \mu)^2$$

$$\frac{d \log f(X_i; \mu)}{d\mu} = \frac{1}{\sigma^2} (X_i - \mu)$$

$$\left(\frac{d \log f(X_i; \mu)}{d\mu} \right)^2 = \frac{1}{\sigma^4} (X_i - \mu)^2$$

$$E \left[\left(\frac{d \log f(X_i; \mu)}{d\mu} \right)^2 \right] = \frac{1}{\sigma^4} E[(X_i - \mu)^2] = \frac{1}{\sigma^2}$$

と計算される。

最尤推定量の性質から、 n が大きいとき、

$$\hat{\mu} \sim N(\mu, \sigma_\mu^2)$$

ただし、

$$\sigma_\mu^2 = \frac{1}{\sum_{i=1}^n E \left[\left(\frac{d \log f(X_i; \mu)}{d\mu} \right)^2 \right]} = \frac{\sigma^2}{n}$$

この場合は、 n の大きさに関わらず、 $\hat{\mu} \sim N(\mu, \sigma_\mu^2)$ が成り立つ。

(2) σ^2 が未知のとき、 μ と σ^2 の尤度関数は、

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

となる。

対数尤度関数は、

$$\begin{aligned} \log l(\mu, \sigma^2) &= -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \sigma^2 \\ &\quad - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \end{aligned}$$

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 $\hat{\mu}, \hat{\sigma}^2$ とすると、最尤推定値は、

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i \equiv \bar{x}$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{\mu})^2 \equiv \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

となる。

観測値 $x_1, x_2, \dots, x_n$ をその確率変数 $X_1, X_2, \dots, X_n$ で置き換えて、 μ, σ^2 の最尤推定量は、

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n X_i \equiv \bar{X}$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \hat{\mu})^2 \equiv \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

となる。

σ^2 の最尤推定量 $\hat{\sigma}^2$ は、 σ^2 の不偏推定量 $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ とは異なることに注意。

$\theta = (\mu, \sigma^2)'$ とする。 n が大きいとき、

$$\hat{\theta} \sim N(\theta, \Sigma_\theta)$$

ただし、

$$\Sigma_\theta = - \left(\sum_{i=1}^n E \left[\frac{\partial^2 \log f(X_i; \theta)}{\partial \theta \partial \theta'} \right] \right)^{-1}$$

$$\log f(X_i; \theta) = -\frac{1}{2} \log(2\pi) - \frac{1}{2} \log(\sigma^2) - \frac{1}{2\sigma^2} (X_i - \mu)^2$$

$$\begin{aligned} \frac{\partial \log f(X_i; \theta)}{\partial \theta} &= \begin{pmatrix} \frac{\partial \log f(X_i; \theta)}{\partial \mu} \\ \frac{\partial \log f(X_i; \theta)}{\partial \sigma^2} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{\sigma^2} (X_i - \mu) \\ -\frac{1}{2\sigma^2} + \frac{1}{2\sigma^4} (X_i - \mu)^2 \end{pmatrix} \end{aligned}$$

$$\begin{aligned}
& \frac{\partial^2 \log f(X_i; \theta)}{\partial \theta \partial \theta'} \\
&= \begin{pmatrix} \frac{\partial^2 \log f(X_i; \theta)}{\partial \mu^2} & \frac{\partial^2 \log f(X_i; \theta)}{\partial \mu \partial \sigma^2} \\ \frac{\partial^2 \log f(X_i; \theta)}{\partial \sigma^2 \partial \mu} & \frac{\partial^2 \log f(X_i; \theta)}{\partial (\sigma^2)^2} \end{pmatrix} \\
&= \begin{pmatrix} -\frac{1}{\sigma^2} & -\frac{1}{\sigma^4}(X_i - \mu) \\ -\frac{1}{\sigma^4}(X_i - \mu) & \frac{1}{2\sigma^4} - \frac{1}{\sigma^6}(X_i - \mu)^2 \end{pmatrix} \\
& \mathbb{E} \left[\frac{\partial^2 \log f(X_i; \theta)}{\partial \theta \partial \theta'} \right] \\
&= \begin{pmatrix} -\frac{1}{\sigma^2} & -\frac{1}{\sigma^4} \mathbb{E}(X_i - \mu) \\ -\frac{1}{\sigma^4} \mathbb{E}(X_i - \mu) & \frac{1}{2\sigma^4} - \frac{1}{\sigma^6} \mathbb{E}[(X_i - \mu)^2] \end{pmatrix} \\
&= \begin{pmatrix} -\frac{1}{\sigma^2} & 0 \\ 0 & -\frac{1}{2\sigma^4} \end{pmatrix}
\end{aligned}$$

よって,

$$\begin{aligned}
\Sigma_\theta &= - \left(\sum_{i=1}^n \mathbb{E} \left[\frac{\partial^2 \log f(X_i; \theta)}{\partial \theta \partial \theta'} \right] \right)^{-1} \\
&= \begin{pmatrix} \frac{\sigma^2}{n} & 0 \\ 0 & \frac{2\sigma^4}{n} \end{pmatrix}
\end{aligned}$$

まとめると, μ, σ^2 の最尤推定量 $\hat{\mu} = (1/n) \sum_{i=1}^n X_i$, $\hat{\sigma}^2 = (1/n) \sum_{i=1}^n (X_i - \bar{X})^2$ の分布は, n が大きいとき,

$$\begin{pmatrix} \hat{\mu} \\ \hat{\sigma}^2 \end{pmatrix} \sim N \left(\begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix}, \begin{pmatrix} \frac{\sigma^2}{n} & 0 \\ 0 & \frac{2\sigma^4}{n} \end{pmatrix} \right)$$

となる。

例 2: $X_1, X_2, \dots, X_n$ 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,

$$f(x; p) = p^x (1-p)^{1-x} \quad x = 0, 1$$

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,

$$\begin{aligned}
\log l(p) &= \sum_{i=1}^n \log f(x_i; p) \\
&= \log(p) \sum_{i=1}^n x_i + \log(1-p) \sum_{i=1}^n (1-x_i) \\
&= \log(p) \sum_{i=1}^n x_i + \log(1-p) \left(n - \sum_{i=1}^n x_i \right)
\end{aligned}$$

となる。

$\log l(p)$ を最大にする p を求める。

$$\frac{d \log l(p)}{dp} = \frac{1}{p} \sum_{i=1}^n x_i - \frac{1}{1-p} \left(n - \sum_{i=1}^n x_i \right) = 0$$

したがって, p について解くと, p の最尤推定値 $\hat{p}$ は,

$$\hat{p} = \frac{1}{n} \sum_{i=1}^n x_i$$

となる。

さらに, x_i を X_i で置き換えて, p の最尤推定量 $\hat{p}$ は,

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

となる。

$\hat{p}$ の分布を求める。

$$\log f(X_i; p) = X_i \log(p) + (1-X_i) \log(1-p)$$

$$\frac{d \log f(X_i; p)}{dp} = \frac{X_i}{p} - \frac{1-X_i}{1-p} = \frac{X_i - p}{p(1-p)}$$

$$\mathbb{E} \left[\left(\frac{d \log f(X_i; p)}{dp} \right)^2 \right] = \frac{\mathbb{E}[(X_i - p)^2]}{p^2(1-p)^2}$$

$$\mathbb{E}[(X_i - p)^2] = \sum_{x_i=0}^1 (x_i - p)^2 f(x_i; p)$$

$$= \sum_{x_i=0}^1 (x_i - p)^2 p^{x_i} (1-p)^{1-x_i}$$

$$= p^2(1-p) + (1-p)^2 p = p(1-p)$$

$$\frac{1}{\sum_{i=1}^n \mathbb{E} \left[\left(\frac{d \log f(X_i; p)}{dp} \right)^2 \right]} = \frac{p(1-p)}{n}$$

