5.4.1　推定量の性質

$\beta_1, \beta_2, \cdots, \beta_k$ の最小二乗推定量は $\hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_k$ とする。

誤差項 (または，撹乱項) u_i の分散 σ^2 の推定量 s^2 は，

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

として表される。

このとき，

$\mathbf{E}(\hat{\beta}_j) = \beta_j$　　(不偏推定量)

$n \longrightarrow \infty$ のとき， $\hat{\beta}_j \longrightarrow \beta_j$　　(一致推定量)

(∗) $\mathbf{plim}\,\hat{\beta}_j = \beta_j$ と書く。

plim (「ピーリム」と読む) = 意味は「**probability limit** (確率極限)」

$\mathbf{E}(s^2) = \sigma^2$　　(不偏推定量)

$n \longrightarrow \infty$ のとき， $s^2 \longrightarrow \sigma^2$　　(一致推定量)

(∗) $\mathbf{plim}\; s^2 = \sigma^2$

を証明することが出来る。(証明略)

(* 注) ベクトルの確率変数の期待値・分散について： $k \times 1$ **ベクトルの確率変数** $X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{pmatrix}$ **の平均・分散を考える。**

$i = 1, 2, \cdots, k$ **について，** $\mathbf{E}(X_i) = \mu_i$ **とする。**

$$\mathbf{E}(X) = \mathbf{E}\begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{pmatrix} = \begin{pmatrix} \mathbf{E}(X_1) \\ \mathbf{E}(X_2) \\ \vdots \\ \mathbf{E}(X_k) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_k \end{pmatrix} = \mu$$

$$\mathbf{V}(X) = \mathbf{E}\Big((X-\mu)(X-\mu)'\Big) = \mathbf{E}\left(\begin{pmatrix} X_1-\mu_1 \\ X_2-\mu_2 \\ \vdots \\ X_k-\mu_k \end{pmatrix}\begin{pmatrix} X_1-\mu_1 & X_2-\mu_2 & \cdots & X_k-\mu_k \end{pmatrix}\right)$$

$$= \mathbf{E}\begin{pmatrix} (X_1-\mu_1)^2 & (X_1-\mu_1)(X_2-\mu_2) & \cdots & (X_1-\mu_1)(X_k-\mu_k) \\ (X_2-\mu_2)(X_1-\mu_1) & (X_2-\mu_2)^2 & \cdots & (X_2-\mu_2)(X_k-\mu_k) \\ \vdots & \vdots & \ddots & \vdots \\ (X_k-\mu_k)(X_1-\mu_1) & (X_k-\mu_k)(X_2-\mu_2) & \cdots & (X_k-\mu_k)^2 \end{pmatrix}$$

$$= \begin{pmatrix} \mathbf{E}\Big((X_1-\mu_1)^2\Big) & \mathbf{E}\Big((X_1-\mu_1)(X_2-\mu_2)\Big) & \cdots & \mathbf{E}\Big((X_1-\mu_1)(X_k-\mu_k)\Big) \\ \mathbf{E}\Big((X_2-\mu_2)(X_1-\mu_1)\Big) & \mathbf{E}\Big((X_2-\mu_2)^2\Big) & \cdots & \mathbf{E}\Big((X_2-\mu_2)(X_k-\mu_k)\Big) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{E}\Big((X_k-\mu_k)(X_1-\mu_1)\Big) & \mathbf{E}\Big((X_k-\mu_k)(X_2-\mu_2)\Big) & \cdots & \mathbf{E}\Big((X_k-\mu_k)^2\Big) \end{pmatrix}$$

$$= \begin{pmatrix} \mathbf{V}(X_1) & \mathbf{Cov}(X_1, X_2) & \cdots & \mathbf{Cov}(X_1, X_k) \\ \mathbf{Cov}(X_2, X_1) & \mathbf{V}(X_2) & \cdots & \mathbf{Cov}(X_2, X_k) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{Cov}(X_k, X_1) & \mathbf{Cov}(X_k, X_2) & \cdots & \mathbf{V}(X_k) \end{pmatrix} = \Sigma$$

このように $\mathrm{E}(X) = \mu$，$\mathrm{V}(X) = \Sigma$ の次元はそれぞれ $k \times 1$ ベクトル，$k \times k$ 行列となる。

