시계열 자료 분석(Time series data analysis)

Part I. 계량 시계열 모형 (Econometric Time Series Model) – 기초 개념

I. 정상성(Stationarity)

A. 시계열 변수 y_t 는 완전한 예측을 할 수 없으므로 확률변수이며, y_t 를 발생시키는 모형을 확률과정(stochastic process or random process)라고 함

i. 그러한 확률과정의 실현(realization)인 y_t 의 값들의 표본이 관측됨

ii. 시계열 자료를 사용하는 회귀모형에서 최소제곱 추정량의 통상적 성질은 그 시계열 변수가 定常的(stationary) 확률과정으로부터 발생되었는가 여부에 의존함

B. 시계열 y_t 가 평균과 분산이 시간에 무관한 상수이고 두 값 간의 자기공분산이 오직 그 두 값간의 시차의 길이에 의존하는 경우 (약) 정상적(공분산(covariance)정상적, 이차(second-order)정상적)이라고 함

cf. 강(strictly) 정상적 – 모든 적률이 시간에 무관하게 일정 :

a. 약정상적 시계열이 정규분포를 한다면 강정상적임

i. 즉, 다음과 같은 조건을 모든 값에 대해 만족하는 시계열 y_t 를 정상 시계열이라고 함

1. $E(y_t) = \mu$ [일정한 평균]

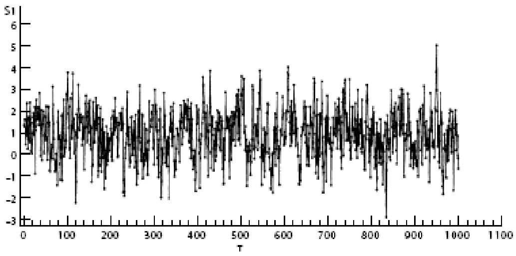
2. $\mathrm{var}(y_t) = \sigma^2$ [일정한 분산]

3. $\mathrm{cov}(y_t, y_{t+s}) = \mathrm{cov}(y_t, y_{t-s}) = \gamma_s$ [공분산은 t 가 아닌 s 에 의존함]

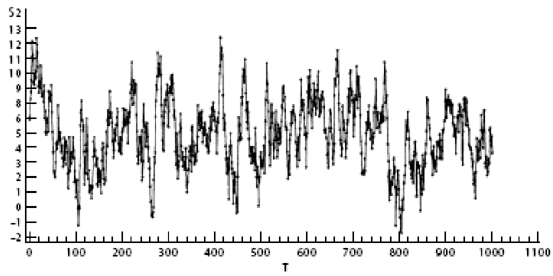
a. 정상 시계열은 어떤 시점에서 평균과 분산 그리고 특정한 시차의 길이를 갖는 자기공분산을 측정하더라도 동일한 값을 지님

b. 정상 시계열은 항상 그 평균값으로 회귀하려는 경향이 있으며, 그 평균값 주변에서의 변동은 대체로 일정한 폭을 가지게 됨

c. 정상 시계열이 아닌 경우 특정 기간의 시계열 자료로부터 얻은 정보를 다른 시기로 일반화 할 수 없음



$y_t = 0.5 + 0.5\, y_{t-1} + e_t$, $e_t \sim$ iid $N(0,1)$



$y_t = 0.5 + 0.9\, y_{t-1} + e_t$

ii. 백색잡음(white noise) 과정

1. 시계열 e_t 의 평균이 0 이고 분산이 일정한 값 σ^2 이고 자기공분산이 0 인 경우 이를 백색잡음 과정이라고 함 (표준모형에서의 오차항의 성질)
   a. 좀 더 강한 조건 즉 시계열간 확률적 독립인 경우 이는 강(strictly) 백색 잡음 과정임
   b. 백색잡음 과정이 정규분포를 따를 경우 이를 가우시안(Gaussian) 백색잡음 과정이라고도 함

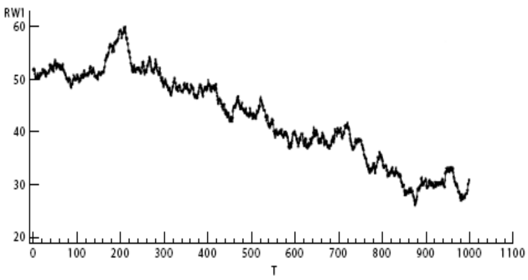
II. 임의보행(random walk, 확률적 보행이라고도 함) 과정

A. 경제 시계열 변수에 있어서 부딪히게 되는 대표적인 비정상 시계열은 임의보행(random walk, 확률적 보행이라고도 함) 과정으로부터 발생함

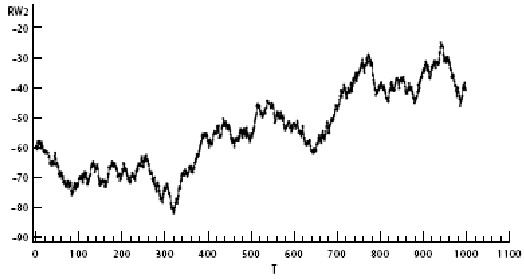
i. $y_t = y_{t-1} + e_t$, , e_t :백색잡음 ⇒ 상수항 혹은 추세선이 없는 임의보행을 따르는 시계열 ((random walk process without a drift)

