

# 第7章 最小二乗法における一般的な仮定 について

最小自乗法の仮定：

- $u_1, u_2, \dots, u_n$  は互いに独立
- $u_i$  の分散は一定 (すなわち, すべての  $i$  について  $\sigma^2$ )

- $u_i \sim N(0, \sigma^2)$  (すなわち, すべての  $i$  について  $u_i$  は正規分布)
- $X_i$  は非確率変数

それぞれの仮定が崩れた場合, どのようなことが起こるかを考察する。

## 7.1 誤差項：系列相関について

### 7.1.1 系列相関の意味

最小自乗法の仮定の一つに, 「攪乱項  $u_1, u_2, \dots, u_n$  は互いに独立に分布する」というものがあつた。

データの順番が重要な場合、すなわち、時系列データの場合を考える。

ただし、時系列データでなくても、順番が重要な場合は考えられる（地理的な位置関係など）が、ここでは時系列データを念頭に置いて欲しい。

$u_1, u_2, \dots, u_n$  に系列相関がある場合、次の2つが考えられる。

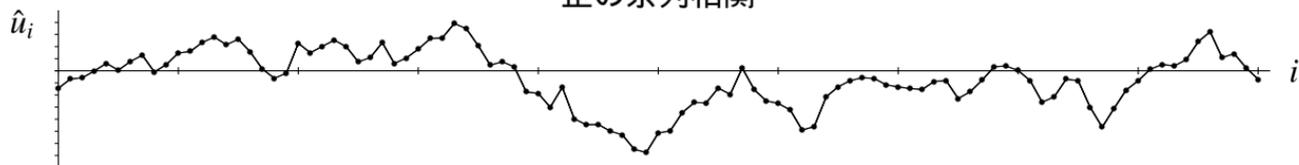
- それぞれの符号が順に、 $+++---++---++\dots$  のように、プラスが連続で続いた後で、マイナスが連続で続くというような場合、 $u_1, u_2, \dots, u_n$  は「正の系列相関がある」と言う。
- また、それぞれの符号が順に、 $+ - + - + - + - + - + - \dots$  のように交互にプラス、マイナスになる場合、 $u_1, u_2, \dots, u_n$  は「負の系列相関がある」と言う。

特徴として、 $u_1, u_2, \dots, u_i$  から  $u_{i+1}$  の符号がある程度予想できる。

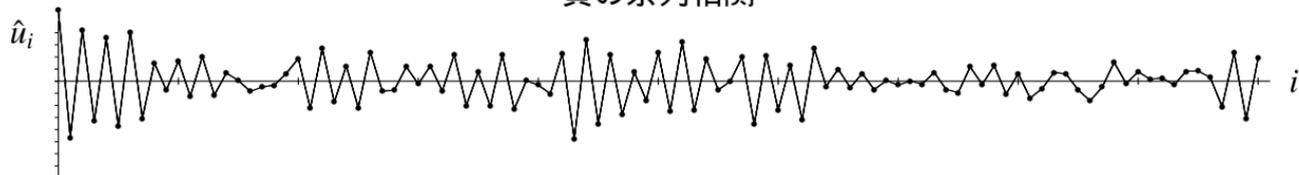
これは「 $u_1, u_2, \dots, u_n$  は互いに独立に分布する」という仮定に反する。

横軸が時間  $i$ 、縦軸が残差  $\hat{u}_i$  としてグラフを描くと次のようになる。

正の系列相関



負の系列相関



## 7.1.2 系列相関の指標：DW について

ダービン・ワトソン (DW) 比とは、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。

すなわち、ダービン・ワトソン (DW) 比とは、回帰式が

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

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

のときに、帰無仮説  $H_0: \rho = 0$ 、対立仮説  $H_1: \rho \neq 0$  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。

- $\rho > 0$  のとき,  $u_1, u_2, \dots, u_n$  は「正の系列相関がある」と言う (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)。
- $\rho < 0$  のとき,  $u_1, u_2, \dots, u_n$  は「負の系列相関がある」と言う (プラスとマイナスが交互になる場合)。
- $\rho = 0$  のとき,  $u_1, u_2, \dots, u_n$  は「系列相関がない」と言う (符号の法則性がない場合)。

