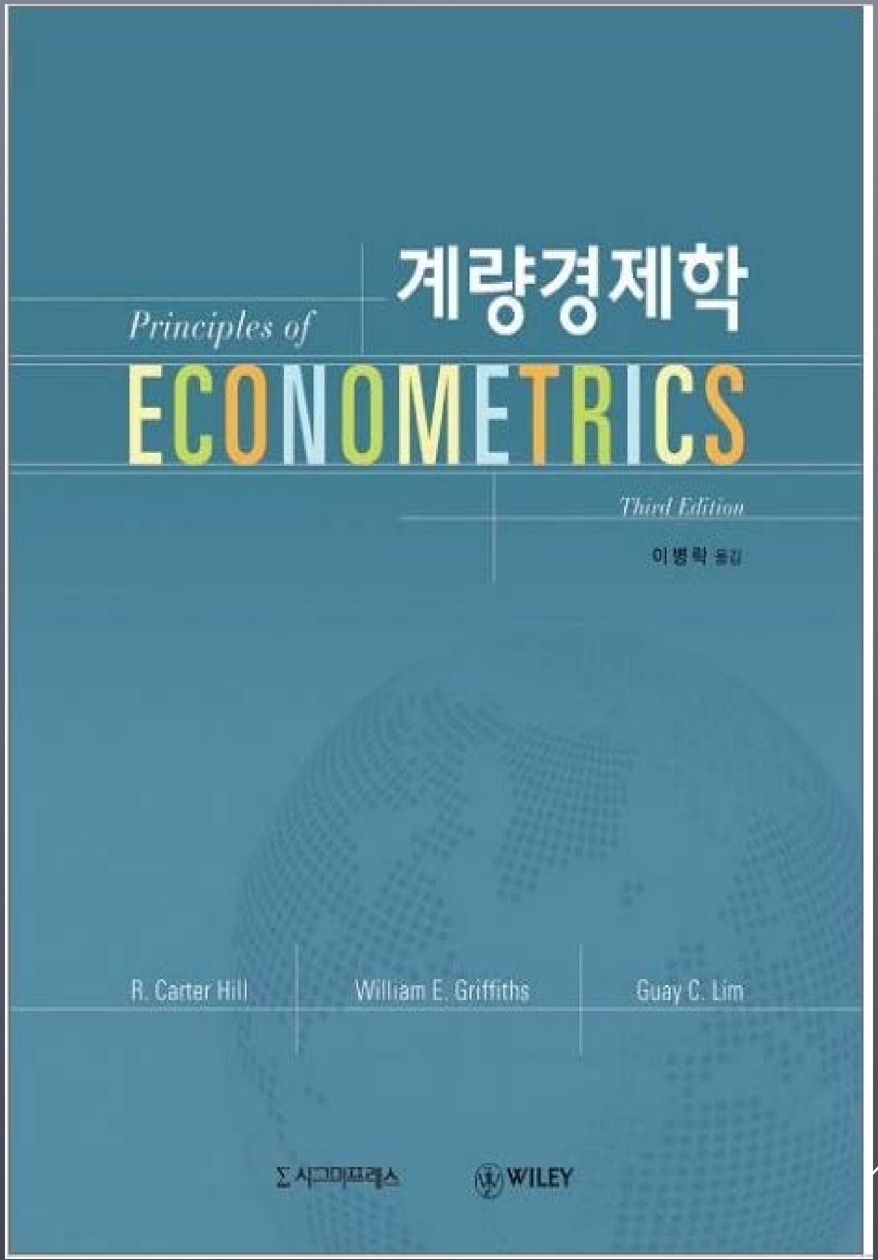
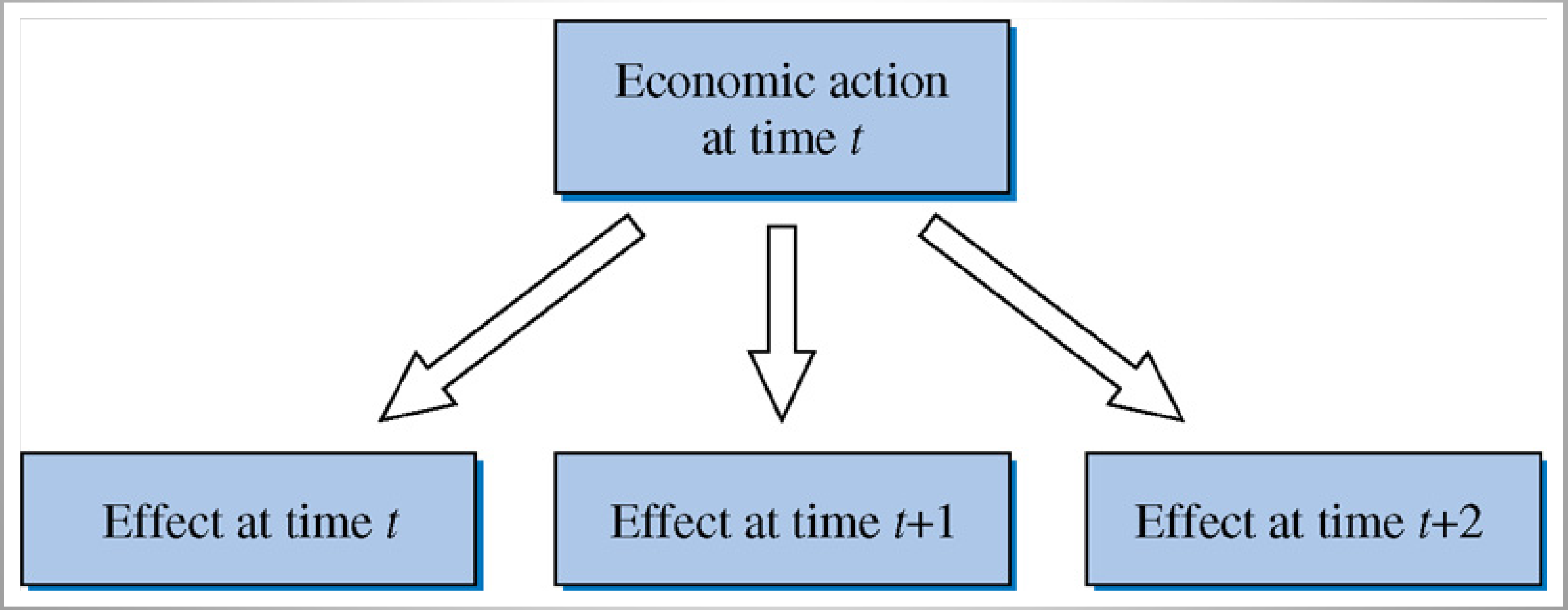


제9장
동태모형, 자기상관, 예측

9.1 서론

- 경제적 인과관계의 시차효과
- x_t 변화는 여러 기간에 걸쳐 종속변수에 영향을 미칠 수 있음

$$x_t \Rightarrow y_t, y_{t+1}, y_{t+2}, \cdots$$



- **경제변수의 동태적 관계를 모형화할 수 있는 세 가지 방법**

- **시차설명변수의 도입**

$x_t \Rightarrow y_t, y_{t+1}, y_{t+2}, \cdots$ 이 관계를 달리 표현하면, $y_t = f(x_t, x_{t-1}, x_{t-2}, \cdots)$

(예) 현재 이자율의 변화는 현재 실업률과 장래의 실업률에 영향 미침

⇒ 현재 실업률은 현재 이자율뿐만 아니라 과거 이자율에도 의존함

- **시차종속변수의 도입**

$$y_t = f(y_{t-1}, x_t)$$

(예) 높은 실업률이 지속되는 경향 있음 (경제의 추세적 변화, 관성)