したがって,

$$\hat{p} \sim N \left(p, \frac{p(1-p)}{n} \right)$$

を得る。

例 3: $X_1, X_2, \dots, X_n$ 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、

$$f(x; \lambda) = \frac{\lambda^x e^{-\lambda}}{x!} \quad x = 0, 1, 2, \dots$$

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 λ を求める。

$$\frac{d \log l(\lambda)}{d\lambda} = \frac{1}{\lambda} \sum_{i=1}^n x_i - n = 0$$

したがって、 λ について解くと、 λ の最尤推定値 $\hat{\lambda}$ は、

$$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^n x_i$$

となる。

さらに、 x_i を X_i で置き換えて、 λ の最尤推定量 $\hat{\lambda}$ は、

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

となる。

$\hat{\lambda}$ の分布を求める。

$$\log f(X_i; \lambda) = X_i \log(\lambda) - \lambda - \log(X_i!)$$

$$\frac{d \log f(X_i; \lambda)}{d\lambda} = \frac{X_i}{\lambda} - 1$$

$$\frac{d^2 \log f(X_i; \lambda)}{d\lambda^2} = -\frac{X_i}{\lambda^2}$$

$$E\left(\frac{d^2 \log f(X_i; \lambda)}{d\lambda^2}\right) = \frac{E(X_i)}{\lambda^2}$$

$$\begin{aligned} E(X_i) &= \sum_{x=0}^{\infty} x f(x; \lambda) \\ &= \sum_{x=0}^{\infty} x \frac{\lambda^x e^{-\lambda}}{x!} \\ &= \sum_{x=1}^{\infty} x \frac{\lambda^x e^{-\lambda}}{x!} \\ &= \sum_{x=1}^{\infty} \lambda \frac{\lambda^{x-1} e^{-\lambda}}{(x-1)!} \\ &= \sum_{x=0}^{\infty} \lambda \frac{\lambda^x e^{-\lambda}}{x!} \\ &= \lambda \\ &\quad - \frac{1}{\sum_{i=1}^n E\left(\frac{d^2 \log f(X_i; \lambda)}{d\lambda^2}\right)} = \frac{\lambda}{n} \end{aligned}$$

したがって、

$$\hat{\lambda} \sim N\left(\lambda, \frac{\lambda}{n}\right)$$

を得る。

例 4: $X_1, X_2, \dots, X_n$ 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、

$$f(x; \lambda) = \lambda e^{-\lambda x} \quad x > 0$$

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 λ を求める。

$$\frac{d \log l(\lambda)}{d\lambda} = \frac{n}{\lambda} - \sum_{i=1}^n x_i = 0$$

したがって、 λ について解くと、 λ の最尤推定値 $\hat{\lambda}$ は、

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^n x_i}$$

となる。

さらに、 x_i を X_i で置き換えて、 λ の最尤推定量 $\hat{\lambda}$ は、

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

となる。

$\hat{\lambda}$ の分布を求める。

$$\log f(X_i; \lambda) = \log \lambda - \lambda X_i$$

$$\frac{d \log f(X_i; \lambda)}{d\lambda} = \frac{1}{\lambda} - X_i$$

$$\frac{d^2 \log f(X_i; \lambda)}{d\lambda^2} = -\frac{1}{\lambda^2}$$

$$-\frac{1}{\sum_{i=1}^n E\left(\frac{d^2 \log f(X_i; \lambda)}{d\lambda^2}\right)} = \frac{\lambda^2}{n}$$

したがって、

$$\hat{\lambda} \sim N\left(\lambda, \frac{\lambda^2}{n}\right)$$

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 $F(x) \equiv P(X < x)$ 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。

$$\begin{aligned} G(y) &= P(Y < y) = P(aX + b < y) \\ &= \begin{cases} P\left(X < \frac{y-b}{a}\right), & a > 0 \text{ のとき} \\ P\left(X > \frac{y-b}{a}\right), & a < 0 \text{ のとき} \end{cases} \\ &= \begin{cases} P\left(X < \frac{y-b}{a}\right), & a > 0 \text{ のとき} \\ 1 - P\left(X < \frac{y-b}{a}\right), & a < 0 \text{ のとき} \end{cases} \\ &= \begin{cases} F\left(\frac{y-b}{a}\right), & a > 0 \text{ のとき} \\ 1 - F\left(\frac{y-b}{a}\right), & a < 0 \text{ のとき} \end{cases} \end{aligned}$$

分布関数と密度関数との関係は、

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

であるので、 Y の密度関数は、

$$\begin{aligned} g(y) &= \frac{dG(y)}{dy} \\ &= \begin{cases} \frac{1}{a} f\left(\frac{y-b}{a}\right), & a > 0 \text{ のとき} \\ -\frac{1}{a} f\left(\frac{y-b}{a}\right), & a < 0 \text{ のとき} \end{cases} \\ &= \left|\frac{1}{a}\right| f\left(\frac{y-b}{a}\right) \end{aligned}$$

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、

$$g(y) = |h'(y)| f(h(y))$$

となる。

12.2.2 回帰分析への応用

回帰モデル

$$Y_i = \alpha + \beta X_i + u_i \quad i = 1, 2, \dots, n$$

$u_1, u_2, \dots, u_n$ は互いに独立で、すべての i について $u_i \sim N(0, \sigma^2)$ を仮定する。

u_i の密度関数は、

$$f(u_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} u_i^2\right)$$

となる。

Y_i の密度関数 $g(Y_i)$ は、

$$g(Y_i) = |h'(Y_i)| f(h(Y_i))$$

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、

$$\begin{aligned} g(Y_i) &= |h'(Y_i)| f(h(Y_i)) \\ &= f(h(Y_i)) \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} (Y_i - \alpha - \beta X_i)^2\right) \end{aligned}$$

となる。

$u_1, u_2, \dots, u_n$ は互いに独立であれば, $Y_1, Y_2, \dots, Y_n$ も互いに独立になるので, $Y_1, Y_2, \dots, Y_n$ の結合密度関数は,

$$\begin{aligned} g(Y_1, Y_2, \dots, Y_n) &= \prod_{i=1}^n g(Y_i) \\ &= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (Y_i - \alpha - \beta X_i)^2\right) \end{aligned}$$

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,

$$l(\alpha, \beta, \sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (Y_i - \alpha - \beta X_i)^2\right)$$

となる。

対数尤度関数は,

$$\begin{aligned} \log l(\alpha, \beta, \sigma^2) &= -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log(\sigma^2) \\ &\quad - \frac{1}{2\sigma^2} \sum_{i=1}^n (Y_i - \alpha - \beta X_i)^2 \end{aligned}$$