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

$$\mathbf{V}\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{pmatrix} = \begin{pmatrix} \mathbf{V}(\hat{\beta}_1) & \mathbf{Cov}(\hat{\beta}_1, \hat{\beta}_2) & \cdots & \mathbf{Cov}(\hat{\beta}_1, \hat{\beta}_k) \\ \mathbf{Cov}(\hat{\beta}_2, \hat{\beta}_1) & \mathbf{V}(\hat{\beta}_2) & \cdots & \mathbf{Cov}(\hat{\beta}_2, \hat{\beta}_k) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{Cov}(\hat{\beta}_k, \hat{\beta}_1) & \mathbf{Cov}(\hat{\beta}_k, \hat{\beta}_2) & \cdots & \mathbf{V}(\hat{\beta}_k) \end{pmatrix}$$

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

最後の等号の右辺の逆行列 A の i 行 j 列目の要素を a_{ij} としたとき，$\hat{\beta}_j$ の分散は，

$$\mathbf{V}(\hat{\beta}_j) = \sigma^2 a_{jj}$$

となる（証明略）。このとき，

$$\hat{\beta}_j \sim N(\beta_j, \sigma^2 a_{jj})$$

となり，標準化すると，

$$\frac{\hat{\beta}_j - \beta_j}{\sigma\sqrt{a_{jj}}} \sim N(0, 1)$$

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 の独立である (証明略)，

さらに，

(* 復習) t 分布について（再掲）：

$Z \sim N(0,1)$，$U \sim \chi^2(k)$，Z と U は独立のとき，

$T = \dfrac{Z}{\sqrt{U/k}} \sim t(k)$ となる。

を利用すると，

$$\frac{\dfrac{\hat{\beta}_j - \beta_j}{\sigma\sqrt{a_{jj}}}}{\sqrt{\dfrac{(n-k)s^2}{\sigma^2}/(n-k)}} = \frac{\hat{\beta}_j - \beta_j}{s\sqrt{a_{jj}}} \sim t(n-k)$$

が得られる。

　このように，$t(n-k)$ を用いることによって，通常の区間推定や仮説検定を行うことが出来る。

$s\sqrt{a_{jj}}$ は $\hat{\beta}_j$ の標準誤差である。

すなわち，$s\sqrt{a_{jj}}$ は，単回帰の場合の $s_{\hat{\alpha}}$，$s_{\hat{\beta}}$ に対応する。

● $\hat{\beta}_j$ の区間推定：　$\dfrac{\hat{\beta}_j-\beta_j}{s\sqrt{a_{jj}}}\sim t(n-k)$ なので，

$$\mathbf{Prob}\Big(-t_{\alpha/2}(n-k)<\frac{\hat{\beta}_j-\beta_j}{s\sqrt{a_{jj}}}<t_{\alpha/2}(n-k)\Big)=1-\alpha$$

ただし，$t_{\alpha/2}(n-k)$ は自由度 $n-k$ の t 分布の $100\times\dfrac{\alpha}{2}$ % 点の値とする。

β_j について解くと，

$$\mathbf{Prob}\left(\hat{\beta}_j - t_{\alpha/2}(n-k) \times s\sqrt{a_{jj}} < \beta_j < \hat{\beta}_j + t_{\alpha/2}(n-k) \times s\sqrt{a_{jj}}\right) = 1 - \alpha$$

を得る。

$\hat{\beta}_j$，s を推定値で置き換えて，信頼係数 $1-\alpha$ の β_j の区間推定は，

$$\left(\hat{\beta}_j - t_{\alpha/2}(n-k) \times s\sqrt{a_{jj}},\ \hat{\beta}_j + t_{\alpha/2}(n-k) \times s\sqrt{a_{jj}}\right)$$

となる。

● $\hat{\beta}_j$ **の仮説検定：　帰無仮説** H_0：$\beta_j = \beta_{*j}$ **を検定することを考える（**β_{*j} **は分析者が設定する値とする）。**

$\dfrac{\hat{\beta}_j - \beta_j}{s\sqrt{a_{jj}}} \sim t(n-k)$ **なので，帰無仮説が正しいもとで（すなわち，**$\beta_j = \beta_{*j}$**），**

$$\frac{\hat{\beta}_j - \beta_{*j}}{s\sqrt{a_{jj}}} \sim t(n-k)$$

となる。

$\hat{\beta}_j$**，**s **を推定値で置き換えて，**

$$\left|\frac{\hat{\beta}_j - \beta_{*j}}{s\sqrt{a_{jj}}}\right| > t_{\alpha/2}(n-k)$$