- **오차항 구조의 동태화**

$$y_t = f(x_t) + e_t \qquad e_t = f(e_{t-1})$$

(예) 테러에 의한 유전 파괴는 현재 유가뿐만 아니라 장래 유가에도 영향

<9장에서 배우는 내용>

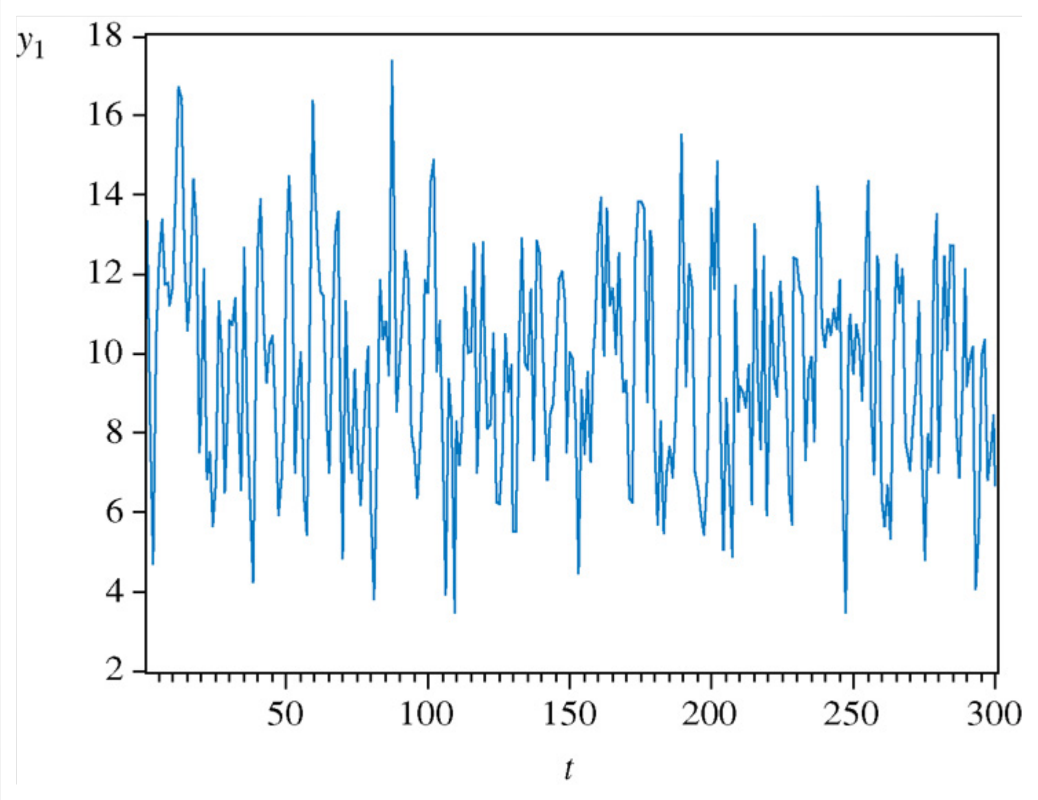
- 경제변수의 동태적 관계를 모형화할 수 있는 세 가지 방법
- 세 가지 방법이 별개의 것이 아님
  - 유사한 효과가 있음
- 동태 모형을 예측을 하는데 활용하는 방법
  - 모형을 동태화하는 주된 목적은 예측

- 이 장에서는 모형에 포함된 변수들이 안정적이라고 가정함

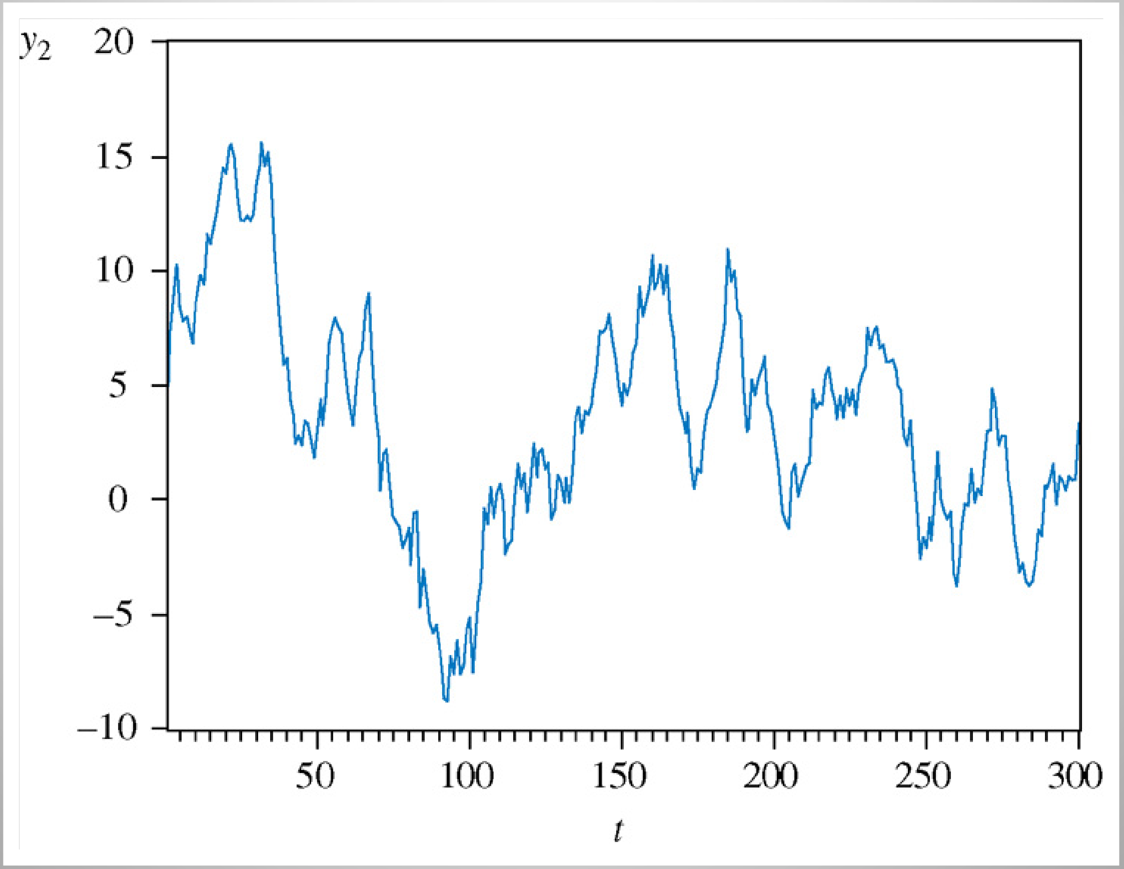
 - 변수가 불안정한 경우는 12장에서 공부함

- **안정적 시계열(stationary variable)**

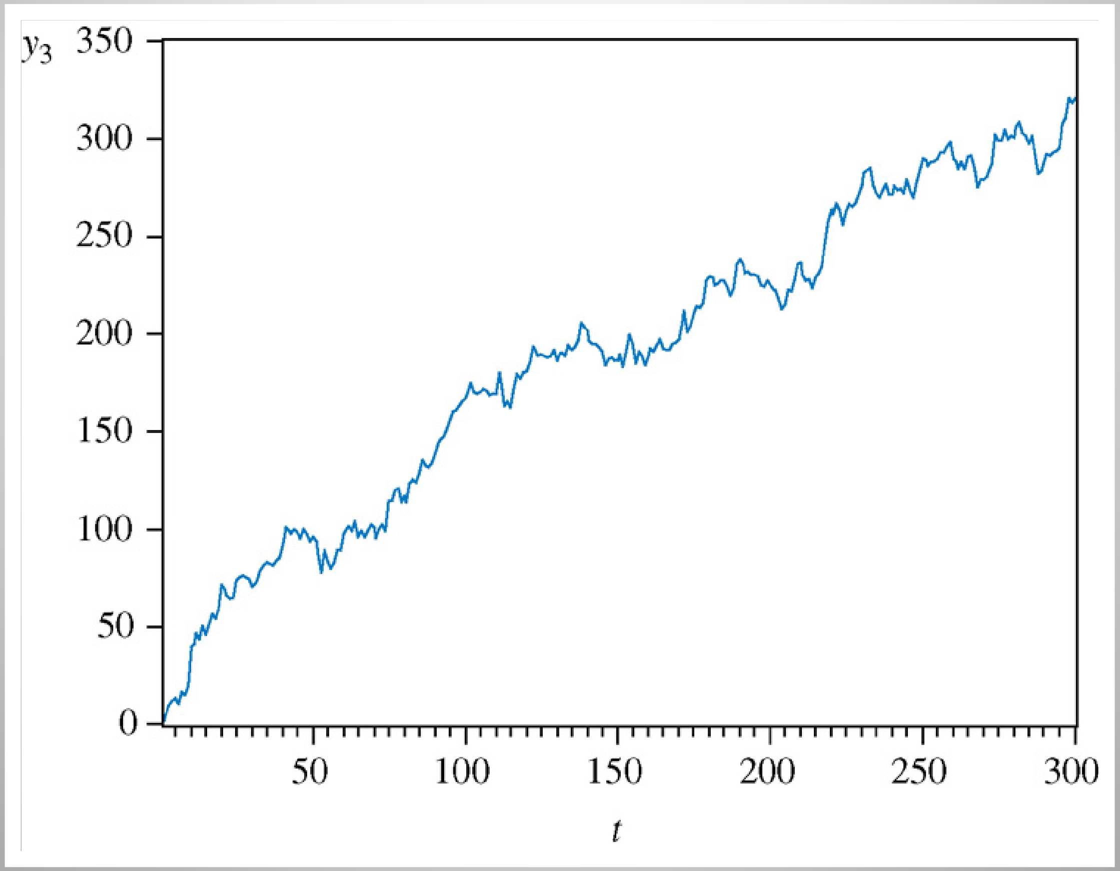
 - 시간이 흘러도 평균과 분산이 일정한 값을 나타내는 시계열



- **불안정한 시계열(a)**
  - Random walk 하는 시계열
  - Nonstationary Variable that is ‘Slow Turning’ or ‘Wandering’



- **불안정한 시계열(b)**
  - Random walk with a drift
  - 'Trends'를 가지는 시계열



9.2 오차항의 시차: 자기상관

- 예상치 못한 충격이 경제에 미치는 영향이 일정 기간 지속되는 경우 있음
- 종속변수는 현재 충격이 미치는 영향뿐만 아니라 이전 충격의 이월된 영향을 받게 됨
- 금기의 충격은 이전 기의 충격과 관련됨
- 오차항에 자기상관(autocorrelation)이 존재한다고 말함

- **자기상관과 관련된 주제**

- 횡단면자료의 경우 보통 표본을 random sampling에 의해 수집하므로 자기상관의 문제가 없음
- 시계열자료의 경우 표본이 시간의 흐름에 따라 수집되므로 연속되는 오차들이 서로 상관될 가능성이 항상 있음

✓ 자기상관이 존재할 때 OLS 추정량의 문제는 무엇인가?