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,

$$\hat{\beta} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta} \bar{X}$$

σ^2 の最尤推定量は,

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

となり, s^2 とは異なる。

$\hat{\theta} = (\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2)'$, $\theta = (\alpha, \beta, \sigma^2)'$ とする。 n が大きいとき,

$$\hat{\theta} \sim N(\theta, \Sigma_\theta)$$

ただし,

$$\begin{aligned} \Sigma_\theta &= \left(\sum_{i=1}^n E \left[\left(\frac{\partial \log g(Y_i; \theta)}{\partial \theta} \right) \left(\frac{\partial \log g(Y_i; \theta)}{\partial \theta} \right)' \right] \right)^{-1} \\ &= - \left(\sum_{i=1}^n E \left[\frac{\partial^2 \log g(Y_i; \theta)}{\partial \theta \partial \theta'} \right] \right)^{-1} \end{aligned}$$

Y_i の密度関数 $g(Y_i; \theta)$ の対数は,

$$\begin{aligned} \log g(Y_i; \theta) &= -\frac{1}{2} \log(2\pi) - \frac{1}{2} \log(\sigma^2) \\ &\quad - \frac{1}{2\sigma^2} (Y_i - \alpha - \beta X_i)^2 \end{aligned}$$

となる。

$$\begin{aligned} \frac{\partial \log g(Y_i; \theta)}{\partial \theta} &= \begin{pmatrix} \frac{\partial \log g(Y_i; \theta)}{\partial \alpha} \\ \frac{\partial \log g(Y_i; \theta)}{\partial \beta} \\ \frac{\partial \log g(Y_i; \theta)}{\partial \sigma^2} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{\sigma^2} (Y_i - \alpha - \beta X_i) \\ \frac{1}{\sigma^2} X_i (Y_i - \alpha - \beta X_i) \\ -\frac{1}{2\sigma^2} + \frac{1}{2\sigma^4} (Y_i - \alpha - \beta X_i)^2 \end{pmatrix} \end{aligned}$$

$$\frac{\partial^2 \log g(Y_i; \theta)}{\partial \theta \partial \theta'}$$

$$\begin{aligned} &= \begin{pmatrix} \frac{\partial^2 \log g(Y_i; \theta)}{\partial \alpha^2} & \frac{\partial^2 \log g(Y_i; \theta)}{\partial \alpha \partial \beta} & \frac{\partial^2 \log g(Y_i; \theta)}{\partial \alpha \partial \sigma^2} \\ \frac{\partial^2 \log g(Y_i; \theta)}{\partial \beta \partial \alpha} & \frac{\partial^2 \log g(Y_i; \theta)}{\partial \beta^2} & \frac{\partial^2 \log g(Y_i; \theta)}{\partial \beta \partial \sigma^2} \\ \frac{\partial^2 \log g(Y_i; \theta)}{\partial \sigma^2 \partial \alpha} & \frac{\partial^2 \log g(Y_i; \theta)}{\partial \sigma^2 \partial \beta} & \frac{\partial^2 \log g(Y_i; \theta)}{\partial (\sigma^2)^2} \end{pmatrix} \\ &= \begin{pmatrix} -\frac{1}{\sigma^2} & -\frac{X_i}{\sigma^2} & -\frac{u_i}{\sigma^4} \\ \frac{X_i}{\sigma^2} & \frac{X_i^2}{\sigma^2} & \frac{X_i u_i}{\sigma^4} \\ -\frac{u_i}{\sigma^4} & -\frac{X_i u_i}{\sigma^4} & \frac{1}{2\sigma^4} - \frac{u_i^2}{\sigma^6} \end{pmatrix} \end{aligned}$$

ただし, $u_i = Y_i - \alpha - \beta X_i$

期待値をとると,

$$\begin{aligned} E\left(\frac{\partial^2 \log g(Y_i; \theta)}{\partial \theta \partial \theta'}\right) &= E\begin{pmatrix} -\frac{1}{\sigma^2} & -\frac{X_i}{\sigma^2} & -\frac{u_i}{\sigma^4} \\ -\frac{X_i}{\sigma^2} & -\frac{X_i^2}{\sigma^2} & -\frac{X_i u_i}{\sigma^4} \\ -\frac{u_i}{\sigma^4} & -\frac{X_i u_i}{\sigma^4} & \frac{1}{2\sigma^4} - \frac{u_i^2}{\sigma^6} \end{pmatrix} \\ &= \begin{pmatrix} -\frac{1}{\sigma^2} & -\frac{X_i}{\sigma^2} & 0 \\ -\frac{X_i}{\sigma^2} & -\frac{X_i^2}{\sigma^2} & 0 \\ 0 & 0 & -\frac{1}{2\sigma^4} \end{pmatrix} \end{aligned}$$

となる。

$$\begin{aligned} \Sigma_\theta &= -\left(\sum_{i=1}^n E\left[\frac{\partial^2 \log g(Y_i; \theta)}{\partial \theta \partial \theta'}\right]\right)^{-1} \\ &= \begin{pmatrix} \frac{n}{\sigma^2} & \frac{\sum_{i=1}^n X_i}{\sigma^2} & 0 \\ \frac{\sum_{i=1}^n X_i}{\sigma^2} & \frac{\sum_{i=1}^n X_i^2}{\sigma^2} & 0 \\ 0 & 0 & \frac{n}{2\sigma^4} \end{pmatrix}^{-1} \\ &= \begin{pmatrix} \sigma^2 \left(\frac{n}{\sum_{i=1}^n X_i} \quad \frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n X_i^2}\right)^{-1} & 0 \\ 0 & \frac{2\sigma^4}{n} \end{pmatrix} \end{aligned}$$

したがって,

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} \sim N\left(\begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \sigma^2 \begin{pmatrix} n & \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i^2 \end{pmatrix}^{-1}\right)$$

となる。

→ 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

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

$$u_i = \rho u_{i-1} + \epsilon_i \quad i = 2, 3, \dots, n$$

$\epsilon_2, \epsilon_3, \dots, \epsilon_n$ は互いに独立で, すべての i について $\epsilon_i \sim N(0, \sigma^2)$ 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。

$$\begin{aligned} \log f(Y_i; \theta) &= -\frac{1}{2} \log(2\pi\sigma^2) \\ &\quad - \frac{1}{2\sigma^2} \left((Y_i - \rho Y_{i-1}) - \alpha(1 - \rho) - \beta(X_i - \rho X_{i-1}) \right)^2 \end{aligned}$$

尤度関数は,

$$\begin{aligned} \log l(\theta) &= \sum_{i=2}^n \log f(Y_i; \theta) = -\frac{n-1}{2} \log(2\pi) - \frac{n-1}{2} \log(\sigma^2) \\ &\quad - \frac{1}{2\sigma^2} \sum_{i=2}^n \left((Y_i - \rho Y_{i-1}) - \alpha(1 - \rho) - \beta(X_i - \rho X_{i-1}) \right)^2 \end{aligned}$$