のとき，有意水準 $100 \times \alpha\%$ **で帰無仮説** H_0：$\beta_j = \beta_{*j}$ **を棄却する（帰無仮説が起こる確率は**

$100 \times \alpha\%$ 以下ということになるので）。

(注)　$u_1, u_2, \cdots, u_n$ は互いに独立で，$u_i \sim N(0, \sigma^2)$ のとき，

$$\sum_{i=1}^{n}\Big(\frac{u_i}{\sigma}\Big)^2 \sim \chi^2(n)$$

となる。u_i をその推定量 $\hat{u}_i$ で置き換えると，

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

ただし，s^2 は σ^2 の推定量で，$s^2 = \dfrac{1}{n-k}\displaystyle\sum_{i=1}^{n}\hat{u}_i^2 = \dfrac{1}{n-k}\sum_{i=1}^{n}(Y_i - \hat{\beta}_1 X_{1i} - \hat{\beta}_2 X_{2i} - \cdots - \hat{\beta}_k X_{ki})^2$ である。$\hat{u}_i$ を得るためには，$\hat{\beta}_1$，$\hat{\beta}_2$，$\cdots$，$\hat{\beta}_k$ （k 個のパラメータ推定量）を求めなければならない。$n-k$ （= データ数 n − パラメータ数 k）を自由度と呼ばれる。

(注)　s^2 が σ^2 の不偏推定量，一致推定量であることは，

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

を利用すれば簡単に証明できる。

(* 復習) カイ二乗分布の平均・分散について（再掲）：

$U \sim \chi^2(k)$ のとき，$\mathbf{E}(U) = k$，$\mathbf{V}(U) = 2k$ となる。

この重回帰の場合は，

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

なので，

$$\mathbf{E}\Big(\frac{(n-k)s^2}{\sigma^2}\Big) = n-k, \qquad \mathbf{V}\Big(\frac{(n-k)s^2}{\sigma^2}\Big) = 2(n-k)$$

すなわち，

$$\mathbf{E}\Big(\frac{(n-k)s^2}{\sigma^2}\Big) = \frac{n-k}{\sigma^2}\mathbf{E}(s^2) = n-k, \qquad \mathbf{V}\Big(\frac{(n-k)s^2}{\sigma^2}\Big) = \Big(\frac{n-k}{\sigma^2}\Big)^2\mathbf{V}(s^2) = 2(n-k)$$

から

$$\mathbf{E}(s^2) = (n-k)\times\frac{\sigma^2}{n-k} = \sigma^2, \qquad \mathbf{V}(s^2) = 2(n-k)\times\Big(\frac{\sigma^2}{n-k}\Big)^2 = \frac{2\sigma^4}{n-k}$$

となる。

$\mathbf{E}(s^2) = \sigma^2$ で，かつ，$n \longrightarrow \infty$ のとき $\mathbf{V}(s^2) \longrightarrow 0$ となるので，s^2 は不偏推定量かつ一致推定量である。

第6章　ダミー変数，関数形，その他

6.1　ダミー変数

6.1.1　異常値ダミー

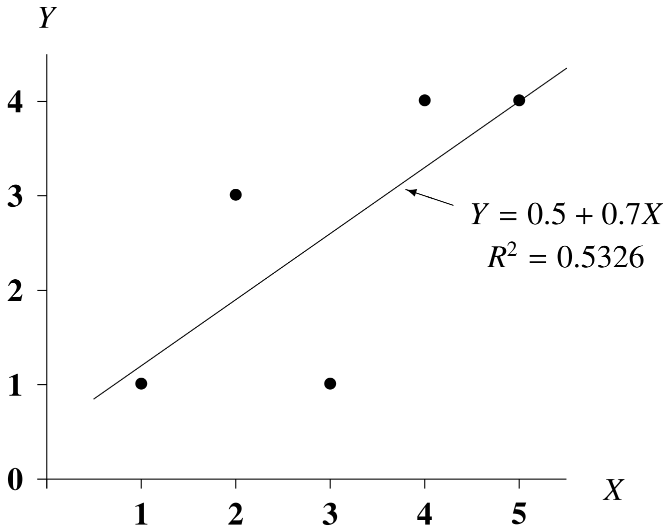
　ダミー変数とは，0 と 1 から成る変数のことである。データに異常値が含まれている場合，ダミー変数を使う。