✓ 자기상관이 존재하는지를 어떻게 탐지할 수 있는가?

✓ 자기상관이 존재하는 경우 어떻게 조치를 취해야 하는가?

Assumptions of the Multiple Regression Model

MR1. $y_t = \beta_1 + \beta_2 x_{t2} + \cdots + \beta_K x_{tK} + e_t,\ t = 1, \ldots, T$

MR2. $E(y_t) = \beta_1 + \beta_2 x_{t2} + \cdots + \beta_K x_{tK} \Leftrightarrow E(e_t) = 0.$

MR3. $\text{var}(y_t) = \text{var}(e_t) = \sigma^2.$

MR4. $\boxed{\text{cov}(y_t, y_s) = \text{cov}(e_t, e_s) = 0}$

MR5. The values of x_{tk} are not random and are not exact linear functions of the other explanatory variables.

MR6. $y_t \sim N\left[(\beta_1 + \beta_2 x_{t2} + \ldots + \beta_K x_{tK}), \sigma^2\right] \Leftrightarrow e_t \sim N(0, \sigma^2),$

- If $Cov(y_t, y_s) = Cov(e_t, e_s) \neq 0$

$\Rightarrow$ 자기상관(autocorrelation)이 존재

9.2.1 사탕수수 가격 변화에 대한 경작면적의 반응

- 사탕수수 경작면적의 가격탄력성을 알면,
  방글라데시 정부는 사탕수수 가격정책 수립에 활용할 수 있음
- 회귀모형 $\ln(A_t) = \beta_1 + \beta_2 \ln(P_t) + e_t$ A : 경작면적, P : 가격

$$y_t = \beta_1 + \beta_2 x_t + e_t \qquad y_t = \ln(A_t) \qquad x_t = \ln(P_t)$$

- 다중회귀모형의 기본가정이 충족된다면 OLS로 추정 가능
- 시계열자료 이용하여 OLS로 추정한 결과

$$\hat{y}_t = 3.893 + 0.776\, x_t$$
$$\text{(se)} \quad (0.061) \quad (0.277)$$

- OLS

Dependent Variable: LNA
Method: Least Squares
Date: 04/24/11 Time: 22:14
Sample: 1 34
Included observations: 34
LNA=C(1)+C(2)*LNP

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	3.893256	0.061345	63.46486	0.0000
C(2)	0.776119	0.277467	2.797154	0.0087

R-squared	0.196466	Mean dependent var	3.980680
Adjusted R-squared	0.171355	S.D. dependent var	0.338123
S.E. of regression	0.307793	Akaike info criterion	0.538245
Sum squared resid	3.031571	Schwarz criterion	0.628031
Log likelihood	-7.150159	Hannan-Quinn criter.	0.568864
F-statistic	7.824072	Durbin-Watson stat	1.168987
Prob(F-statistic)	0.008653		

