

1 系列相関：DW について

1.1 回帰モデルの仮定

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）

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， $Cov(u_i, u_j) = 0$ とする。（ $Cov(u_i, u_j) = E(u_i u_j) = 0$ 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。

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。

1.2 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_T 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_t と u_{t-1} との間の相関の有無を検定するために考案された。
 \Rightarrow 時系列データのときのみ有効

u_1, u_2, \dots, u_T の系列について，それぞれの符号が， $++ + \dots + + \dots + +$ 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_T 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_T 負の系列相関があると言う。

特徴： u_1, u_2, \dots, u_t から u_{t+1} の符号が予想できる。 \Rightarrow 「 u_1, u_2, \dots, u_T 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

$$y_t = \alpha + \beta x_t + u_t, \\ u_t = \rho u_{t-1} + \epsilon_t,$$

のときに， $H_0: \rho = 0$ ， $H_1: \rho \neq 0$ の検定である。ただし， $\epsilon_1, \epsilon_2, \dots, \epsilon_T$ 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。

$$DW = \frac{\sum_{t=2}^T (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^T \hat{u}_t^2}$$

DW は近似的に，次のように表される。

$$DW = \frac{\sum_{t=2}^T (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^T \hat{u}_t^2} \\ = \frac{\sum_{t=2}^T \hat{u}_t^2 - 2 \sum_{t=2}^T \hat{u}_t \hat{u}_{t-1} + \sum_{t=2}^T \hat{u}_{t-1}^2}{\sum_{t=1}^T \hat{u}_t^2} \\ = \frac{2 \sum_{t=1}^T \hat{u}_t^2 - (\hat{u}_1^2 + \hat{u}_T^2)}{\sum_{t=1}^T \hat{u}_t^2} - 2 \frac{\sum_{t=2}^T \hat{u}_t \hat{u}_{t-1}}{\sum_{t=1}^T \hat{u}_t^2} \\ \approx 2(1 - \hat{\rho}),$$

以下の2つの近似が用いられる。

$$\frac{\hat{u}_1^2 + \hat{u}_T^2}{\sum_{t=1}^T \hat{u}_t^2} \approx 0, \\ \frac{\sum_{t=2}^T \hat{u}_t \hat{u}_{t-1}}{\sum_{t=1}^T \hat{u}_t^2} = \frac{\sum_{t=2}^T \hat{u}_t \hat{u}_{t-1}}{\sum_{t=2}^T \hat{u}_{t-1}^2 + \hat{u}_T^2} \\ \approx \frac{\sum_{t=2}^T \hat{u}_t \hat{u}_{t-1}}{\sum_{t=2}^T \hat{u}_{t-1}^2} \\ = \hat{\rho},$$

すなわち, $\hat{\rho}$ は \hat{u}_t と \hat{u}_{t-1} の回帰係数である。 $u_t = \rho u_{t-1} + \epsilon_t$ において, u_t, u_{t-1} の代わりに \hat{u}_t, \hat{u}_{t-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, $DW \approx 0$)。
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 T 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	$\sum Y_i$ 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				

$$DW = \frac{\sum_{t=2}^T (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^T \hat{u}_t^2}$$

$$= \frac{(-0.8 - 0.9)^2 + (0.9 - 0.6)^2 + (0.6 - (-0.7))^2}{(-0.8)^2 + 0.9^2 + 0.6^2 + (-0.7)^2}$$

$$= \frac{4.67}{2.30} = 2.03$$

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,

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

$$R^2 = 0.786, \quad \bar{R}^2 = 0.679,$$

$$s = 1.07, \quad DW = 2.03,$$

ただし, 係数の推定値の下の括弧内は t 値を表すものとする。

または,

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

$$R^2 = 0.786, \quad \bar{R}^2 = 0.679,$$

$$s = 1.07, \quad DW = 2.03,$$

ただし, 係数の推定値の下の括弧内は標準誤差を表すものとする。

のように書く。 $s = \sqrt{1.15} = 1.07$ に注意。

図 4: 正の系列相関

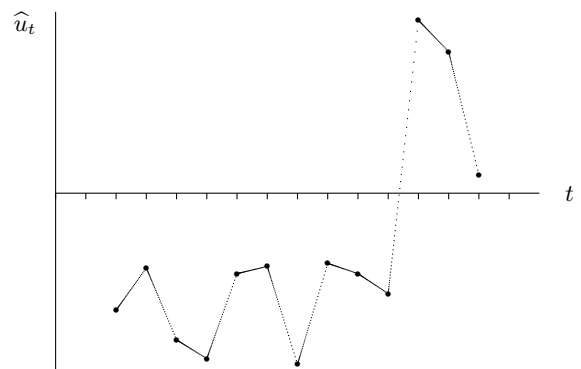
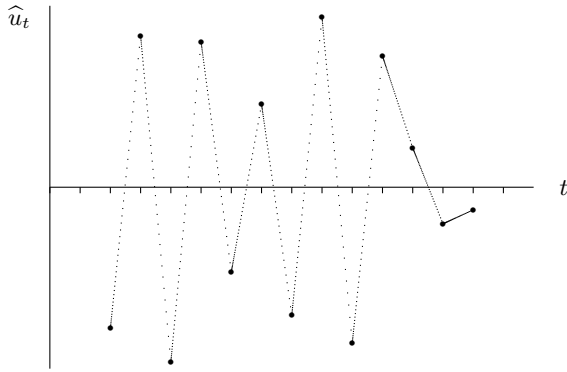