となる。

尤度関数をそれぞれ $\alpha, \beta, \sigma^2, \rho$ について微分し, ゼロとおく。

$$\begin{aligned} \frac{\partial \log l(\theta)}{\partial \alpha} &= \frac{1-\rho}{\sigma^2} \sum_{i=2}^n \left((Y_i - \rho Y_{i-1}) \right. \\ &\quad \left. - \alpha(1 - \rho) - \beta(X_i - \rho X_{i-1}) \right) = 0 \end{aligned}$$

$$\begin{aligned} \frac{\partial \log l(\theta)}{\partial \beta} &= \frac{1}{\sigma^2} \sum_{i=2}^n (X_i - \rho X_{i-1}) \left((Y_i - \rho Y_{i-1}) \right. \\ &\quad \left. - \alpha(1 - \rho) - \beta(X_i - \rho X_{i-1}) \right) = 0 \end{aligned}$$

$$\begin{aligned} \frac{\partial \log l(\theta)}{\partial \sigma^2} &= -\frac{n-1}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=2}^n \left((Y_i - \rho Y_{i-1}) \right. \\ &\quad \left. - \alpha(1 - \rho) - \beta(X_i - \rho X_{i-1}) \right)^2 = 0 \end{aligned}$$

$$\begin{aligned} \frac{\partial \log l(\theta)}{\partial \rho} &= \frac{1}{\sigma^2} \sum_{i=2}^n (Y_{i-1} - \alpha - \beta X_{i-1}) \left((Y_i - \alpha - \beta X_i) \right. \\ &\quad \left. - \rho(Y_{i-1} - \alpha - \beta X_{i-1}) \right) = 0 \end{aligned}$$

$(Y_i - \alpha - \beta X_i) - \rho(Y_{i-1} - \alpha - \beta X_{i-1})$ は $(Y_i - \rho Y_{i-1}) - \alpha(1 - \rho) - \beta(X_i - \rho X_{i-1})$ を書き直したものの。

4 つの連立方程式を解いて, 最尤推定量 $\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2, \hat{\rho}$ が得られる。

→ 下記のように収束計算によって求める。

(i) 初期段階では、 $\hat{\rho} = 0$ とする。

$$(ii) X_i^* = X_i - \hat{\rho}X_{i-1}$$

$$Y_i^* = Y_i - \hat{\rho}Y_{i-1}$$

$$(iii) \begin{pmatrix} \tilde{\alpha} \\ \hat{\beta} \end{pmatrix} = \begin{pmatrix} n-1 & \sum_{i=2}^n X_i^* \\ \sum_{i=2}^n X_i^* & \sum_{i=2}^n X_i^{*2} \end{pmatrix}^{-1} \begin{pmatrix} \sum_{i=2}^n Y_i^* \\ \sum_{i=2}^n X_i^* Y_i^* \end{pmatrix}$$

$$(iv) \hat{\alpha} = \frac{\tilde{\alpha}}{1 - \hat{\rho}}$$

$$(v) \hat{u}_i = Y_i - \hat{\alpha} - \hat{\beta}X_i$$

$$(vi) \hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=2}^n (\hat{u}_i - \hat{\rho}\hat{u}_{i-1})^2$$

$$(vii) \hat{\rho} = \frac{\sum_{i=2}^n \hat{u}_i \hat{u}_{i-1}}{\sum_{i=2}^n \hat{u}_{i-1}^2}$$

(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 $X_1, X_2, \dots, X_n$ 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 $\tilde{\theta}$ 、制約無し最尤推定量を $\hat{\theta}$ とする。

制約の数を G 個とする。

$\frac{l(\tilde{\theta})}{l(\hat{\theta})}$ 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、

$$\frac{l(\tilde{\theta})}{l(\hat{\theta})} < c$$

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。

$$\int \cdots \int \prod_{i=1}^n f(x_i; \tilde{\theta}) dx_1 \cdots dx_n = \alpha$$

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、 $n \rightarrow \infty$ のとき、

$$-2 \log \frac{l(\tilde{\theta})}{l(\hat{\theta})} \rightarrow \chi^2(G)$$

となる。

この検定を尤度比検定と呼ぶ。

例 1： 正規母集団 $N(\mu, \sigma^2)$ からの標本値 $x_1, x_2, \dots, x_n$ 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、

$$l(\mu) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\right)$$

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、

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n X_i \equiv \bar{X}$$

となる。

尤度比検定統計量は、

$$\begin{aligned} \frac{l(\mu_0)}{l(\bar{X})} &= \frac{\exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (X_i - \mu_0)^2\right)}{\exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2\right)} \\ &= \exp\left(-\frac{1}{2\sigma^2/n} (\bar{X} - \mu_0)^2\right) < c \end{aligned}$$

となる c を求める。

H_0 が正しいときに、 $\sqrt{n}(\bar{X} - \mu_0)/\sigma \sim N(0, 1)$ となるので、

$$P\left(\left|\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}\right| > z_{\alpha/2}\right) = \alpha$$

すなわち、

$$P\left(\exp\left(-\frac{1}{2\sigma^2/n} (\bar{X} - \mu_0)^2\right) < \exp\left(-\frac{1}{2} z_{\alpha/2}^2\right)\right) = \alpha$$

と変形できる。したがって、

$$c = \exp\left(-\frac{1}{2} z_{\alpha/2}^2\right)$$

とすればよい。

例 2: $X_1, X_2, \dots, X_n$ 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、

$$f(x; p) = p^x (1-p)^{1-x} \quad x = 0, 1$$

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 $\hat{p}$ は、

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

である。

次の仮説検定を考える。

$$H_0: p = p_0 \quad H_1: p \neq p_0$$

→ 制約数は 1 つ。 ($G = 1$)

尤度比は、

$$\frac{l(p_0)}{l(\hat{p})} = \frac{\prod_{i=1}^n p_0^{X_i} (1-p_0)^{1-X_i}}{\prod_{i=1}^n \hat{p}^{X_i} (1-\hat{p})^{1-X_i}}$$

したがって、 $n \rightarrow \infty$ のとき、

$$-2 \log \frac{l(p_0)}{l(\hat{p})} = -2 \log \frac{p_0}{\hat{p}} \sum_{i=1}^n X_i - 2 \log \frac{1-p_0}{1-\hat{p}} \sum_{i=1}^n (1-X_i)$$

→ $\chi^2(1)$

$\chi^2(1)$ 分布の上側 100 $\alpha\%$ 点を $\chi_\alpha^2(1)$ とするとき、

$$-2 \log \frac{l(p_0)}{l(\hat{p})} > \chi_\alpha^2(1)$$

のとき、帰無仮説 $H_0: p = p_0$ を棄却する。

例 3: 回帰モデル

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

$$u_i \sim N(0, \sigma^2) \quad i = 1, 2, \dots, n$$

について、 $\beta_1, \dots, \beta_k$ 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

$\theta = (\beta_1, \dots, \beta_k, \sigma^2)$ とする。

尤度関数は、

$$l(\theta) = \prod_{i=1}^n f(Y_i; \theta) \\ = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (Y_i - \beta_1 X_{1i} - \dots - \beta_k X_{ki})^2\right)$$

となる。

H_0 の制約つき最尤推定量を $\tilde{\theta} = (\tilde{\beta}_1, \dots, \tilde{\beta}_k, \tilde{\sigma}^2)$ とする。

この仮説に含まれる制約数を G とする。

制約なし最尤推定量を $\hat{\theta} = (\hat{\beta}_1, \dots, \hat{\beta}_k, \hat{\sigma}^2)$ とする。