<OLS 잔차 계산 결과>

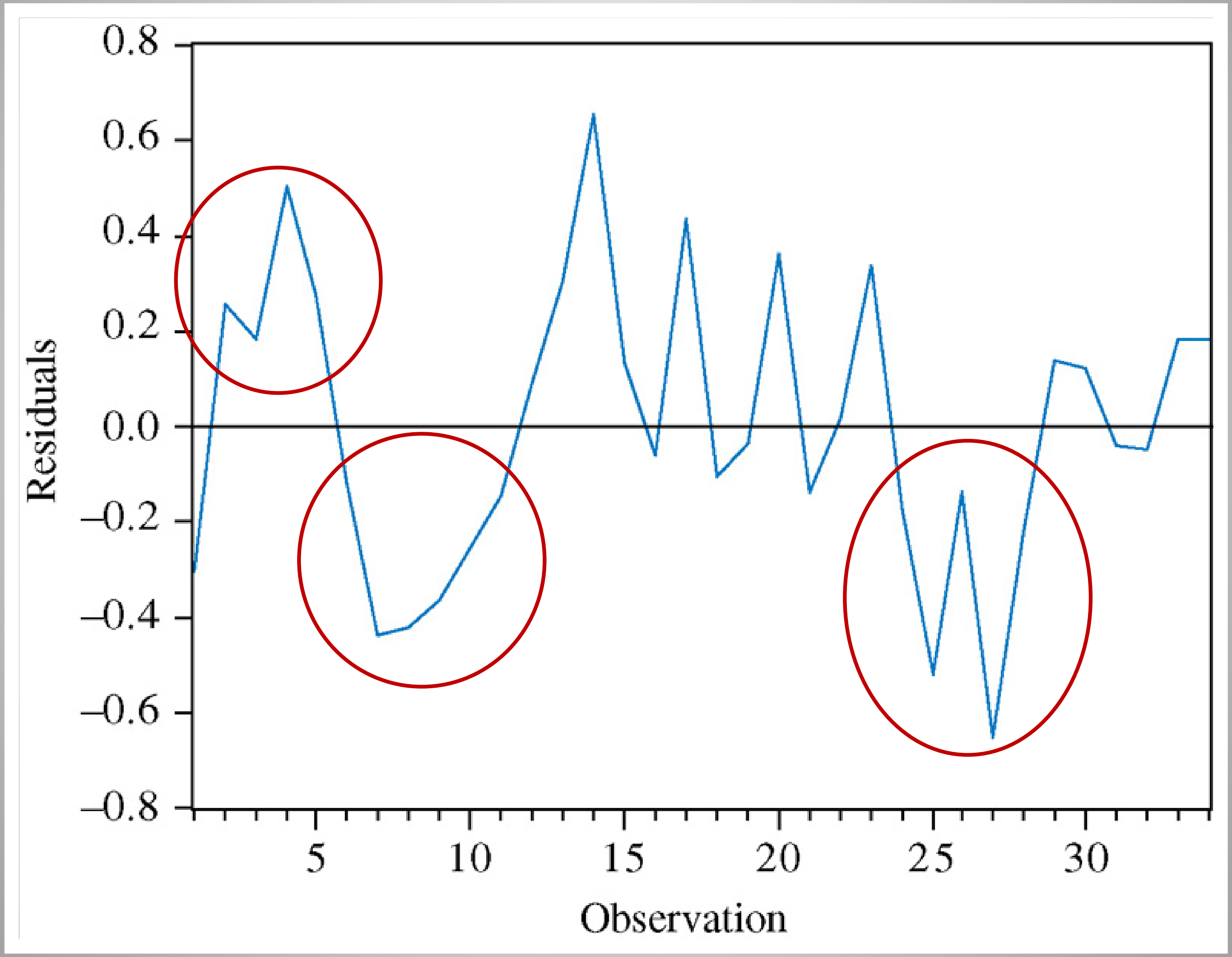
$$\hat{y}_t = 3.893 + 0.776\,x_t$$

$$\hat{e}_t = y_t - \hat{y}_t$$

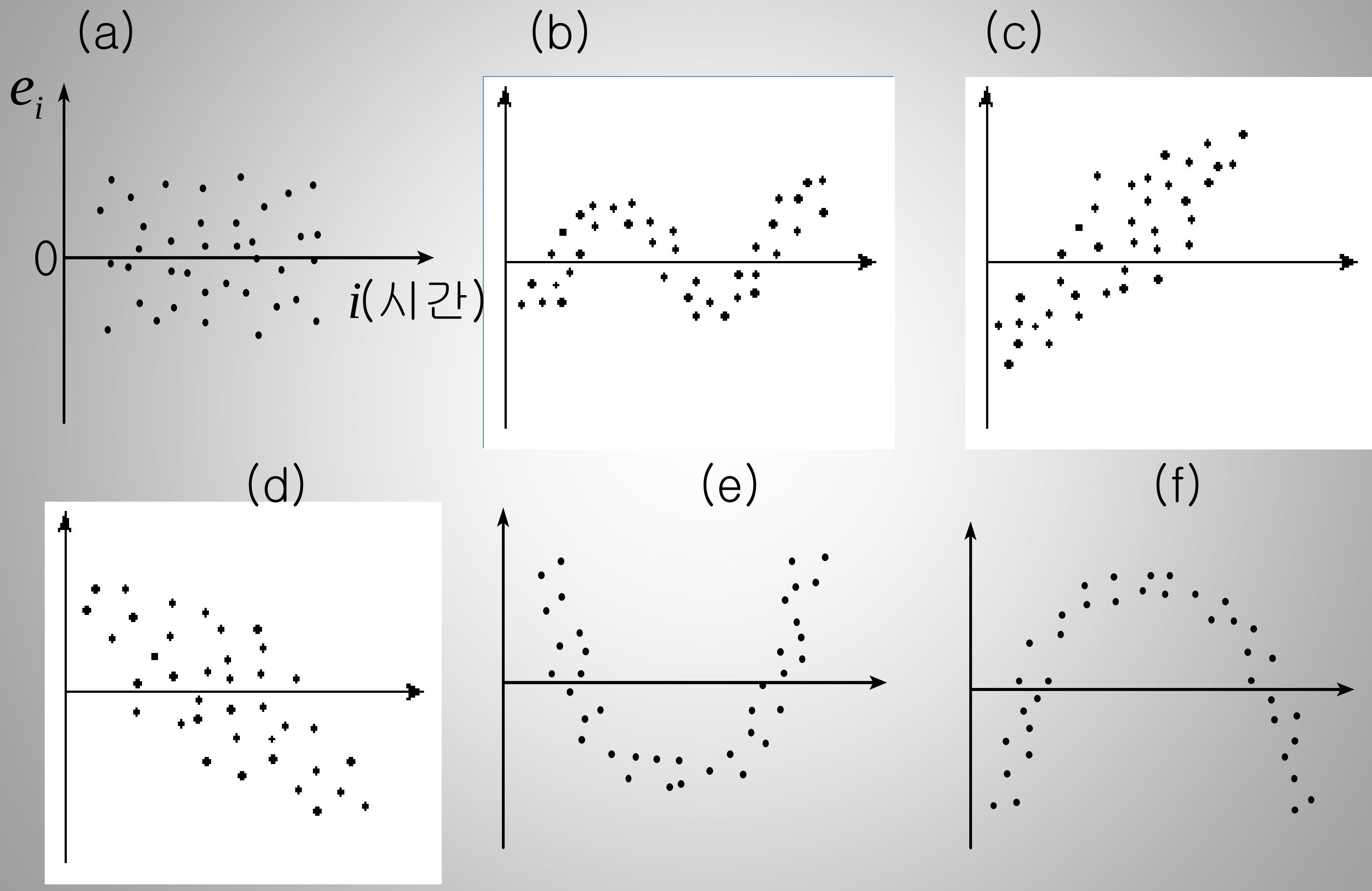
Table 9.1 Least Squares Residuals for the Sugarcane Example

Time	$\hat{e}_t$	Time	$\hat{e}_t$	Time	$\hat{e}_t$	Time	$\hat{e}_t$
1	−0.303	10	−0.254	19	−0.036	27	−0.651
2	0.254	11	−0.145	20	0.361	28	−0.218
3	0.182	12	0.091	21	−0.138	29	0.137
4	0.503	13	0.304	22	0.017	30	0.121
5	0.275	14	0.656	23	0.336	31	−0.040
6	−0.115	15	0.134	24	−0.175	32	−0.048
7	−0.437	16	−0.059	25	−0.517	33	0.183
8	−0.423	17	0.435	26	−0.137	34	0.184
9	−0.367	18	−0.106				

<OLS 잔차의 그림: Residuals plotted against time>



- **잔차 그래프**



✓ **현재의 오차항이 이전의 오차에 관련되는 것을 모형화할 수 있는 방법은 무엇일까?**

- 가장 간단한 방법은 다음의 AR(1) 모형으로 나타내는 것임

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

(a) ρe_{t-1} 부분은 경제활동의 관성으로 인해 이전 기간의 오차에서 이월된 부분

(b) v_t 부분은 새로운 (금기의) 충격이 만들어 낸 무작위 오차

- **현재의 오차항이 이전의 오차와 관련되는 것을 모형화하는 방법**
- AR(1) 모형

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

- ρ 는 충격의 영향이 얼마나 빨리 소멸되는지를 나타내는 모수
- ρ 의 크기가 클수록 충격의 영향이 다음 기간으로 많이 이월됨

- 종속변수에 대한 충격의 영향이 시간의 흐름에 따라 서서히 확산됨

- AR(p) 모형

$$e_t = \phi_1 e_{t-1} + \phi_2 e_{t-2} + \cdots + \phi_p e_{t-p} + v_t$$

- MA(q) 모형

$$e_t = v_t - \theta_1 v_{t-1} - \theta_2 v_{t-2} - \cdots - \theta_q v_{t-q}$$

9.2.2 1차 자기회귀오차, AR(1)

- 오차항이 AR(1) 모형을 따른다고 하자 (이를 AR(1) 과정이라고 함)

$$y_t = \beta_1 + \beta_2 x_t + e_t , \qquad e_t = \rho e_{t-1} + v_t$$

- v_t는 무작위 오차이고, *iid* 확률변수임 (independently and identically distributed)

$$E(v_t) = 0, \quad \mathrm{var}(v_t) = \sigma_v^2 , \quad \mathrm{cov}(v_t , v_s) = 0 \quad \text{for } t \neq s$$

(a) ρe_{t-1} 부분은 경제활동의 관성으로 인해
이전 기간의 오차에서 이월된 부분

(b) v_t 부분은 새로운 (금기의) 충격이 만들어 낸 오차

- **AR(1) 오차의 특성** $e_t = \rho e_{t-1} + v_t$

- AR(1) 모형 오차의 특성을 결정하는 요소는 ρ 임
- ρ 에 대해 다음과 같이 가정함

$$-1 < \rho < 1$$

- 만약 ρ 의 값이 이 범위를 넘어서면,
e_t 는 시간이 경과함에 따라 점점 더 커져
궁극적으로는 무한대가 됨 (상식 밖의 결과)

- **AR(1) 오차의 특성** $e_t = \rho e_{t-1} + v_t \qquad -1 < \rho < 1$

- AR(1) 모형의 오차는 평균이 0임 $E(e_t) = 0$
- AR(1) 모형의 오차는 동분산임 $\operatorname{var}(e_t) = \sigma_e^2 = \dfrac{\sigma_v^2}{1-\rho^2}$
- AR(1) 모형의 오차들 사이의 독립성은? (공분산을 알아보자)

$\operatorname{cov}\left(e_t, e_{t-k}\right) = \sigma_e^2 \rho^k, \quad k > 0$ (k 는 오차간의 시차, AR(1) 경우에는 $k = 1$)

$$\operatorname{corr}(e_t, e_{t-k}) = \frac{\operatorname{cov}(e_t, e_{t-k})}{\sqrt{\operatorname{var}(e_t)\operatorname{var}\left(e_{t-k}\right)}} = \frac{\operatorname{cov}(e_t, e_{t-k})}{\operatorname{var}(e_t)} = \frac{\sigma_e^2 \rho^k}{\sigma_e^2} = \rho^k$$

- 공분산은 ρ 와 k 에 의존

- $\rho = 0$ 이면 공분산은 $Cov(e_t, e_{t-k}) = 0$ (오차들은 독립적)
- $\rho \neq 0$ 이면 공분산은 k 가 증가함에 따라 급격히 감소

$$\rho,\ \rho^2,\ \rho^3,\ \rho^4,\ \cdots$$

9.3 AR(1) 오차의 추정

9.3.1 자기상관이 있는 경우의 OLS 추정의 문제점

- 회귀모형 오차에 자기상관이 존재하지만,
  이를 무시하거나 알지 못하여 OLS 이용하여 추정한 경우에는
  어떤 문제가 발생하는가?

- 오차항이 AR(1) 과정을 따를 경우 $\operatorname{cov}\left(e_t, e_{t-k}\right) = \sigma_e^2 \rho^k, \quad k > 0$

- 이것은 OLS 추정의 기본가정인 독립성 가정에 위배됨

$$\operatorname{cov}\left(e_t, e_s\right) = 0, \quad t \neq s$$

▪ 자기상관이 있는 경우의 OLS 추정의 문제점

- OLS 추정량은 선형 불편 추정량이지만,
  최소분산을 가지지는 않음 (BLUE 아님)
  ⇒ 더 분산이 작은 추정량을 찾을 수 있을 것임 (GLS 추정량)

- 추정치의 표준오차는 부정확하므로(클 수도 작을 수도 있음),
  이를 이용한 가설검정 및 신뢰구간은 오류를 범할 수 있음
  ⇒ OLS 추정량에 대한 정확한 표준오차 구하는 방법 있음
  (HAC 표준오차, Newey-West 표준오차)

$$
\begin{aligned}
\hat{y}_t &= 3.893 + 0.776\,x_t \\
&\quad\ (0.061)\quad (0.277) \qquad \text{'incorrect' se (OLS)} \\
&\quad\ (0.062)\quad (0.378) \qquad \text{'correct' se}
\end{aligned}
$$

- OLS + Newey-West 표준오차

Dependent Variable: LNA
Method: Least Squares
Date: 05/19/11 Time: 15:02
Sample: 1 34
Included observations: 34
Newey-West HAC Standard Errors & Covariance (lag truncation=3)
LNA=C(1)+C(2)*LNP

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	3.893256	0.062444	62.34761	0.0000
C(2)	0.776119	0.378207	2.052102	0.0484

R-squared	0.196466	Mean dependent var	3.980680
Adjusted R-squared	0.171355	S.D. dependent var	0.338123
S.E. of regression	0.307793	Akaike info criterion	0.538245
Sum squared resid	3.031571	Schwarz criterion	0.628031
Log likelihood	-7.150159	Hannan-Quinn criter.	0.568864
F-statistic	7.824072	Durbin-Watson stat	1.168987
Prob(F-statistic)	0.008653		