　例えば，今までの数値例を使って説明する。

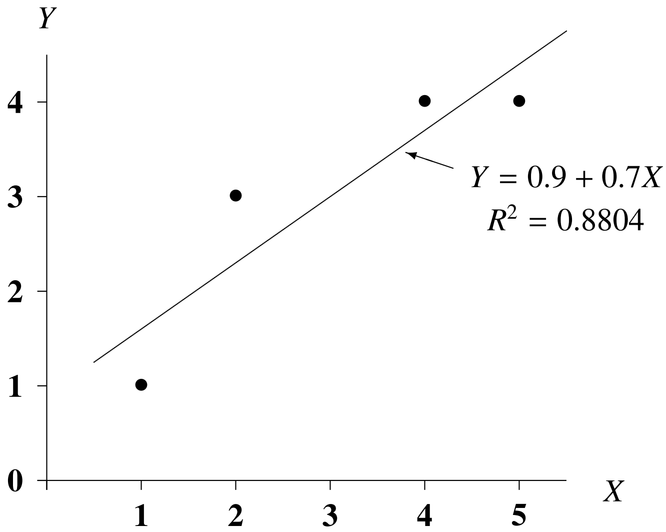
i	X_i	Y_i	X_i^2	X_iY_i	$\hat{Y}_i$	$\hat{u}_i$
1	**5**	**4**	**25**	**20**	**4.0**	**0.0**
2	**1**	**1**	**1**	**1**	**1.2**	-0.2
3	**3**	**1**	**9**	**3**	**2.6**	-1.6
4	**2**	**3**	**4**	**6**	**1.9**	**1.1**
5	**4**	**4**	**16**	**16**	**3.3**	**0.7**
合計	$\sum X_i$	$\sum Y_i$	$\sum X_i^2$	$\sum X_iY_i$	$\sum \hat{Y}_i$	$\sum \hat{u}_i$
	15	**13**	**55**	**46**	**13**	**0.0**
平均	$\overline{X}$	$\overline{Y}$				
	3	**2.6**				

$i = 3$ **のデータ** $(X_3, Y_3) = (3, 1)$ **について，直線** $Y = 0.5 + 0.7X$ **との縦軸方向の垂直距離，す**

なわち，残差 $\hat{u}_3 = -1.6$ が絶対値で最も大きくなっている。



$i = 3$ のデータを除いて，$n = 4$ 個のデータを用いて最小二乗法で推定してみる。


Y
4
3
2
1
0
1
2
3
4
5
X
$Y = 0.9 + 0.7X$
$R^2 = 0.8804$

今まで見てきた通り，$i = 1, 2, 3, 4, 5$ の全部のデータを使って，

$$
\begin{array}{cccccc}
Y_i = & \mathbf{0.5} & + & \mathbf{0.7} & X_i, \\
 & \mathbf{(0.398)} & & \mathbf{(1.849)} & \\
\end{array}
$$

$$R^2 = 0.5326, \qquad \overline{R}^2 = 0.3768, \qquad s^2 = 1.197^2$$

と推定される。ただし，係数の推定値の下の括弧内は t 値を表すものとする。

一方，$i = 3$ を除いて，$i = 1, 2, 4, 5$ の 4 組のデータを使うと，

$$
\begin{array}{cccccc}
Y_i = & \mathbf{0.9} & + & \mathbf{0.7} & X_i, \\
 & \mathbf{(1.132)} & & \mathbf{(2.985)} & \\
\end{array}
$$

$$R^2 = 0.8804, \qquad \overline{R}^2 = 0.7609, \qquad s^2 = 0.742^2$$

となる。ただし，係数の推定値の下の括弧内は t 値を表すものとする。

このように，定数項の結果変わる（傾きの値が変化しなかったのは，単なる偶然）。

3 番目のデータが，回帰直線から離れている（すなわち，異常値）ものとして考えて，

$$D_i = \begin{cases} 0, & i = 1, 2, 4, 5 \text{ のとき} \\ 1, & i = 3 \text{ のとき} \end{cases}$$