尤度比

$$\frac{l(\tilde{\theta})}{l(\hat{\theta})} = \frac{(2\pi\tilde{\sigma}^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\tilde{\sigma}^2} \sum_{i=1}^n (Y_i - \tilde{\beta}_1 X_{1i} - \dots - \tilde{\beta}_k X_{ki})^2\right)}{(2\pi\hat{\sigma}^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\hat{\sigma}^2} \sum_{i=1}^n (Y_i - \hat{\beta}_1 X_{1i} - \dots - \hat{\beta}_k X_{ki})^2\right)} \\ = \frac{(\tilde{\sigma}^2)^{-\frac{n}{2}} \exp\left(-\frac{n-G}{2}\right)}{(\hat{\sigma}^2)^{-\frac{n}{2}} \exp\left(-\frac{n-k}{2}\right)} \\ = \left(\frac{\frac{1}{n-G} \sum_{i=1}^n \tilde{u}_i^2}{\frac{1}{n-k} \sum_{i=1}^n \hat{u}_i^2}\right)^{-n/2} \exp\left(-\frac{k-G}{2}\right) \\ = \exp\left(-\frac{k-G}{2}\right) \left(\frac{n-k}{n-G}\right)^{-n/2} \left(\frac{\sum_{i=1}^n \tilde{u}_i^2}{\sum_{i=1}^n \hat{u}_i^2}\right)^{-n/2} \\ = \exp\left(-\frac{k-G}{2}\right) \left(\frac{n-k}{n-G}\right)^{-n/2} \\ \times \left(1 + \frac{\sum_{i=1}^n \tilde{u}_i^2 - \sum_{i=1}^n \hat{u}_i^2}{\sum_{i=1}^n \hat{u}_i^2}\right)^{-n/2} \\ = \exp\left(-\frac{k-G}{2}\right) \left(\frac{n-k}{n-G}\right)^{-n/2} \\ \times \left(1 + \frac{G}{n-k} \frac{(\sum_{i=1}^n \tilde{u}_i^2 - \sum_{i=1}^n \hat{u}_i^2)/G}{\sum_{i=1}^n \hat{u}_i^2/(n-k)}\right)^{-n/2} \\ < c$$

のとき仮説を棄却する。

$$\frac{(\sum_{i=1}^n \tilde{u}_i^2 - \sum_{i=1}^n \hat{u}_i^2)/G}{\sum_{i=1}^n \tilde{u}_i^2/(n-k)} \sim F(G, n-k)$$

を利用すると c が求まる。

ただし、途中で以下を利用

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

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

近似的には、

$$\begin{aligned} -2 \log \frac{l(\tilde{\theta})}{l(\hat{\theta})} &= -2 \log \frac{(\tilde{\sigma}^2)^{-\frac{n}{2}} \exp\left(-\frac{n-G}{2}\right)}{(\hat{\sigma}^2)^{-\frac{n}{2}} \exp\left(-\frac{n-k}{2}\right)} \\ &= n \log\left(\frac{\tilde{\sigma}^2}{\hat{\sigma}^2}\right) + (k-G) \\ &\rightarrow \chi^2(G) \end{aligned}$$

例 4： 回帰モデル

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

$$u_i = \rho u_{i-1} + \epsilon_i$$

$$\epsilon_i \sim N(0, \sigma^2) \quad i = 2, 3, \dots, n$$

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、

$$\begin{aligned} \log l(\theta) &= \sum_{i=2}^n \log f(Y_i; \theta) = -\frac{n-1}{2} \log(2\pi) - \frac{n-1}{2} \log(\sigma^2) \\ &\quad - \frac{1}{2\sigma^2} \sum_{i=2}^n \left((Y_i - \rho Y_{i-1}) - \alpha(1-\rho) - \beta(X_i - \rho X_{i-1}) \right)^2 \end{aligned}$$

となる。

対数尤度関数をそれぞれ $\alpha, \beta, \sigma^2, \rho$ 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 $\hat{\theta} = (\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2, \hat{\rho})$ 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 $\tilde{\theta} = (\tilde{\alpha}, \tilde{\beta}, \tilde{\sigma}^2, 0)$ が得られる。

すなわち、

$$\frac{\max_{\alpha, \beta, \sigma^2} l(\alpha, \beta, \sigma^2, 0)}{\max_{\alpha, \beta, \sigma^2, \rho} l(\alpha, \beta, \sigma^2, \rho)} = \frac{l(\tilde{\alpha}, \tilde{\beta}, \tilde{\sigma}^2, 0)}{l(\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2, \hat{\rho})} = \frac{l(\tilde{\theta})}{l(\hat{\theta})}$$

$\log l(\hat{\theta})$ は、 $\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=2}^n \left((Y_i - \hat{\rho} Y_{i-1}) - \hat{\alpha}(1-\hat{\rho}) - \hat{\beta}(X_i - \hat{\rho} X_{i-1}) \right)^2$ に注意して、

$$\begin{aligned} \log l(\hat{\theta}) &= -\frac{n-1}{2} \log(2\pi) - \frac{n-1}{2} \log(\hat{\sigma}^2) \\ &\quad - \frac{1}{2\hat{\sigma}^2} \sum_{i=2}^n \left((Y_i - \hat{\rho} Y_{i-1}) - \hat{\alpha}(1-\hat{\rho}) - \hat{\beta}(X_i - \hat{\rho} X_{i-1}) \right)^2 \\ &= -\frac{n-1}{2} \log(2\pi) - \frac{n-1}{2} \log(\hat{\sigma}^2) - \frac{n-1}{2} \end{aligned}$$

となる。

同様に、 $\log l(\tilde{\theta})$ は、 $\tilde{\sigma}^2 = \frac{1}{n-1} \sum_{i=2}^n (Y_i - \tilde{\alpha} - \tilde{\beta} X_i)^2$ に注意して、

$$\begin{aligned} \log l(\tilde{\theta}) &= -\frac{n-1}{2} \log(2\pi) - \frac{n-1}{2} \log(\tilde{\sigma}^2) \\ &\quad - \frac{1}{2\tilde{\sigma}^2} \sum_{i=2}^n (Y_i - \tilde{\alpha} - \tilde{\beta} X_i)^2 \\ &= -\frac{n-1}{2} \log(2\pi) - \frac{n-1}{2} \log(\tilde{\sigma}^2) - \frac{n-1}{2} \end{aligned}$$

となる。

したがって、尤度比検定統計量

$$-2 \log \frac{l(\tilde{\theta})}{l(\hat{\theta})} = (n-1) \log \frac{\tilde{\sigma}^2}{\hat{\sigma}^2}$$

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、

$$\text{加法型モデル: } X_t = T_t + C_t + S_t + I_t$$

または、

$$\text{乗法型モデル: } X_t = T_t \times C_t \times S_t \times I_t$$

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) から季節変動を取り除いたデータ

季節調整済みデータを作成する簡単な方法 → 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:

$$\begin{aligned} Y_t &= \frac{1}{2} \left(\underbrace{\frac{1}{12}(X_{t-6} + \cdots + X_{t+5})}_{(t-0.5) \text{ 期の季節調整値}} + \underbrace{\frac{1}{12}(X_{t-5} + \cdots + X_{t+6})}_{(t+0.5) \text{ 期の季節調整値}} \right) \\ &\quad \underbrace{\hspace{10em}}_{t \text{ 期の季節調整値}} \\ &= \frac{1}{24} (X_{t-6} + 2(X_{t-5} + \cdots + X_{t+5}) + X_{t+6}) \end{aligned}$$