9.3.2 AR(1) 오차모형의 비선형 최소제곱 추정법 (NLS)

- AR(1) 오차모형의 경우,
- OLS 추정량에 대한 정확한 표준오차를 구하는 것이 가능하기는 함
- 그럴지만 더 나은 추정방법이 있음

▪ 기본 회귀모형의 변형

$$y_t = \beta_1 + \beta_2 x_t + e_t \qquad e_t = \rho e_{t-1} + v_t \qquad y_t = \beta_1 + \beta_2 x_t + \rho e_{t-1} + v_t$$

$$e_{t-1} = y_{t-1} - \beta_1 - \beta_2 x_{t-1}$$

$$\rho e_{t-1} = \rho y_{t-1} - \rho\beta_1 - \rho\beta_2 x_{t-1} \qquad \boxed{y_t = \beta_1 + \beta_2 x_t + \rho y_{t-1} - \rho\beta_1 - \rho\beta_2 x_{t-1} + v_t}$$

- v_t는 *iid* random error ⇒ NLS 적용하여 추정하면 됨
- 사탕수수 사례

$$\widehat{\ln(A_t)} = 3.899 + 0.888\ln(P_t) \qquad e_t = 0.422 e_{t-1} + v_t$$

(se) (0.092) (0.259) (0.166)

- AR(1) 추정 (NLS 이용)

$$y_t = \beta_1 + \beta_2 x_t + \rho y_{t-1} - \rho\beta_1 - \rho\beta_2 x_{t-1} + v_t$$

$$\widehat{\ln(A_t)} = 3.899 + 0.888\ln(P_t) \qquad e_t = 0.422 e_{t-1} + v_t$$
$$(\text{se}) \quad (0.092) \; (0.259) \qquad\qquad (0.166)$$

Dependent Variable: LNA
Method: Least Squares
Date: 05/19/11 Time: 15:17
Sample (adjusted): 2 34
Included observations: 33 after adjustments
Convergence achieved after 6 iterations
LNA=C(1)+C(2)*LNP+C(3)*LNA(-1)-C(3)*C(1)-C(3)*C(2)*LNP(-1)

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	3.898771	0.092166	42.30167	0.0000
C(2)	0.888371	0.259298	3.426056	0.0018
C(3)	0.422138	0.166047	2.542274	0.0164

R-squared	0.277777	Mean dependent var	3.999309
Adjusted R-squared	0.229629	S.D. dependent var	0.325164
S.E. of regression	0.285399	Akaike info criterion	0.416650
Sum squared resid	2.443575	Schwarz criterion	0.552696
Log likelihood	-3.874725	Hannan-Quinn criter.	0.462425
Durbin-Watson stat	1.820557		

9.3.2a AR(1) 오차모형의 GLS 추정법

- 기본 회귀모형의 변형

$$y_t = \beta_1 + \beta_2 x_t + e_t \;,\; e_t = \rho e_{t-1} + v_t \quad \Rightarrow \quad y_t = \beta_1 + \beta_2 x_t + \rho e_{t-1} + v_t$$

$$e_{t-1} = y_{t-1} - \beta_1 - \beta_2 x_{t-1}$$

$$\rho e_{t-1} = \rho y_{t-1} - \rho\beta_1 - \rho\beta_2 x_{t-1}$$

$$y_t = \beta_1 + \beta_2 x_t + \rho y_{t-1} - \rho\beta_1 - \rho\beta_2 x_{t-1} + v_t$$

$$\boxed{y_t - \rho y_{t-1} = \beta_1(1-\rho) + \beta_2(x_t - \rho x_{t-1}) + v_t}$$

$y_t^* = y_t - \rho y_{t-1}$ $\quad x_{t1}^* = 1-\rho$ $\quad x_{t2}^* = x_t - \rho x_{t-1}$ 으로 바꾸어 표시하면

$\boxed{y_t^* = x_{t1}^*\beta_1 + x_{t2}^*\beta_2 + v_t}$ 이 변형된 식의 오차항(v_t)는 동분산, 독립

⇒ 따라서 OLS 이용하면 β_1, β_2 에 대한 BLUE 추정량 구할 수 있음

✓ 그렇지만 자기상관계수 ρ 는 어떻게 알 수 있는가?

- **GLS를 이용한 추정절차 (Cochrane-Orcutt 추정법)**

• ρ 대신 그것의 추정치 $\hat{\rho}$ 를 이용하는 방법

(1) $y_t = \beta_1 + \beta_2 x_t + e_t$ 를 추정하여 $\hat{e}_t = y_t - b_1 - b_2 x_t$ 를 구함

(2) 이 잔차 자료를 이용하여 OLS 방법으로 $\hat{\rho}$ 를 추정함

$$\hat{\rho} = \frac{\sum_{t=2}^{T} \hat{e}_t \hat{e}_{t-1}}{\sum_{t=2}^{T} \hat{e}_{t-1}^2}$$

(3) $\hat{\rho}$ 를 이용하여 변형된 자료 만듦

$$y_t^* = y_t - \hat{\rho} y_{t-1} \qquad x_{t1}^* = 1 - \hat{\rho} \qquad x_{t2}^* = x_t - \hat{\rho} x_{t-1}$$

(4) $y_t^* = x_{t1}^* \beta_1 + x_{t2}^* \beta_2 + v_t$ 를 OLS로 추정

$\Rightarrow$ (1) 단계부터 다시 반복, 추정치가 수렴하면 종료

❖ 이 추정방법은 앞의 NLS 추정법과 사실상 동일함 (부록 9A)

- **AR(1) (Cochrane-Orcutt 추정법)**

Quick 클릭,
Estimate Equation 클릭,
lna c lnp ar(1) 입력

Dependent Variable: LNA
Method: Least Squares
Date: 04/24/11 Time: 22:34
Sample (adjusted): 2 34
Included observations: 33 after adjustments
Convergence achieved after 7 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.898771	0.092165	42.30197	0.0000
LNP	0.888370	0.259299	3.426048	0.0018
AR(1)	0.422140	0.166047	2.542284	0.0164

R-squared	0.277777	Mean dependent var	3.999309
Adjusted R-squared	0.229629	S.D. dependent var	0.325164
S.E. of regression	0.285399	Akaike info criterion	0.416650
Sum squared resid	2.443575	Schwarz criterion	0.552696
Log likelihood	-3.874725	Hannan-Quinn criter.	0.462425
F-statistic	5.769216	Durbin-Watson stat	1.820560
Prob(F-statistic)	0.007587		

- **AR(1) 오차모형의 GLS 추정법에서 첫번째 표본의 추정 방법**

- 아래의 변형에서는 첫번째 표본 관찰치를 구할 수 없음

$$y_t^* = y_t - \hat{\rho} y_{t-1} \qquad x_{t1}^* = 1 - \hat{\rho} \qquad x_{t2}^* = x_t - \hat{\rho} x_{t-1}$$

$t = 2, 3, \cdots, T$ 에 대해서만 $(y_t^*, x_{t1}^*, x_{t2}^*)$ 를 구할 수 있음

- $t = 1$ 에 대응되는 $(y_1^*, x_{11}^*, x_{12}^*)$ 를 구하는 방법

(Prais-Winstern 추정량)

$$y_1^* = \sqrt{1-\rho^2}\, y_1 \qquad x_{11}^* = \sqrt{1-\rho^2}$$

$$x_{12}^* = \sqrt{1-\rho^2}\, x_1 \qquad e_1^* = \sqrt{1-\rho^2}\, e_1$$

Cf. $\text{var}(e_1^*) = (1-\rho^2)\,\text{var}(e_1) = (1-\rho^2)\dfrac{\sigma_v^2}{1-\rho^2} = \sigma_v^2$ 동분산, 독립성 충족

▪ 동태적 모형 사이의 관계

- 오차항의 자기상관을 명시적으로 모형화하는 방법

$$\boxed{y_t = \beta_1 + \beta_2 x_t + e_t \quad e_t = \rho e_{t-1} + v_t}$$

$$\widehat{\ln(A_t)} = 3.899 + 0.888\ln(P_t) \qquad e_t = 0.422 e_{t-1} + v_t$$
$$\text{(se)} \quad (0.092) \; (0.259) \qquad\qquad (0.166)$$

- 정태적 선형모형에 시차종속변수와 시차설명변수를 포함시키는 방법

$$y_t = \beta_1 + \beta_2 x_t + \rho y_{t-1} - \rho\beta_1 - \rho\beta_2 x_{t-1} + v_t$$

$$y_t = \beta_1(1-\rho) + \beta_2 x_t - \rho\beta_2 x_{t-1} + \rho y_{t-1} + v_t$$

$$\boxed{y_t = \delta + \delta_0 x_t + \delta_1 x_{t-1} + \theta_1 y_{t-1} + v_t}$$

$$\hat{y}_t = 2.366 + 0.777 x_t - 0.611\, x_{t-1} + 0.404 y_{t-1}$$
$$\text{(se)} \; (0.656) \; (0.280) \quad (0.297) \quad (0.167)$$