ダービン・ワトソン ( $DW$ ) 比の定義は次の通りである。

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

ただし,  $\hat{u}_i = Y_i - \hat{\beta}_1 X_{1i} - \hat{\beta}_2 X_{2i} - \dots - \hat{\beta}_k X_{ki}$ ,  $j = 1, 2, \dots, k$  について  $\hat{\beta}_j$  は  $\beta_j$  の最小二乗推定値とする。

DW 比は，近似的に次のように書き換えられる。

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

以下の 2 つの近似が用いられる。

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

すなわち， $\hat{\rho}$  は， $\hat{u}_i$  を被説明変数， $\hat{u}_{i-1}$  を説明変数としたときの  $\hat{u}_{i-1}$  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， $\rho$  の最小二乗推定値  $\hat{\rho}$  を求めることになる。

具体的には，下記の最小化問題を解いたときの  $\hat{\rho}$  の解となる。

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

DW 比には，次のような性質がある。

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

$u_i = \rho u_{i-1} + \epsilon_i$  について， $\rho$  は  $u_i$  と  $u_{i-1}$  との相関係数：

$$\rho = \frac{\text{Cov}(u_i, u_{i-1})}{\sqrt{\mathbf{V}(u_i)} \sqrt{\mathbf{V}(u_{i-1})}} = \frac{\text{Cov}(u_i, u_{i-1})}{\mathbf{V}(u_i)}$$

$V(u_i) = V(u_{i-1})$  に注意。すなわち ,  $-1 < \rho < 1$  となる。

$$\hat{\rho} = \frac{\sum_{i=2}^n \hat{u}_i \hat{u}_{i-1}}{\sum_{i=2}^n \hat{u}_{i-1}^2} = \frac{\frac{1}{n-1} \sum_{i=2}^n \hat{u}_i \hat{u}_{i-1}}{\frac{1}{n-1} \sum_{i=2}^n \hat{u}_{i-1}^2} \longrightarrow \frac{\mathbf{Cov}(u_i, u_{i-1})}{\mathbf{V}(u_i)} = \rho$$

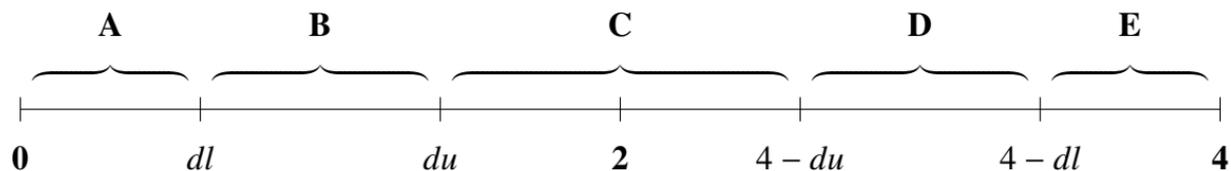
$\hat{\rho}$  は  $\rho$  の一致推定量となっている (証明略)。

$$DW \approx 2(1 - \hat{\rho}) \longrightarrow 2(1 - \rho)$$

よって ,  $0 < DW < 4$  となる。

$\rho = 0$  (系列相関なし) のとき ,  $DW = 2$  となる。

DW テストでは ,  $dl$  と  $du$  によって下記のように A , B , C , D , E の 5 つの領域に分類される。



$dl$  と  $du$  は,  $n$  と  $k$  に依存し, 統計表から得られる数値である (統計表は後述)。

- $DW$  が 2 前後の場合 (領域 C),  $\rho = 0$  を意味し, 誤差項に系列相関はない ( $\rho = 0$ ) と判定される。
- $DW$  がゼロに近い場合 (領域 A),  $\rho$  は 1 に近いことを意味し, 誤差項に正の系列相関がある ( $\rho > 0$ ) と判定される。
- $DW$  が 4 に近い場合 (領域 E),  $\rho$  は  $-1$  に近いことを意味し, 誤差項に負の系列相関がある ( $\rho < 0$ ) と判定される。