四半期データの季節調整:

$$\begin{aligned} Y_t &= \frac{1}{2} \left(\underbrace{\frac{1}{4}(X_{t-2} + \cdots + X_{t+1})}_{(t-0.5) \text{ 期の季節調整値}} + \underbrace{\frac{1}{4}(X_{t-1} + \cdots + X_{t+2})}_{(t+0.5) \text{ 期の季節調整値}} \right) \\ &\quad \underbrace{\hspace{10em}}_{t \text{ 期の季節調整値}} \\ &= \frac{1}{8} (X_{t-2} + 2(X_{t-1} + X_t + X_{t+1}) + X_{t+2}) \end{aligned}$$

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、

$$\log T_t = A' + bt$$

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。

$$C_t = T_t - \hat{\alpha} - \hat{\beta}t$$

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 時系列分析と定常性

時系列データが生成された確率的構造を把握するために、その生成過程を確率的に表す。

観察された時系列データの集合を

$$\{x_t\} = \{x_1, x_2, \dots, x_n\}$$

によって表す。

各期の観察値を確率変数の実現値として捉える。次のような確率変数の集合（確率過程と呼ぶ）を考える。

$$\{X_t\} = \{\dots, X_1, X_2, \dots, X_n, \dots\}$$

第 t 期の観察値 x_t は、第 1 期の確率変数 X_t の実現値と考える。

各期の確率変数はそれぞれ分布関数によって特徴づけられており、平均、分散をもつ。

異時点間の確率変数について、相互に依存し、共分散を定義する。

時系列データが生成された確率的構造を明らかにすることは、これらの確率変数の特性値を推定すること。

一般に各期の平均、分散、さらに異時点間の共分散は異なった値をとりうるであろう。

しかし、各時点の確率変数の平均、分散を推計するためのデータはその期の実現値 1 つしか利用できない。

そこで推定のためには、確率的構造に制約を課す必要が出てくる。

その制約として重要なのが弱定常性（weak stationary）という性質である。

弱定常性は次の 3 つの式によって表される。

$$E[X_t] = \mu$$

$$V(X_t) = E[(X_t - \mu)^2] = \phi(0)$$

$$\text{Cov}(X_t, X_{t-s}) = E[(X_t - \mu)(X_{t-s} - \mu)] = \phi(s)$$

各期の平均、分散が時間に依存せずに一定であることを示している。

X_t と X_{t-s} の間の共分散（自己共分散、auto-covariance）が 2 時点間の時間的距離 s にのみに依存している。分散（自己分散）は $s = 0$ の場合であることに注意。

定常性の制約を課することにより、推定すべきパラメータの数を大幅に減らすことができる。

平均、分散、共分散の推定には、以下の標本平均、標本分散、標本共分散が用いられる。

$$\bar{x} = \frac{1}{n} \sum_{t=1}^n x_t$$

$$\hat{\phi}(0) = \frac{1}{n} \sum_{t=1}^n (x_t - \bar{x})^2$$

$$\hat{\phi}(s) = \frac{1}{n-s} \sum_{t=s+1}^n (x_t - \bar{x})(x_{t-s} - \bar{x})$$

経済データには定常性を満たさないものが多い。

この場合は、階差をとる（2 回階差をとる場合もある）

時系列データの背後の確率過程を特定化する際に役立つのが自己相関（係数）（auto-correlation）、あるいは、コレログラム（correlogram）

自己相関は自己共分散 $\phi(s)$ を自己分散 $\phi(0)$ で割った値である。

$$\rho(s) = \frac{\phi(s)}{\phi(0)}$$

自己相関は異時点間の確率変数の相関関係を表しており、その値は -1 と 1 の間にある。

自己相関係数の全体をコレログラムという。

各期の確率変数が独立に分布している場合には、異時点間の自己相関はすべてゼロとなる。

平均がゼロで分散が一定の定常性をもつ確率過程で自己相関がゼロのものを、ホワイト・ノイズ (white noise)、または、白色雑音と呼ぶ。

標本について見れば、標本自己相関は次式で計算される。

$$\hat{\rho}(s) = \frac{\hat{\phi}(s)}{\hat{\phi}(0)}$$

14.1 時系列モデルの特定化

14.1.1 自己回帰 (AR) モデル

自己回帰モデルの基本型は AR(1) といわれる次のようなモデルである。

$$X_t = \mu + \theta X_{t-1} + u_t$$

ただし、 μ , θ は定数、 u_t は平均 0、分散 σ^2 をもつホワイト・ノイズである。

また AR(1) の () 内の 1 はモデルの次数と呼ばれる。

上記のモデルは 1 次の自己回帰モデルである。

確率過程 $\{X_t\}$ が定常性を満たすには $|\theta| < 1$ でなければならない (理由は等比数列の収束の条件によっている)。

X_t の平均は $\mu/(1 - \theta)$,

自己分散 $\phi(0)$ は $\sigma^2/(1 - \theta^2)$,

自己共分散 $\phi(s)$ は $\theta^s \phi(0)$,

自己相関は $\rho(s) = \theta^s$ となる。

$|\theta| < 1$ であるから、自己相関は確率変数間の時間的距離 (すなわち、 s) が長ければ長いほど小さくなっていく。

14.1.2 移動平均 (MA) モデル

移動平均モデルの基本型は、MA(1) といわれる次のようなモデルである。

$$X_t = \mu + u_t + \psi u_{t-1}$$

ホワイト・ノイズの移動平均によって表現されている。

X_t の平均は μ で、

分散 $\phi(0)$ は $\sigma^2(1 + \psi^2)$,

自己共分散 $\phi(s)$ は、 s が 1 のときには $\psi\sigma^2$ 、それ以外のときには 0 となる。

自己相関も s が 1 のときには $\psi/(1 + \psi^2)$ 、それ以外のときには 0 となる。

ある時点から自己相関が 0 になるのが移動平均モデルの特徴である。

14.1.3 より複雑なモデル

より複雑な確率過程を描写するには、自己回帰 (AR) モデルの次数をさらに高くする。

移動平均 (MA) モデルではホワイト・ノイズの移動平均をより過去のホワイト・ノイズを含める。

さらに、自己回帰モデルと移動平均モデルを結合した自己回帰移動平均 (ARMA) モデルを構築することもできる。

$$X_t = \mu + \theta_1 X_{t-1} + \theta_2 X_{t-2} + \cdots + \theta_p X_{t-p} + u_t + \psi_1 u_{t-1} + \cdots + \psi_q u_{t-q}$$

このモデルを ARMA(p, q) モデルと呼ぶ。

X_t を

$$1 \text{ 階階差 } \Delta X_t = X_t - X_{t-1}$$

$$2 \text{ 階階差 } \Delta^2 X_t = \Delta X_t - \Delta X_{t-1}$$

$\vdots$

$$d \text{ 階階差 } \Delta^d X_t = \Delta^{d-1} X_t - \Delta^{d-1} X_{t-1}$$

で置き換えて、

$$\Delta^d X_t = \mu + \theta_1 \Delta^d X_{t-1} + \theta_2 \Delta^d X_{t-2} + \cdots + \theta_p \Delta^d X_{t-p} + u_t + \psi_1 u_{t-1} + \cdots + \psi_q u_{t-q}$$