- 두 방법은 동일한 효과

- 시차종속변수와 시차설명변수가 포함된 모형

Dependent Variable: LNA
Method: Least Squares
Date: 05/19/11 Time: 15:32
Sample (adjusted): 2 34
Included observations: 33 after adjustments
LNA=C(1)+C(2)*LNP+C(3)*LNP(-1)+C(4)*LNA(-1)

$$\hat{y}_t = 2.366 + 0.777x_t - 0.611\,x_{t-1} + 0.404y_{t-1}$$
$$(\text{se})\ (0.656)\ (0.280)\quad (0.297)\qquad (0.167)$$

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	2.366173	0.655701	3.608615	0.0011
C(2)	0.776629	0.279813	2.775530	0.0095
C(3)	-0.610862	0.296644	-2.059245	0.0485
C(4)	0.404284	0.166624	2.426323	0.0217

R-squared	0.304297	Mean dependent var	3.999309
Adjusted R-squared	0.232328	S.D. dependent var	0.325164
S.E. of regression	0.284899	Akaike info criterion	0.439845
Sum squared resid	2.353848	Schwarz criterion	0.621240
Log likelihood	-3.257444	Hannan-Quinn criter.	0.500879
F-statistic	4.228155	Durbin-Watson stat	1.843321
Prob(F-statistic)	0.013492		

- **AR(1) 추정을 위한 SAS program**

```
data sugar ;                                  * create dataset;
infile 'C:₩tmp₩table12-1.prn' ;               * open data file;
input a p ;                                   * input variables;
y = log(a) ;                                  * take logs;
x = log(p) ;

proc autoreg data=sugar ;          * estimate autoregressive model;
model y = x / nlag=1 ;             * GLS for AR(1) via 2-step estimator;
run ;
```

9.4 자기상관에 대한 검정

- 추정된 식에 자기상관이 존재하는지를 검정하는 세 가지 방법
  - 잔차 상관도표(residual correlogram)를 그려 살펴보는 방법
  - LM 검정 (Lagrange multiplier test)
  - DW 검정 (Durbin-Watson test)

9.4.1 잔차 상관도표를 이용한 자기상관 검정법 (1 시차 경우)

- AR(1) 모형

$$y_t = \beta_1 + \beta_2 x_t + e_t \qquad e_t = \rho e_{t-1} + v_t$$

v_t는 무작위 오차이고, *iid* 확률변수임

- 가설: $H_0 : \rho = 0 \quad H_1 : \rho \neq 0$
- 검정통계량: $z = \sqrt{T} r_1 \sim N(0,1) \qquad r_1 = \mathrm{corr}(\hat{e}_t, \hat{e}_{t-1})$
- 5% 유의수준의 채택역: $-1.96 \le z \le +1.96$
- 사탕수수 예 경우: $z = \sqrt{34} \times 0.404 = 2.36 \ge 1.96$
- 결론: 오차가 자기상관되어 있음. AR(1) 모형이 적절함.

9.4.1 잔차 상관도표를 이용한 자기상관 검정법 (모든 시차 경우)

- 앞에서는 잔차의 1시차 자기상관만을 검정하였음

- 이 경우 기각역은 다음과 같이 두 가지로 표현할 수 있음

$$\sqrt{T}r_1 \le -1.96,\quad \sqrt{T}r_1 \ge 1.96 \quad \text{혹은} \quad r_1 \le -\frac{1.96}{\sqrt{T}},\quad r_1 \ge \frac{1.96}{\sqrt{T}}$$

- 잔차 상관도표: 모든 시차에서의 잔차의 표본 자기상관 함수

$$r_k = \text{corr}(\hat{e}_t, \hat{e}_{t-k}),\quad k = 1, 2, ..., k$$

- k 시차의 자기상관을 검정하는 기각역: $r_k \le -\dfrac{1.96}{\sqrt{T}},\quad r_k \ge \dfrac{1.96}{\sqrt{T}}$ (5% 유의수준 경우)

9.4.1 잔차 상관도표를 이용한 자기상관 검정법

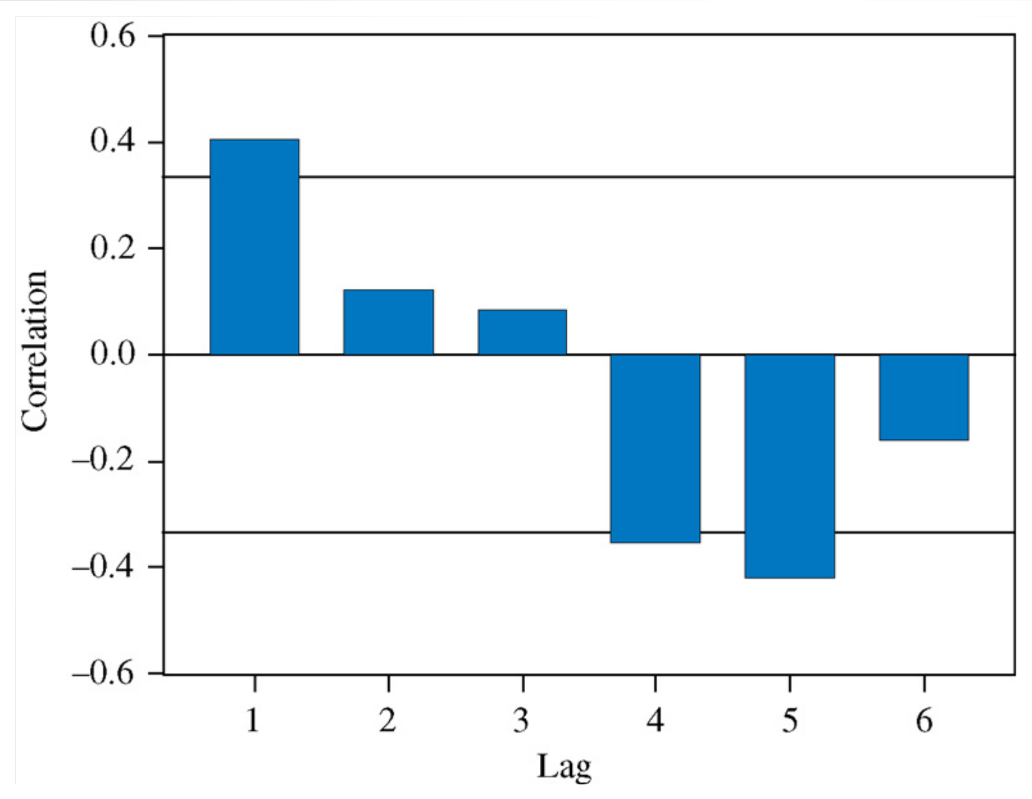
- 대부분 SW에서는 잔차 상관도표와 $\pm\frac{1.96}{\sqrt{T}}$ 를 함께 보여줌
- 잔차 자기상관 그림이 $\pm\frac{1.96}{\sqrt{T}}$ 보다 크거나 작은 경우 나타나면,

$\Rightarrow$ 5% 유의수준에서 자기상관이 존재한다고 결론 내리면 됨

9.4.1 잔차 상관도표를 이용한 자기상관 검정 (사탕수수 경우)

- $y_t = \beta_1 + \beta_2 x_t + e_t$ 의 경우 (6 시차까지만 검정한다면)

$$r_1 = 0.404,\ r_2 = 0.122,\ r_3 = 0.084,\ r_4 = -0.353,\ r_5 = -0.420,\ r_6 = -0.161$$

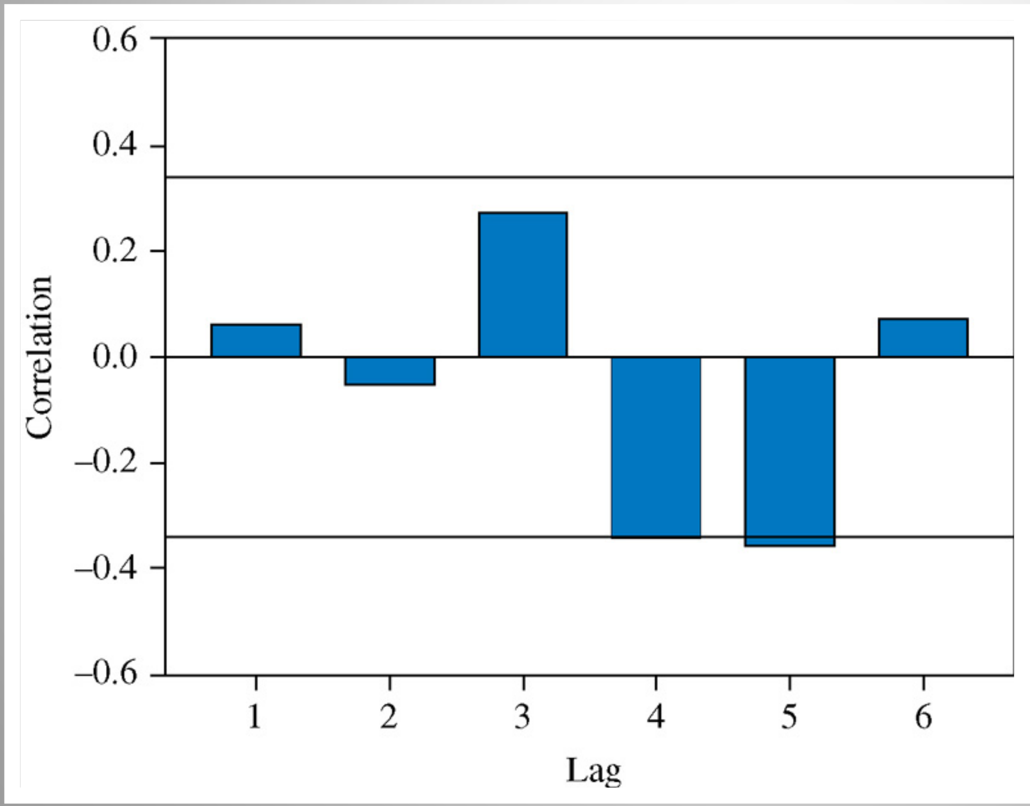


자기상관 존재함

(1, 4, 5 시차에서)