- $DW$  は 2 より小さいがゼロに近いとは言えない場合 (領域 **B**) , 誤差項に正の系列相関があるとは言えない ( $\rho > 0$  という傾向はある)。
- $DW$  は 2 より大きいが 4 に近いとは言えない場合 (領域 **D**) , 誤差項に負の系列相関があるとは言えない ( $\rho < 0$  という傾向はある)。

正確な判定には, データ数  $n$  とパラメータ数  $k$  (正確には, 定数項を除くパラメータ数  $k'$ ) に依存して,  $dl$  と  $du$  の値が決まり, 次ページの表 7.1 のように 5 つの領域に分類される。

表 7.1 : ダービン・ワトソン統計量の 5 % 点の上限と下限

(1)  $k' = 1$

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

(2)  $k' = 2$

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

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

ただし、 $DW$  は 2 を中心に対象になっているので、データ数  $n$  と定数項を除くパラメータ数  $k'$  によって、 $dl$  と  $du$  が得られれば、5 つの領域を求めることができる。

次ページの表 7.2 に、より多くのデータ数  $n$  と定数項を除くパラメータ数  $k'$  との組み合わせを載せている。

表 7.2 : ダービン・ワトソン統計量の 5% 点の上限と下限

| n  | k' = 1 |       | k' = 2 |       | k' = 3 |       | k' = 4 |       | k' = 5 |       | k' = 6 |       | k' = 7 |       | k' = 8 |       | k' = 9 |       | k' = 10 |       | k' = 11 |       | k' = 12 |       |
|----|--------|-------|--------|-------|--------|-------|--------|-------|--------|-------|--------|-------|--------|-------|--------|-------|--------|-------|---------|-------|---------|-------|---------|-------|
|    | dl     | du    | dl      | du    | dl      | du    | dl      | du    |
| 6  | 0.610  | 1.400 | —      | —     | —      | —     | —      | —     | —      | —     | —      | —     | —      | —     | —      | —     | —      | —     | —       | —     | —       | —     | —       | —     |
| 7  | 0.700  | 1.356 | 0.467  | 1.896 | —      | —     | —      | —     | —      | —     | —      | —     | —      | —     | —      | —     | —      | —     | —       | —     | —       | —     | —       | —     |
| 8  | 0.763  | 1.332 | 0.559  | 1.777 | 0.367  | 2.287 | —      | —     | —      | —     | —      | —     | —      | —     | —      | —     | —      | —     | —       | —     | —       | —     | —       | —     |
| 9  | 0.824  | 1.320 | 0.629  | 1.699 | 0.455  | 2.128 | 0.296  | 2.588 | —      | —     | —      | —     | —      | —     | —      | —     | —      | —     | —       | —     | —       | —     | —       | —     |
| 10 | 0.879  | 1.320 | 0.697  | 1.641 | 0.525  | 2.016 | 0.376  | 2.414 | 0.243  | 2.822 | —      | —     | —      | —     | —      | —     | —      | —     | —       | —     | —       | —     | —       | —     |
| 11 | 0.927  | 1.324 | 0.758  | 1.604 | 0.595  | 1.928 | 0.444  | 2.283 | 0.315  | 2.645 | 0.203  | 3.004 | —      | —     | —      | —     | —      | —     | —       | —     | —       | —     | —       | —     |
| 12 | 0.971  | 1.331 | 0.812  | 1.579 | 0.658  | 1.864 | 0.512  | 2.177 | 0.380  | 2.506 | 0.268  | 2.832 | 0.171  | 3.149 | —      | —     | —      | —     | —       | —     | —       | —     | —       | —     |