を ARIMA(p, d, q) モデルと呼ぶ。

14.2 時系列モデルの作成手順と予測

残差 $\hat{u}_t$ から判断して、 u_t がホワイト・ノイズになるように、 p, q を選んで、何度も推定しなおす。

標本の自己相関パターンから判断して適当と考えられるモデル（すなわち、 p, d, q ）を選択する。その場合にはできるだけ少ないパラメータ数（ $= p + q$ ）となるモデルを選択する。

14.3 非定常時系列

14.3.1 単位根

経済変数で、特に重要な非定常時系列モデルは、ランダム・ウォーク過程（random walk process）である。

国内総生産（GDP）、消費、投資、マネーサプライなどほとんどの経済変数は、ランダム・ウォーク過程に従う非定常時系列であるといわれている。

ある時系列がランダム・ウォーク過程に従っているかどうかの検定を単位根検定（unit root test）と呼ぶ。

この単位根検定における検定統計量は、 t 分布や正規分布には従わないことがわかっている。

ランダム・ウォーク過程とは、 $X_1, X_2, \dots, X_n$ の系列が、

$$X_t = X_{t-1} + u_t$$

と表現される。ただし、 u_t は、平均ゼロ、分散 σ^2 の分布に従うものとする。

1 階の階差をとることによって定常過程となるとき、その系列は 1 次の和分過程（integrated process of order one）に従っていると呼ばれ、 $I(1)$ 過程と表現される（定常過程は $I(0)$ と表される）。

X_t が $I(d)$ 過程に従っているとき、 $X_t \sim I(d)$ と表記することがある。

この記号を使うと、

$$X_t \sim I(1) \iff \Delta X_t \sim I(0)$$

という関係がある。

上記の $I(1)$ 過程 X_t が非定常であることを示す。

定常性（または、弱定常性）の条件は、

- (i) X_t の平均は一定、
- (ii) X_t の分散も一定、
- (iii) X_t と X_{t-s} の共分散（すなわち、自己共分散）は時

点の差のみに依存する、
ということである。

初期値 X_0 を定数（非確率）と仮定し、繰り返し代入することによって、

$$X_t = X_0 + u_1 + u_2 + \dots + u_t$$

という式が得られる。

これを用いると、 X_t の平均と分散は、

$$E[X_t] = X_0 \quad V(X_t) = t\sigma^2$$

となる。また、

$$X_t = X_{t-s} + u_{t-s+1} + u_{t-s+2} + \dots + u_t$$

を用いると、 X_t と X_{t-s} の共分散は、

$$\begin{aligned} \text{Cov}(X_t, X_{t-s}) &= E[(X_t - X_0)(X_{t-s} - X_0)] \\ &= E[(X_{t-s} + u_{t-s+1} + u_{t-s+2} + \dots + u_t - X_0)(X_{t-s} - X_0)] \\ &= E[(X_{t-s} - X_0)^2] \\ &\quad + E[(u_{t-s+1} + u_{t-s+2} + \dots + u_t)(X_{t-s} - X_0)] \\ &= V(X_{t-s}) = (t-s)\sigma^2 \end{aligned}$$

となる。

X_{t-s} は、過去の誤差項（ $u_1, u_2, \dots, u_{t-s}$ ）とは相関をもつが、将来の誤差項（ $u_{t-s+1}, u_{t-s+2}, \dots, u_t$ ）とは無相関である。

以上から、

$V(X_t)$ は定数でなく時点 t に依存し、

$\text{Cov}(X_t, X_{t-s})$ も時点の差 s だけでなく時点 t にも依存することがわかる。

t が大きくなると、 $V(X_t)$ 、 $\text{Cov}(X_t, X_{t-s})$ とともに無限大になる。

したがって、ランダム・ウォーク過程 X_t は非定常であるといえる。

さらに、 X_t と X_{t-s} の相関係数（すなわち、自己相関係数）は、

$$\begin{aligned} \rho(s) = \rho(X_t, X_{t-s}) &= \frac{\text{Cov}(X_t, X_{t-s})}{\sqrt{V(X_t)}\sqrt{V(X_{t-s})}} \\ &= \frac{(t-s)\sigma^2}{\sqrt{t\sigma^2}\sqrt{(t-s)\sigma^2}} = \sqrt{\frac{t-s}{t}} \end{aligned}$$

と計算される。

時系列に単位根が含まれている場合、 t を一定としたとき、 s が大きくなるにつれて、 $\rho(s)$ はゆっくりと減衰していくという性質がある。

この性質を利用して、従来はグラフを描いて視覚的に単位根があるかどうかを判断していた。

しかし、 ρ の値が 1 に近い定常過程の場合も、 $\rho = 1$ の非定常過程の場合と同様に、自己相関関数が減衰的なものとなることが知られている。

したがって、自己相関関数の形状からだけでは定常過程と非定常過程との区別をすることが実際には困難である。

それを避けるために統計的な検定を行うことが必要とされる。

1970 年代後半に単位根の検定が開発され、近年では統計的に定常か非定常かの検定が行われるようになった。

以下に、単位根検定の方法を簡単に説明する。

X_t が AR(1) 過程

$$X_t = \theta X_{t-1} + u_t$$

に従っているとすると。

ただし、 u_t は、 $t = 1, 2, \dots, n$ についてそれぞれ独立に、平均ゼロ、分散 σ^2 の分布に従うものとする。

もし $|\theta| < 1$ であれば $X_t \sim I(0)$ であるが、

もし $\theta = 1$ であれば X_t は単位根をもち、 $X_t \sim I(1)$ となる。

特に、経済時系列では、単位根があれば $\theta = 1$ となり、そうでなければ $\theta < 1$ となる。

そのため、単位根検定では左片側検定が行われる。

実際に、検定するためには、

$$\Delta X_t = \beta X_{t-1} + u_t$$

と変形して、帰無仮説 $H_0 : \beta = 0$ 、対立仮説 $H_1 : \beta < 0$ の左片側検定を行うことになる ($\beta = \theta - 1$ という関係がある)。

この場合、 β の最小 2 乗推定量

$$\hat{\beta} = \frac{\sum_{t=1}^n t = 1nX_{t-1}\Delta X_t}{\sum_{t=1}^n t = 1nX_{t-1}^2}$$

は、正規分布には従わないことが知られている。

したがって、帰無仮説 $H_0 : \beta = 0$ が正しいとき、検定統計量

$$t_{\hat{\beta}} = \frac{\hat{\beta}}{Se(\hat{\beta})}$$

も t 分布には従わず、通常の t 検定を適用することはできない。

検定統計量 $t_{\hat{\beta}}$ の分布は、明示的に求めることができず、特殊な関数形をしているが、 t 分布よりも左側の裾野が広い分布となっていることが知られている。

したがって、検定統計量 $t_{\hat{\beta}}$ を誤って通常の t 分布表 (付表 3) を用いて検定すると、対立仮説 $H_1 : \beta < 0$ を過剰に採択してしまうことになる。

言いかえると、実際は $I(1)$ 変数であるにもかかわらず、誤って、 $I(0)$ 変数と判定されてしまう可能性が高くなるのである。

検定統計量 $t_{\hat{\beta}}$ の分布の下側 1 パーセント点と 5 パーセント点の値は、下の表に示されている (検定統計量 $t_{\hat{\beta}}$ の分布は標本数 n に依存することに注意せよ)。