9.4.1 잔차 상관도표를 이용한 자기상관 검정 (사탕수수 경우)

- AR(1) 모형의 경우 $y_t = \beta_1 + \beta_2 x_t + e_t \qquad e_t = \rho e_{t-1} + v_t$
- 혹은 $y_t = \beta_1(1-\rho) + \beta_2 x_t + \rho y_{t-1} - \rho\beta_2 x_{t-1} + v_t$ 모형 경우



자기상관 존재함

(4, 5 시차에서)

그렇지만 선형모형보다 크게 완화되었음을 알 수 있음

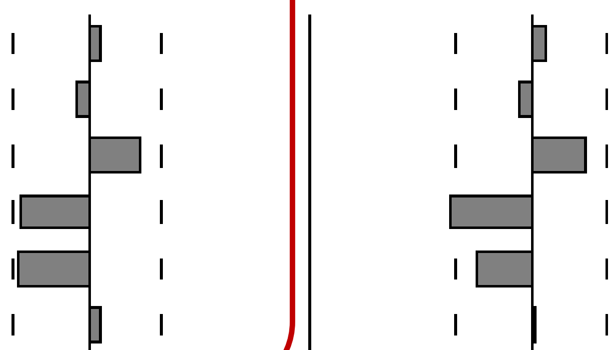
▪ 잔차 상관도표 (사탕수수 경우): Eviews output

Correlogram of Residuals

Date: 04/24/11 Time: 22:49
Sample: 2 34
Included observations: 33
Q-statistic probabilities adjusted for 1 ARMA term(s)

OLS 추정 후,
View-Residual Tests 클릭,
Correlogram-Q-Statistics 클릭,
적절한 시차 입력

Autocorrelation | Partial Correlation



	AC	PAC	Q-Stat	Prob
1	0.061	0.061	0.1352	
2	-0.048	-0.052	0.2200	0.639
3	0.252	0.260	2.6741	0.263
4	-0.312	-0.378	6.5429	0.088
5	-0.320	-0.258	10.775	0.029
6	0.064	0.022	10.952	0.052

9.4.2 LM 검정을 이용한 자기상관 검정법

- 자기상관 탐지를 위한 AR(1) 모형과 가설

$$y_t = \beta_1 + \beta_2 x_t + \rho e_{t-1} + v_t \qquad H_0 : \rho = 0, \qquad H_1 : \rho > 0$$

- 만약 e_{t-1} 를 관찰할 수 있다면,
  통상적인 모수 추정치에 대한 t-검정 실시하면 됨
- e_{t-1} 는 관찰될 수 없기 때문에, 대신 $\hat{e}_{t-1}$ 이용해 검정 가능
  이렇게 하는 것을 LM 검정이라 함 $(\hat{e}_{t-1} = y_{t-1} - b_1 - b_2 x_{t-1})$
- 사탕수수 예에서의 검정 결과는 다음과 같음

  t=2.439 F=5.949 p-값=0.021
- 5% 유의수준에서 귀무가설 $H_0 : \rho = 0$ 을 기각함
  (즉 자기상관 존재함)

9.4.2 LM 검정을 이용한 자기상관 검정법 ($LM = T \times R^2$ 이용)

- 자기상관 탐지를 위한 AR(1) 모형을 변형

$$y_t = \beta_1 + \beta_2 x_t + \rho e_{t-1} + v_t$$

$$b_1 + b_2 x_t + \hat{e}_t = \beta_1 + \beta_2 x_t + \rho \hat{e}_{t-1} + \hat{v}_t$$

- 다음의 보조회귀 수행 $\hat{e}_t = (\beta_1 - b_1) + (\beta_2 - b_2) x_t + \rho \hat{e}_{t-1} + \hat{v}_t$

$$= \gamma_1 + \gamma_2 x_t + \rho \hat{e}_{t-1} + \hat{v}_t$$

- $\gamma_1 = (\beta_1 - b_1)$ 과 $\gamma_2 = (\beta_2 - b_2)$ 는 거의 0에 가까우므로,
  위 회귀식의 설명력은 $\hat{e}_{t-1}$ 에서 비롯됨

- $H_0 : \rho = 0$ 이 참인 경우, (R^2는 위 식의 결정계수)
  $LM = T \times R^2$ 은 대략 $\chi^2(1)$ 분포를 따름

9.4.2 LM 검정을 이용한 자기상관 검정법 (사탕수수 예)

$$LM = T \times R^2 = 34 \times 0.16101 = 5.474$$

- $\chi^2(1)$ 분포에서 유의수준 5%일 경우 임계값은 3.84
- 따라서 귀무가설 기각 (즉, 자기상관 존재함: p-값=0.019)

❖ LM 검정은 다양한 변종과 일반화된 방법이 있음

- 교과서 p. 338에 일부 소개되어 있음

- **LM test, SAS program**

```
data sugar ;                          * create dataset;
infile 'C:₩tmp₩table12-1.prn' ;       * open data file;
input a p ;                           * input variables;
y = log(a) ;                          * take logs;
x = log(p) ;

proc autoreg data=sugar;              * estimate autoregressive model;
model y = x  ;
output out=sugarout r=ehat;           * output residuals;

data lmtest;                          * create data set;
set sugarout;                         * read sugar output;
ehat_1=lag(ehat);                     * lagged residuals;
if t=1 then ehat_1=0;

proc autoreg data=lmtest;             * estimate lm model;
lmtest:model y = x ehat_1;            * ehat(-1) added to model;
test ehat_1=0;
run ;
```

- **Godfrey (1988)의 LM test, SAS program**

```
data sugar ;                              * create dataset;
infile 'C:₩tmp₩table12-1.prn' ;           * open data file;
input a p ;                               * input variables;
y = log(a) ;                              * take logs;
x = log(p) ;

proc autoreg data=sugar ;          * estimate autoregressive model;
model y = x / godfrey=1 ;          * automatic LM test for AR(1) errors;
run ;
```

▪ 사탕수수 예: LM test 결과, Eviews output

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	5.949152	Prob. F(1,31)	0.0206
Obs*R-squared	5.474312	Prob. Chi-Square(1)	0.0193

Test Equation:
Dependent Variable: RESID
Method: Least Squares
Date: 04/24/11 Time: 23:03
Sample: 1 34
Included observations: 34
Presample missing value lagged residuals set to zero.

OLS 추정 후,
View-Residual Tests 클릭,
Serial Correlation LM test 클릭,
적절한 시차 (1) 입력

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.008116	0.057186	-0.141927	0.8881
LNP	0.091601	0.260934	0.351052	0.7279
RESID(-1)	0.407821	0.167202	2.439088	0.0206

R-squared	0.161009	Mean dependent var	2.45E-17
Adjusted R-squared	0.106881	S.D. dependent var	0.303094
S.E. of regression	0.286439	Akaike info criterion	0.421513
Sum squared resid	2.543460	Schwarz criterion	0.556192
Log likelihood	-4.165715	Hannan-Quinn criter.	0.467442
F-statistic	2.974576	Durbin-Watson stat	1.905835
Prob(F-statistic)	0.065802		

9.4.3 더빈-왓슨 검정 (Durbin-Watson test)

$$y_t = \beta_1 + \beta_2 x_t + e_t \qquad e_t = \rho e_{t-1} + v_t$$

$H_0 : \rho = 0, \quad H_1 : \rho > 0$ (양의 자기상관이 일반적인 점을 고려)

- DW 통계량: $$d = \frac{\sum_{t=2}^{T} \left(\hat{e}_t - \hat{e}_{t-1}\right)^2}{\sum_{t=1}^{T} \hat{e}_t^2} \qquad d \approx 2\left(1 - \hat{\rho}\right)$$

  $(\hat{e}_t = y_t - b_1 - b_2 x_t)$

- If $\hat{\rho} = 0 \quad \Rightarrow \quad d \approx 2$ DW 값이 2에 가까우면, "자기상관 없음"
- If $\hat{\rho} \to +1 \quad \Rightarrow \quad d \to 0$ DW 값이 2보다 작을수록 "양의 자기상관"
- If $\hat{\rho} \to -1 \quad \Rightarrow \quad d \to 4$ DW 값이 2보다 클수록 "음의 자기상관"