| 13 | 1.010  | 1.340 | 0.861  | 1.562 | 0.715  | 1.816 | 0.574  | 2.094 | 0.444  | 2.390 | 0.328  | 2.692 | 0.230  | 2.985 | 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  | 1.750 | 0.685  | 1.977 | 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  | 1.539 | 0.857  | 1.728 | 0.734  | 1.935 | 0.615  | 2.157 | 0.502  | 2.388 | 0.398  | 2.624 | 0.304  | 2.860 | 0.222  | 3.090 | 0.155   | 3.304 | 0.098   | 3.503 | —       | —     |
| 17 | 1.133  | 1.381 | 1.015  | 1.536 | 0.897  | 1.710 | 0.779  | 1.900 | 0.664  | 2.104 | 0.554  | 2.318 | 0.451  | 2.537 | 0.356  | 2.757 | 0.272  | 2.975 | 0.198   | 3.184 | 0.138   | 3.378 | 0.087   | 3.557 |
| 18 | 1.158  | 1.391 | 1.046  | 1.535 | 0.933  | 1.696 | 0.820  | 1.872 | 0.710  | 2.060 | 0.603  | 2.257 | 0.502  | 2.461 | 0.407  | 2.668 | 0.321  | 2.873 | 0.244   | 3.073 | 0.177   | 3.265 | 0.123   | 3.441 |
| 19 | 1.180  | 1.401 | 1.074  | 1.536 | 0.967  | 1.685 | 0.859  | 1.848 | 0.752  | 2.023 | 0.649  | 2.206 | 0.549  | 2.396 | 0.456  | 2.589 | 0.369  | 2.783 | 0.290   | 2.974 | 0.220   | 3.159 | 0.160   | 3.335 |
| 20 | 1.201  | 1.411 | 1.100  | 1.537 | 0.998  | 1.676 | 0.894  | 1.828 | 0.792  | 1.991 | 0.691  | 2.162 | 0.595  | 2.339 | 0.502  | 2.521 | 0.416  | 2.704 | 0.336   | 2.885 | 0.263   | 3.063 | 0.200   | 3.234 |
| 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 | 0.546  | 2.461 | 0.461  | 2.633 | 0.380   | 2.806 | 0.307   | 2.976 | 0.240   | 3.141 |
| 22 | 1.239  | 1.429 | 1.147  | 1.541 | 1.053  | 1.664 | 0.958  | 1.797 | 0.863  | 1.940 | 0.769  | 2.090 | 0.677  | 2.246 | 0.588  | 2.407 | 0.504  | 2.571 | 0.424   | 2.735 | 0.349   | 2.897 | 0.281   | 3.057 |
| 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  | 2.360 | 0.545  | 2.514 | 0.465   | 2.670 | 0.391   | 2.826 | 0.322   | 2.979 |
| 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 |
| 25 | 1.288  | 1.454 | 1.206  | 1.550 | 1.123  | 1.654 | 1.038  | 1.767 | 0.953  | 1.886 | 0.868  | 2.013 | 0.784  | 2.144 | 0.702  | 2.280 | 0.621  | 2.419 | 0.544   | 2.560 | 0.470   | 2.702 | 0.400   | 2.844 |
| 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 |
| 27 | 1.316  | 1.469 | 1.240  | 1.556 | 1.162  | 1.651 | 1.084  | 1.753 | 1.004  | 1.861 | 0.925  | 1.974 | 0.845  | 2.093 | 0.767  | 2.216 | 0.691  | 2.342 | 0.616   | 2.470 | 0.544   | 2.600 | 0.475   | 2.730 |
| 28 | 1.328  | 1.476 | 1.255  | 1.560 | 1.181  | 1.650 | 1.104  | 1.747 | 1.028  | 1.850 | 0.951  | 1.958 | 0.874  | 2.071 | 0.798  | 2.188 | 0.723  | 2.309 | 0.650   | 2.431 | 0.578   | 2.555 | 0.510   | 2.680 |
| 29 | 1.341  | 1.483 | 1.270  | 1.563 | 1.198  | 1.650 | 1.124  | 1.743 | 1.050  | 1.841 | 0.975  | 1.944 | 0.900  | 2.052 | 0.826  | 2.164 | 0.753  | 2.278 | 0.682   | 2.396 | 0.612   | 2.515 | 0.544   | 2.634 |
| 30 | 1.352  | 1.489 | 1.284  | 1.567 | 1.214  | 1.650 | 1.143  | 1.739 | 1.071  | 1.833 | 0.998  | 1.931 | 0.926  | 2.034 | 0.854  | 2.141 | 0.782  | 2.251 | 0.712   | 2.363 | 0.643   | 2.477 | 0.577   | 2.592 |