というダミー変数を作り，

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

を推定する。γ の推定値 $\hat{\gamma}$ の有意性を調べることによって，3 番目のデータが異常値かどうかを検定することができる。

この回帰式の意味は，

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

となる。3 番目のデータのときに定数項（切片）が γ だけシフトする。

推定結果は，

$$Y_i = \underset{(1.132)}{\mathbf{0.9}} + \underset{(2.985)}{\mathbf{0.7}} X_i - \underset{(2.412)}{\mathbf{2.0}} D_i,$$

$$R^2 = 0.8804, \qquad \overline{R}^2 = 0.7609, \qquad s^2 = 0.742^2$$

となる。ただし，係数の推定値の下の括弧内は t 値を表すものとする。

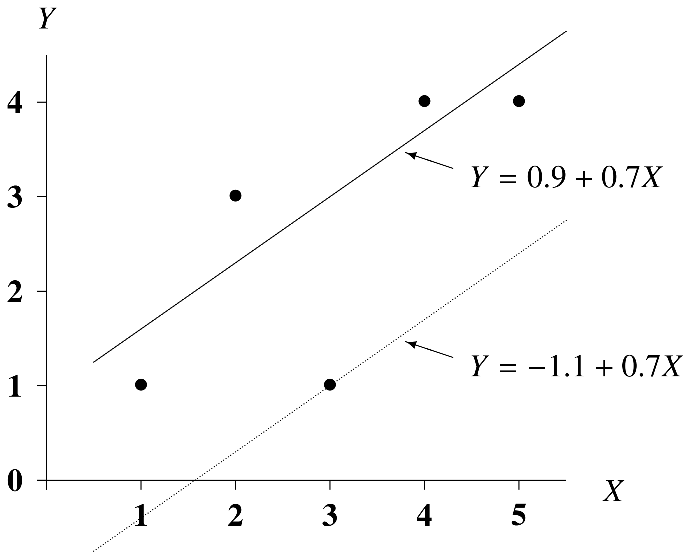
推定結果からみると，D_i の係数推定値の t 値は 2.412 で，この場合，自由度 $n-k=5-3=2$ の t 分布の 2.5 % 点 $t_{0.025}(2)=4.3027$ と比較することになる。

$2.412<4.3027$ なので，有意水準 5 % で $H_0:\ \gamma=0$ を棄却できない。

すなわち，$i=3$ のデータは異常値とは認められない。

この場合，$\hat{Y}_3=Y_3$，すなわち，$\hat{u}_3=0$ となることに注意。

グラフに描くと，$i=3$ のデータを通る平行移動した直線が追加される。


Y
4
3
2
1
0
1
2
3
4
5
X
Y = 0.9 + 0.7X
Y = −1.1 + 0.7X

6.1.2　構造変化ダミー

経済構造がある時期から変化した場合もダミー変数を使って，処理することができる。

この場合，添え字 i は時間を表す。

$n = 20$ として，例えば，9 期目以前と以降とで，経済構造が変化している場合を考える。

$$D_i = \begin{cases} 0, & i = 1, 2, \cdots, 9 \text{ のとき} \\ 1, & i = 10, 11, \cdots, 20 \text{ のとき} \end{cases}$$

という変数を作り，

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

$$
= \begin{cases} \alpha + \beta X_i + u_i, & i = 1, 2, \cdots, 9 \text{ のとき} \\ (\alpha + \delta) + \beta X_i + u_i, & i = 10, 11, \cdots, 20 \text{ のとき} \end{cases}
$$

を推定する（定数項だけが変化したと考えた場合）。または，

$$
\begin{aligned}
Y_i &= \alpha + \delta D_i + \beta X_i + \gamma D_i X_i + u_i \\
&= \begin{cases} \alpha + \beta X_i + u_i, & i = 1, 2, \cdots, 9 \text{ のとき} \\ (\alpha + \delta) + (\beta + \gamma) X_i + u_i, & i = 10, 11, \cdots, 20 \text{ のとき} \end{cases}
\end{aligned}
$$

を推定する（定数項も係数も変化）。

δ や γ の推定値の有意性を調べることによって，構造変化の検定を行うことができる。

上の例でデータを示すと，

i	Y_i	X_i	D_i	D_iX_i
1	Y_1	X_1	0	0
2	Y_2	X_2	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
9	Y_9	X_9	0	0
10	Y_{10}	X_{10}	1	X_{10}
11	Y_{11}	X_{11}	1	X_{11}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
20	Y_{20}	X_{20}	1	X_{20}