▪ 사탕수수 예: OLS 추정결과

Dependent Variable: LNA
Method: Least Squares
Date: 04/24/11 Time: 22:14
Sample: 1 34
Included observations: 34
LNA=C(1)+C(2)*LNP

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	3.893256	0.061345	63.46486	0.0000
C(2)	0.776119	0.277467	2.797154	0.0087

R-squared	0.196466	Mean dependent var	3.980680
Adjusted R-squared	0.171355	S.D. dependent var	0.338123
S.E. of regression	0.307793	Akaike info criterion	0.538245
Sum squared resid	3.031571	Schwarz criterion	0.628031
Log likelihood	-7.150159	Hannan-Quinn criter.	0.568864
F-statistic	7.824072	Durbin-Watson stat	1.168987
Prob(F-statistic)	0.008653		

▪ 사탕수수 예: AR(1) 추정결과

Dependent Variable: LNA
Method: Least Squares
Date: 04/24/11 Time: 22:34
Sample (adjusted): 2 34
Included observations: 33 after adjustments
Convergence achieved after 7 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.898771	0.092165	42.30197	0.0000
LNP	0.888370	0.259299	3.426048	0.0018
AR(1)	0.422140	0.166047	2.542284	0.0164

R-squared	0.277777	Mean dependent var	3.999309
Adjusted R-squared	0.229629	S.D. dependent var	0.325164
S.E. of regression	0.285399	Akaike info criterion	0.416650
Sum squared resid	2.443575	Schwarz criterion	0.552696
Log likelihood	-3.874725	Hannan-Quinn criter.	0.462425
F-statistic	5.769216	Durbin-Watson stat	1.820560
Prob(F-statistic)	0.007587		

- **Durbin-Watson test의 임계치**

$$y_t = \beta_1 + \beta_2 x_t + e_t$$

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

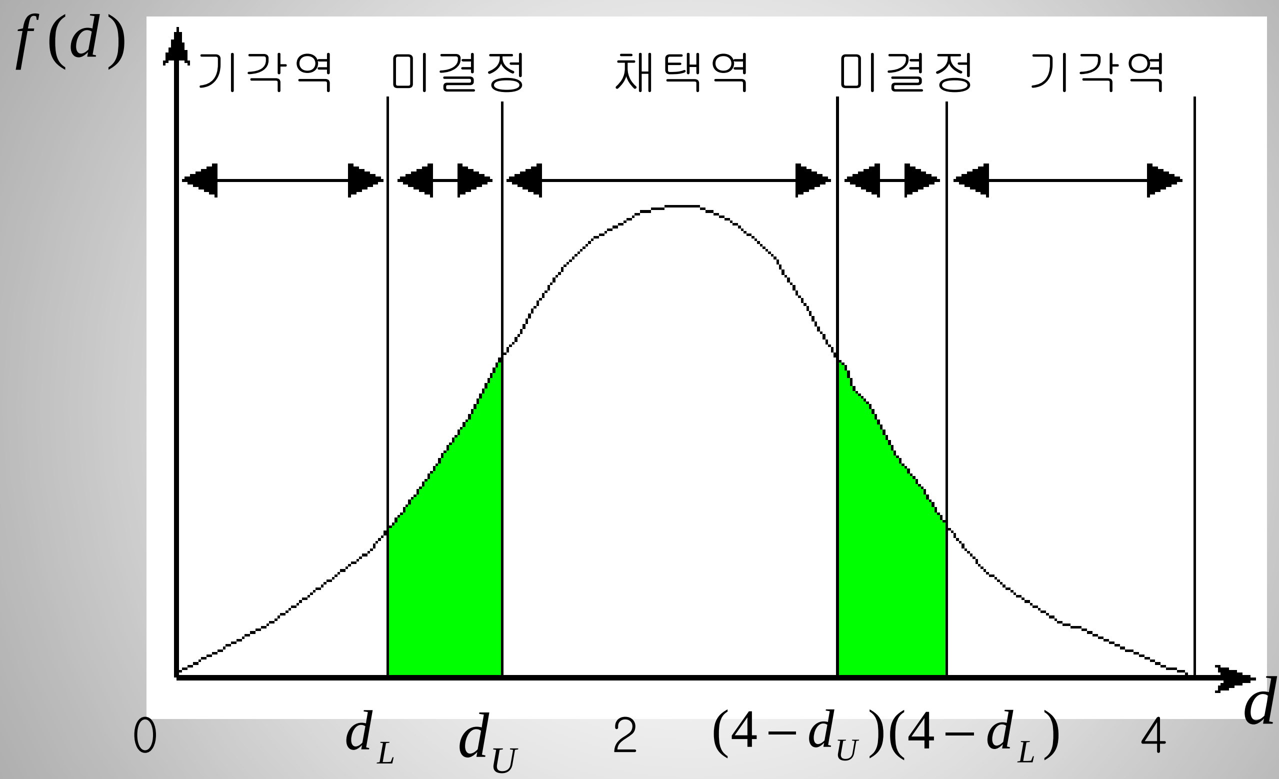
$$H_0 : \rho = 0, \qquad H_1 : \rho > 0$$

$$d \approx 2\left(1 - \hat{\rho}\right)$$

- 교과서(2판) <표 5> 이용, 두 개의 임계치 이용하여 검정함

- $d < d_{LC}$ 이면, $H_0 : \rho = 0$ 기각하고 $H_1 : \rho > 0$ 채택함
- $d > d_{UC}$ 이면, $H_0 : \rho = 0$ 기각하지 않음
- $d_{LC} < d < d_{UC}$ 인 경우에는 결론을 내리지 못함

- **Durbin-Watson test의 결정규칙**



Durbin-Watson Statistic: 5 Per Cent Significance Points of dL and dU

	k'*=1		k'=2		k'=3		k'=4		k'=5		k'=6		k'=7		k'=8		k'=9		k'=10	
n	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU
6	0.610	1.400	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
7	0.700	1.356	0.467	1.896	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
8	0.763	1.332	0.559	1.777	0.367	2.287	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
9	0.824	1.320	0.629	1.699	0.455	2.128	0.296	2.588	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
10	0.879	1.320	0.697	1.641	0.525	2.016	0.376	2.414	0.243	2.822	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
11	0.927	1.324	0.758	1.604	0.595	1.928	0.444	2.283	0.315	2.645	0.203	3.004	-----	-----	-----	-----	-----	-----	-----	-----
12	0.971	1.331	0.812	1.579	0.658	1.864	0.512	2.177	0.380	2.506	0.268	2.832	0.171	3.149	-----	-----	-----	-----	-----	-----
13	1.010	1.340	0.861	1.562	0.715	1.816	0.574	2.094	0.444	2.390	0.328	2.692	0.230	2.985	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
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
18	1.158	1.391	1.046	1.535	0.933	1.696	0.820	1.872	0.710	2.060	0.603	2.258	0.502	2.461	0.407	2.668	0.321	2.873	0.244	3.073
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
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
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
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
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
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
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
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
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
28	1.328	1.476	1.255	1.560	1.181	1.650	1.104	1.747	1.028	1.850	0.951	1.959	0.874	2.071	0.798	2.188	0.723	2.309	0.649	2.431
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	0.681	2.396
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
31	1.363	1.496	1.297	1.570	1.229	1.650	1.160	1.735	1.090	1.825	1.020	1.920	0.950	2.018	0.879	2.120	0.810	2.226	0.741	2.333
32	1.373	1.502	1.309	1.574	1.244	1.650	1.177	1.732	1.109	1.819	1.041	1.909	0.972	2.004	0.904	2.102	0.836	2.203	0.769	2.306
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
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
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	0.971	2.054	0.908	2.144	0.845	2.236

*k' is the number of regressors excluding the intercept

- 사탕수수 재배면적 사례의 **Durbin-Watson test**

- DW 임계치 표에서 T=34, K*=1을 찾으면

$$\ln(A_t) = \beta_1 + \beta_2 \ln(P_t) + e_t$$

$d_{LC} = 1.393$, $d_{UC} = 1,514$

- OLS 추정결과에서 DW 통계량 계산값은 $d = 1.1689$

- $d(=1.1689) < d_{LC}(=1.393)$ 이므로, $H_0 : \rho = 0$ 를 기각함

- 유의수준 5%에서 사탕수수 재배면적 모형에는
  자기상관이 존재한다고 결론내릴 수 있음

❖ AR(1) 모형의 경우 DW 통계량 계산값은 $d = 1.8206$

- 자기상관 없다고 결론 내릴 수 있음

❖ DW test와 LM test 비교

- DW test 검정 결과와 LM test 검정결과가 다를 수 있음
  - LM test가 제2종 오류를 범했을 가능성이 높음 (검정력 약함)
- DW test는 오차항이 정규분포할 경우 소표본에도 적용 가능,
  LM test는 대규모 표본일 때 적용 가능
- 설명변수에 시차 종속변수 y_{t-1} 가 포함되는 경우는
  DW test는 부적절, LM test 사용해야 함
- e_{t-2}, e_{t-3} 등 고차적 시차를 포함하는 자기상관 구조를 검정하기 위해서는 F-검정을 이용하여 LM test 사용해야 함