単位根検定

α	n	25	50	100	250	500	∞
0.01		-2.66	-2.62	-2.60	-2.58	-2.58	-2.58
0.05		-1.95	-1.95	-1.95	-1.95	-1.95	-1.95

具体的には、 $t_{\hat{\beta}}$ の統計値と表中の数値とを比較する。

もし $t_{\hat{\beta}}$ の統計値が、表中の対応する数値より大きければ、帰無仮説 $H_0 : \beta = 0$ を採択することになる。

この場合、 X_t は単位根をもつ非定常時系列であると判断する。

14.3.2 見せかけ回帰

経済時系列の場合、全く関係のない変数間で回帰分析を行った結果、決定係数が高く t 値も高いという結果を得る傾向がある。この回帰は見せかけ回帰 (spurious regression) と呼ばれ、これは経済時系列変数で回帰分析を行う場合に非常に重要な問題となる。

次の単回帰モデルを考える。

$$Y_t = \alpha + \beta X_t + u_t$$

Y_t と X_t が次のように単位根をもつ非定常時系列であるとする。

$$Y_t = Y_{t-1} + \epsilon_t$$

$$X_t = X_{t-1} + \eta_t$$

ただし、 ϵ_t , η_t は、すべての $t = 1, 2, \dots, n$ について、互いに独立であるとする。

以上の設定のもとで、 X_t は Y_t に影響を与えない変数であったとしても、次のような結果を得ることが知られている。

1. 回帰係数 β に関する t 統計量 (t 値) は、観測数が大きければ、大きな値となりやすい。無関係な変数間の回帰なので、 t 値は絶対値で小さい値となるべきである。しかし、逆に大きな t 値を示すことがあり、間違っただ推論になる傾向がある。
2. 決定係数 R^2 は変数間の関係がないのでゼロに近づくべきであるが、見せかけ回帰では 0 と 1 の何らかの間の値を取る。ゼロに近いとは限らず、逆に 1 に近い値を取る可能性も十分にある。
3. DW 比はゼロに近い値となる。すなわち、誤差項 u_t の強い正の系列相関があると判断される。

このように、全く意味のない回帰を、得られた推定結果から、あたかも意味のあるように解釈してしまうことが起こりうるのである。このような状況のもとでは、最小 2 乗推定量 $\hat{\beta}$ は β の一致推定量にはならないということも知られている。

見せかけ回帰を避けるためには、誤差項 u_t が定常である必要がある。これは、次節で述べる共和分という概念に密接に関連する。

14.3.3 共和分

Y_t , X_t ともに単位根をもつ、すなわち、 $Y_t \sim I(1)$, $X_t \sim I(1)$ とする。このとき、 Y_t と X_t の線形結合が定常、すなわち、 $Y_t - \beta X_t \sim I(0)$ であれば、 Y_t と X_t とは共和分 (cointegration) の関係にあるという。

重要な点は、次の回帰

$$Y_t = \beta X_t + u_t$$

を考えた場合、以下のようにまとめられる。

- (i) $u_t \sim I(1)$ のとき、 β の最小 2 乗推定量 $\hat{\beta}$ は一致性をもたず、しかも、回帰式の結果を評価すること自体に意味がなくなる。 → 見せかけ回帰
- (ii) $u_t \sim I(0)$ のとき、 β の最小 2 乗推定量 $\hat{\beta}$ は一致性を持ち、 Y_t と X_t との間には安定的な関係がある。 → 共和分回帰

(ii) の場合は、 $\beta \neq 0$ となることに注意せよ。もし $\beta = 0$ であれば、 $Y_t = u_t$ となり、 $Y_t \sim I(1)$ なので、 $u_t \sim I(1)$ となる必要がある。よって、 $u_t \sim I(0)$ となるためには、 $\beta \neq 0$ でなければならない。

このように、誤差項 u_t が定常かどうかによって、回帰式の意味が大きく異なることになる。したがって、回帰式に意味があるかどうかを調べるために、共和分の検定を行う必要がある。さまざまな検定方法が提案されているが、本書では、最も簡単な方法を紹介する (本書で解説する検定は、**エングル=グレンジャー検定**と呼ばれる)。

定数項を含めて、次の回帰式

$$Y_t = \alpha + \beta X_t + u_t$$

を考える。

共和分分析を行う前に、まず最初に、行わなければならないことは、単位根検定である。

すなわち、まず、 Y_t , X_t ともに、 $Y_t \sim I(1)$, $X_t \sim I(1)$ であることを検定によって確認する。

次に、最小 2 乗法によって α , β を推定し、残差 $e_t = Y_t - \hat{\alpha} - \hat{\beta} X_t$ を求める。

最後に、

$$\Delta e_t = \gamma e_{t-1} + v_t$$

から、帰無仮説 $H_0: \gamma = 0$ 、対立仮説 $H_1: \gamma < 0$ の単位根検定を行う。このとき、検定統計量

$$t_{\hat{\gamma}} = \frac{\hat{\gamma}}{Se(\hat{\gamma})}$$

の分布は、前節の単位根検定で用いた表に示された分布表とも t 分布表とも異なり、この場合は下記の表を用いる。

残差の単位根検定

k'	α	1		2	
		0.01	0.05	0.01	0.05
50		-4.32	-3.67	-4.84	-4.11
100		-4.07	-3.37	-4.45	-3.93
200		-4.00	-3.37	-4.35	-3.78
k'	α	3		4	
		0.01	0.05	0.01	0.05
50		-4.94	-4.35	-5.41	-4.76
100		-4.75	-4.22	-5.18	-4.58
200		-4.70	-4.18	-5.02	-4.48

計量経済学 練習問題

表では、標本数 (n) と説明変数の個数 ($k' = k - 1$) に依存して、

検定統計量 $t_{\hat{\gamma}}$ の分布の下側 1 パーセント点と 5 パーセント点が表示されている。

$t_{\hat{\gamma}}$ の統計値と表中の対応する数値とを比較して、残差 e_t の単位根検定が行われる。

分布表に上の表を用いることを除いて、前節で行った単位根検定をそのまま当てはめればよい。

すなわち、帰無仮説 $H_0 : \gamma = 0$ が棄却されれば、 $u_t \sim I(0)$ と判定され、回帰式の分析 ($\hat{\beta}$, R^2 , DW 比等) を行うことができる。

逆に、帰無仮説 $H_0 : \gamma = 0$ を棄却できなければ、 $u_t \sim I(1)$ と判定され、これ以上の回帰式の分析には全く意味がなくなる。

この場合は、 X_t とは別の $I(1)$ 変数を回帰式に含めなかったために、このような見せかけ回帰が生じたと考えられる。

1 n 組のデータ $(Y_1, X_1), (Y_2, X_2), \dots, (Y_n, X_n)$ 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、 $u_1, u_2, \dots, u_n$ 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	$Y_i X_i$	$\hat{Y}_i$	$\hat{u}_i$	$\hat{u}_i^2$
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 $\hat{\alpha}, \hat{\beta}$ 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 $\hat{\alpha}, \hat{\beta}$ の分散 $V(\hat{\alpha}), V(\hat{\beta})$ 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 $\bar{R}^2$ をそれぞれ求めよ。
- ⑥ $\hat{\beta}$ 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。