となる。

数値例：

i	X_i	Y_i	D_i	D_iX_i
1	1	1	0	0
2	1	2	0	0
3	1	0	0	0
4	2	1	0	0
5	2	2	0	0
6	2	3	0	0
7	3	2	0	0
8	3	3	0	0
9	3	4	0	0
10	4	4	1	4
11	4	5	1	4
12	4	6	1	4
13	5	5	1	5
14	5	6	1	5
15	5	7	1	5
16	6	5	1	6
17	6	6	1	6
18	6	7	1	6
19	7	6	1	7
20	7	7	1	7

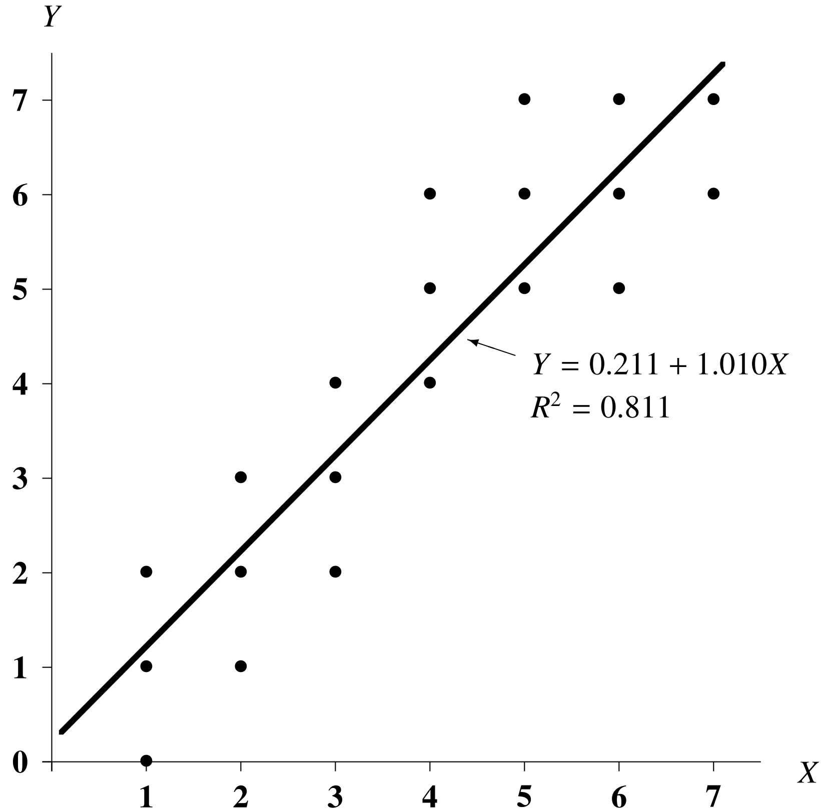
20 組全部のデータを用いて推定すると，推定結果は，ダミー変数を用いなければ，

$$Y_i = \underset{(0.427)}{\mathbf{0.211}} + \underset{(8.784)}{\mathbf{1.010}}\ X_i$$