図 5: 負の系列相関



1.3 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) $k' = 1$

T	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) $k' = 2$

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

(3) $k' = 3$

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

(4) $k' = 4$

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

(5) $k' = 5$

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

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

表 2: ダービン・ワトソン統計量の 5% 点の上限と下限

T	k' = 1		k' = 2		k' = 3		k' = 4		k' = 5	
	dl	du	dl	du	dl	du	dl	du	dl	du
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

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

1. ダービン・ワトソン比がゼロに近い場合, $\rho = 1$ と近似して,

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

として推定する。ただし, $Y_i^* = Y_i - Y_{i-1}$, $X_{1i}^* = X_{1i} - X_{1,i-1}$ とする。

2. 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。

3. 収束計算によって求める。

(i) まず,

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

に最小二乗法を適用し, $\hat{u}_1, \hat{u}_2, \dots, \hat{u}_n$ を求める。

(ii) 次に,

$$\hat{u}_i = \rho \hat{u}_{i-1} + \epsilon_i,$$

に最小二乗法を適用し, $\hat{\rho}$ を求める。

(iii) データを変換する。すなわち,

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

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

$$X_{2i}^* = (X_{2i} - \hat{\rho} X_{2,i-1}),$$

...

$$X_{ki}^* = (X_{ki} - \hat{\rho} X_{k,i-1})$$

を新たな変数として,

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

を計算して, $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k$ を求める。

(iv) さらに,

$$\hat{u}_i = Y_i - \hat{\beta}_1 X_{1i} - \hat{\beta}_2 X_{2i} - \cdots - \hat{\beta}_k X_{ki},$$

から \hat{u}_i を再計算して,

$$\hat{u}_i = \rho \hat{u}_{i-1} + \epsilon_i,$$

から, 最小二乗法により $\hat{\rho}$ を求める。

(v) (iii) と (iv) を収束するまで繰り返す。

(iv) で得られた $\hat{\rho}$ の有意性から, $H_0: \rho = 0$ の検定を行う方法もある。

1.4 遅れのある変数 (再考)

習慣的効果を考慮に入れたモデル:

$$y_t = \alpha + \beta x_t + \gamma y_{t-1} + u_t,$$

ラグ付き内生変数が説明変数に用いられる。

x_t の y_t 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, $y_t = y_{t-1}$ となるときの, x_t から y_t への影響を示す効果である。

$$y_t = \alpha + \beta x_t + \gamma y_t + u_t,$$

として, y_t について解くと,

$$y_t = \frac{\alpha}{1-\gamma} + \frac{\beta}{1-\gamma} + \frac{1}{1-\gamma} u_t,$$

となり, $\frac{\beta}{1-\gamma}$ が x_t の y_t 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_t と x_t とは, 経済理論的に考えると, 相関が高いはず。 y_t と y_{t-1} は相関が高い。当然, y_{t-1} と x_t 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 (詳細略)。実際には誤差項に系列相関があるにもかかわらず, 標本数 (データ数) が増えるにつれて, DW は 2 に近づいてしまう。すなわち, 系列相関なしと判定されてしまう。

⇒ 代わりに, h 統計量を使う。

h 統計量は次のように求められる。

$$h = \hat{\rho} \sqrt{\frac{T}{1 - T s_{\hat{\rho}}^2}},$$

T は標本数 (データ数), $s_{\hat{\rho}}$ は $\hat{\rho}$ の標準誤差, $\hat{\rho}$ は $\hat{\rho} = \frac{\sum_{t=2}^T \hat{u}_t \hat{u}_{t-1}}{\sum_{t=2}^T \hat{u}_{t-1}^2}$ で, \hat{u}_t はもとのモデルの最小二乗法による残差である。

帰無仮説は, $u_t = \rho u_{t-1} + \epsilon_t$ のモデルにおいて, $H_0: \rho = 0$ であり, 対立仮説は $H_1: \rho \neq 0$ である。帰無仮説 $H_0: \rho = 0$ のもとで, 標本数 T が増加するにつれて, h は標準正規分布に近づくことが知られている。よって,

(a) $|h| > z_{\alpha/2}$ のとき, 有意水準 100 % で H_0 を棄却する。

(b) $|h| < z_{\alpha/2}$ のとき, 有意水準 100 % で H_0 を採択する。

ただし, $z_{\alpha/2}$ は標準正規分布表から得られた $100 \times \frac{\alpha}{2}$ % 点の値である。