n は標本数, k' は定数項を除く説明変数の数とする。

数値例： 今までと同じ数値例で， $DW$  比を計算する。

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

$$\begin{aligned}
 DW &= \frac{\sum_{i=2}^n (\hat{u}_i - \hat{u}_{i-1})^2}{\sum_{i=1}^n \hat{u}_i^2} \\
 &= \frac{(0.0 - (-0.2))^2 + (-0.2 - (-1.6))^2 + (-1.6 - 1.1)^2 + (1.1 - 0.7)^2}{0.0^2 + (-0.2)^2 + (-1.6)^2 + 1.1^2 + 0.7^2} \\
 &= \frac{0.04 + 1.96 + 7.29 + 0.16}{0.00 + 0.04 + 2.56 + 1.21 + 0.49} = \frac{9.45}{4.3} = 2.198
 \end{aligned}$$

推定結果の表記方法（数値例）： 回帰モデル：

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

の推定の結果， $\hat{\alpha} = 0.5$ ， $\hat{\beta} = 0.7$ ， $s_{\hat{\alpha}} = \sqrt{1.5766667} = 1.25565$ ， $s_{\hat{\beta}} = \sqrt{0.1433333} = 0.3786$ ， $\frac{\hat{\alpha}}{s_{\hat{\alpha}}} = 0.398$ ， $\frac{\hat{\beta}}{s_{\hat{\beta}}} = 1.849$ ， $s^2 = 1.433333$ （すなわち， $s = 1.197$ ）， $R^2 = 0.5326$ ， $\bar{R}^2 = 0.3768$ ， $DW = 2.198$ を得た。

これらをまとめて，

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

$$R^2 = 0.5326, \quad \bar{R}^2 = 0.3768, \quad s = 1.197, \quad DW = 2.198$$

ただし，係数の推定値の下の括弧内は  $t$  値を表すものとする。

または，

$$Y_i = \begin{array}{c} \mathbf{0.5} \\ \mathbf{(1.256)} \end{array} + \begin{array}{c} \mathbf{0.7} \\ \mathbf{(0.379)} \end{array} X_i$$

$$R^2 = 0.5326, \quad \bar{R}^2 = 0.3768, \quad s = 1.197, \quad DW = 2.198$$

ただし，係数の推定値の下の括弧内は標準誤差を表すものとする。

のように書く。

### 7.1.3 系列相関は何が問題か?: 最小二乗推定量の分散について

簡単化のために, 下記のように単回帰を考える。

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

仮定:  $\mathbf{E}(u_i) = 0$

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

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

系列相関を考慮せずに、今までで扱ってきた通常の最小二乗推定量は、

$$\hat{\beta} = \frac{\sum_i (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_i (X_i - \bar{X})^2} = \sum_i \omega_i Y_i = \beta + \sum_i \omega_i u_i$$

である。ただし、 $\omega_i = \frac{X_i - \bar{X}}{\sum_j (X_j - \bar{X})^2}$

$\hat{\beta}$  の期待値  $\mathbf{E}(\hat{\beta})$  は、

$$\mathbf{E}(\hat{\beta}) = \mathbf{E}(\beta + \sum_i \omega_i u_i) = \beta + \sum_i \omega_i \mathbf{E}(u_i) = \beta$$

となるので、誤差項  $u_1, u_2, \dots, u_n$  に系列相関があっても、 $\hat{\beta}$  は不偏推定量となる。

$\hat{\beta}$  の分散  $\mathbf{V}(\hat{\beta})$  は ,

$$\begin{aligned}\mathbf{V}(\hat{\beta}) &= \mathbf{E}((\hat{\beta} - \beta)^2) = \mathbf{E}\left(\left(\sum_i \omega_i u_i\right)^2\right) = \mathbf{E}\left(\left(\sum_i \omega_i u_i\right)\left(\sum_j \omega_j u_j\right)\right) \\ &= \mathbf{E}\left(\sum_i \sum_j \omega_i \omega_j u_i u_j\right) = \sum_i \sum_j \omega_i \omega_j \mathbf{E}(u_i u_j) = \sum_i \sum_j \sigma_{ij} \omega_i \omega_j\end{aligned}$$