$$R^2 = 0.8108, \qquad \overline{R}^2 = 0.8003, \qquad s^2 = 0.9928^2$$

となる（自由度は $n - k = 20 - 2 = 18$**）。ただし，係数の推定値の下の括弧内は** t **値を表すものとする。**

散布図と回帰直線は次ページであるが，一見何の問題もないように見える。

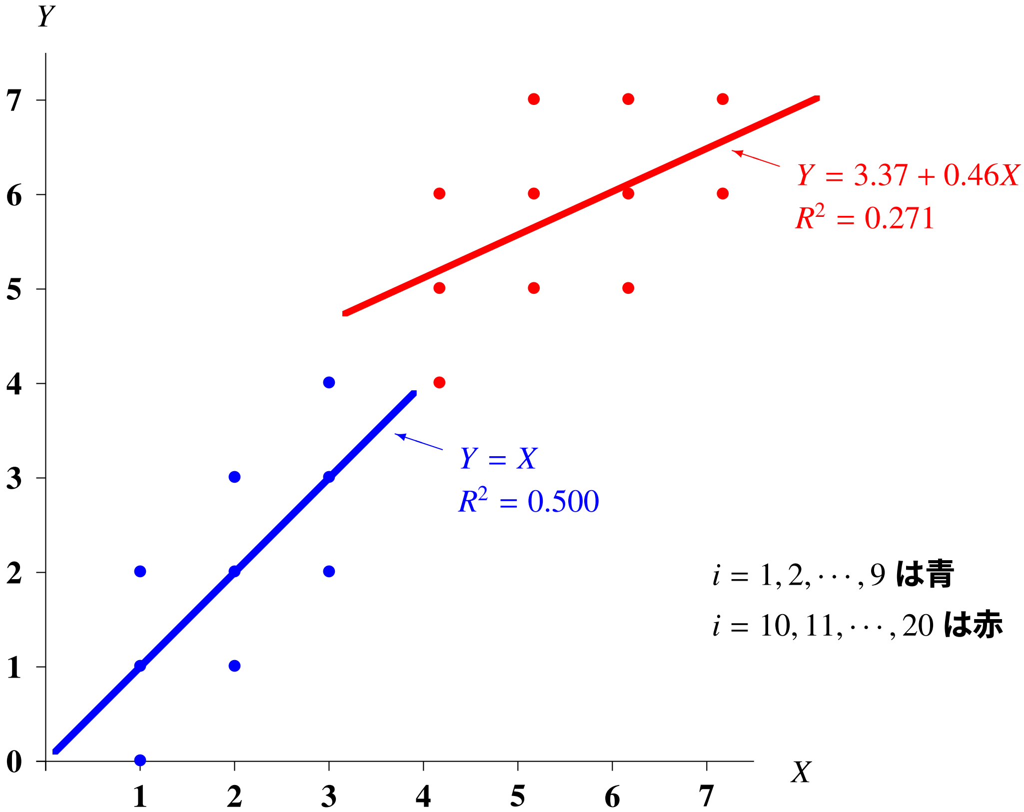

Y
7
6
5
4
3
2
1
0
1
2
3
4
5
6
7
X
Y = 0.211 + 1.010X
R2 = 0.811

数値例の表によると，左下の方のデータは推定期間の前半のデータ（$i = 1, 2, \cdots, 9$），右上のデータは後半のデータ（$i = 10, 11, \cdots, 20$）と想定している。

2 つの部分の回帰式は同じになるかどうか？

2 つの期に分けて，別々に推定する。 $\Longrightarrow$ 次ページ

この場合，前半部分の自由度は $n - k = 9 - 2 = 7$，後半部分は $n - k = 11 - 2 = 9$ となる。



Y
7
6
5
4
3
2
1
0
1
2
3
4
5
6
7
X
$Y = 3.37 + 0.46X$
$R^2 = 0.271$
$Y = X$
$R^2 = 0.500$
$i = 1, 2, \cdots, 9$ は青
$i = 10, 11, \cdots, 20$ は赤

推定結果は，ダミー変数を用いなければ，

$$Y_i = \underset{(0.427)}{\mathbf{0.211}} + \underset{(8.784)}{\mathbf{1.010}} X_i$$

$$R^2 = 0.8108, \qquad \overline{R}^2 = 0.8003, \qquad s^2 = 0.9928^2$$

となる（自由度は $n - k = 20 - 2 = 18$）。ただし，係数の推定値の下の括弧内は t 値を表すものとする。

また，ダミー変数を用いて，切片と傾きの両方が変化したと考えて推定すると，

$$Y_i = \underset{(0.000)}{\mathbf{0.000}} + \underset{(2.715)}{\mathbf{1.000}} X_i + \underset{(2.101)}{\mathbf{3.370}} D_i - \underset{(1.214)}{\mathbf{0.543}} D_iX_i,$$

$$R^2 = 0.8612, \qquad \overline{R}^2 = 0.8351, \qquad s^2 = 0.9021^2$$

となる（自由度は $n - k = 20 - 4 = 16$）。ただし，係数の推定値の下の括弧内は t 値を表すものとする。

教科書『計量経済学』の付表 (p.352) から，$t_{0.025}(16) = 2.120$ である。

$2.101 < 2.120$，$1.214 < 2.120$ なので，$H_0:\ \delta = 0$，$H_0:\ \gamma = 0$ の帰無仮説を共に有意水準 5 % で棄却できない。

したがって，切片・傾き共に構造変化があったとは言えない。