最初の等式は ,  $\mathbf{E}(\hat{\beta}) = \beta$  を利用して , 分散の定義である。2つ目の等式では , 添字一つを  $i$  から  $j$  に変更している。最後の等式では ,  $\mathbf{E}(u_i u_j) = \sigma_{ij}$  を利用している。

$i = j$  のとき ,  $\sigma_{ij} = \sigma^2$  とする。すなわち ,

$$\sum_i \sum_j \sigma_{ij} \omega_i \omega_j = \sigma^2 \sum_i \omega_i^2 + \sum_i \sum_{i \neq j} \omega_i \omega_j \sigma_{ij} \neq \sigma^2 \sum_i \omega_i^2$$

となる。

以上から， $\hat{\beta}$  の分布は，

$$\hat{\beta} \sim N(\beta, \sum_i \sum_j \sigma_{ij} \omega_i \omega_j)$$

となる。

さらに，標準化すると，

$$\frac{\hat{\beta} - \beta}{\sqrt{\sum_i \sum_j \sigma_{ij} \omega_i \omega_j}} \sim N(0, 1)$$

となる。

したがって、 $u_1, u_2, \dots, u_n$  に系列相関があるとき、共分散  $\sigma_{ij}$  の項を無視して、

$$\frac{\hat{\beta} - \beta}{\sigma \sqrt{\sum_i \omega_i^2}} \sim N(0, 1) \quad \Rightarrow \quad \text{これは間違い!!}$$

とはならないことに注意せよ（ただし、 $\sigma^2 = \sigma_{ii}$  としている）。

同様に、 $\sigma^2$  を  $s^2$  に置き換えて、

$$\frac{\hat{\beta} - \beta}{s \sqrt{\sum_i \omega_i^2}} \sim t(n-2) \quad \Rightarrow \quad \text{これも間違い!!}$$

ともならない。

ただし、 $s^2$  は今まで通り、 $s^2 = \frac{1}{n-2} \sum_i \hat{u}_i^2 = \frac{1}{n-2} \sum_i (Y_i - \hat{\alpha} - \hat{\beta}X_i)^2$  としている。

$u_1, u_2, \dots, u_n$  に系列相関があるとき，通常 of 最小二乗推定量  $\hat{\beta}$  の分散の推定量は，

$$\sum_i \sum_j s_{ij} \omega_i \omega_j = s^2 \sum_i \omega_i^2 + \sum_{i \neq j} \sum_j \omega_i \omega_j s_{ij} \neq s^2 \sum_i \omega_i^2$$

とならなければならない。

$s^2, s_{ij}$  は  $\sigma^2, \sigma_{ij}$  の推定量とする。

しかし，計量ソフトは  $s^2 \sum_i \omega_i^2$  と計算する。

さらに， $u_1, u_2, \dots, u_n$  が互いに独立のときは  $E(s^2) = \sigma^2$  となるが，系列相関があるときは  $E(s^2) \neq \sigma^2$  となる。

まとめると， $u_1, u_2, \dots, u_n$  に系列相関があるとき，通常 of 計量ソフトで得られる最小二乗法による推定結果について，

- 係数推定量  $\hat{\beta}_j$  は  $\beta_j$  の不偏推定量，一致推定量となっている（正しく推定される）。
- 回帰式の標準誤差  $s$  は正しく推定されない。
- 係数推定値  $\hat{\beta}_j$  の標準誤差  $s\sqrt{a_{jj}}$  も正しく推定されない。
- 係数推定値  $\hat{\beta}_j$  の  $t$  値も正しくない。

となる。

このように， $u_1, u_2, \dots, u_n$  に系列相関があるとき，最小二乗法による推定結果を用いると，間違っただ信賴区間，間違っただ検定結果が得られる。