1. $E(y_t) = E(e_t + e_{t-1} + \ldots) = 0$
2. $V(y_t) = V(e_t + e_{t-1} + \ldots) = \sum_{t=1}^{\infty} \sigma^2 = \infty$
   a. ⇒ 비정상 시계열임 (분산이 무한히 커짐)
   b. ⇒ 확률적 충격의 지속성(the persistence of random shocks) 혹은 영구적 기억(infinite memory)를 가짐
   c. ⇒ 1차 차분값은 백색잡음 과정이 됨
      i. $\Delta y_t = y_t - y_{t-1} = e_t$,
3. 임의보행(random walk)과정의 모습

$y_t = y_{t-1} + 0.5e_t$, e_t~iid N(0,1)



$y_t = y_{t-1} + e_t$,



ii. $y_t = a + y_{t-1} + e_t$, , e_t :백색잡음 ⇒ 상수항을 갖는 (추세선을 갖는) 임의보행 과정(random walk process with a drift)

1. $E(y_t) = E(a + e_t + a + e_{t-1} + \ldots) = \sum_{t=1}^{\infty} a = \infty$

2. $V(y_t)= V(y_0+a+e_1+a+e_2+ \dots a+e_t)= \sum_{t=1}^{\infty}\sigma^2 = \infty \Rightarrow$ 비정상 시계열임 (분산, 평균 모두 무한히 커짐)

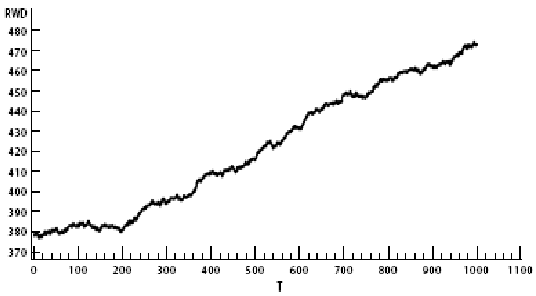
   a. $\Rightarrow$ a (상수항)이 0 보다 클 경우 상방으로 0 보다 작을 경우 하방으로 흘러감(drift)

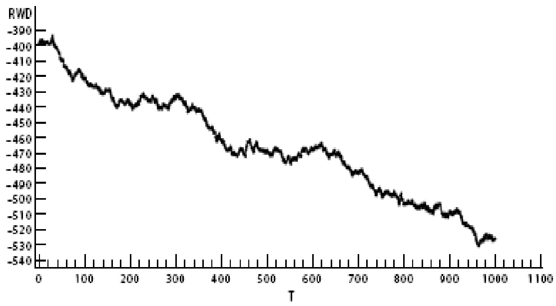
      i. $\Delta y_t = y_t - y_{t-1} = a+ e_t$,

      ii. 확률적 추세(stochastic trend)를 갖는 시계열임

3. 상수함을 갖는 임의 보행 혹은 확률적 보행(random walk with a drift)

$y_t = 0.1+y_{t-1} + 0.5e_t$, $e_t \sim iid\ N(0,1)$ $\qquad$ $y_t =-0.1+y_{t-1} + e_t$,





II. 단위근(unit root)

B. 이러한 임의보행 과정은 단위근(unit root)을 갖는 시계열의 한 예임

i. $y_t = \beta y_{t-1} + e_t, -1 \le \beta \le 1$

1. $\Rightarrow \beta=1$ 인 경우, 즉 임의보행과정은 단위근을 가짐

   a. $(1 - \beta L)y_t = a + e_t$, L is the lag operator : $L^n y_t = y_{t-n}$

   b. $\Rightarrow$ 단위근이라는 용어는 lag operator 의 다항식의 근을 의미하는 것임. $(1-\beta L)=0$

      i. 시계열에 따라서 한 개 이상의 단위근을 갖는 경우도 존재함

2. $\Rightarrow$ 비정상성, 단위근, 임의보행은 유사하게 취급될 수 있는 용어임 (비정상성>단위근>임의보행의 순으로 개념이 포괄하는 바가 좁아짐)

3. $\Rightarrow$ $-1<\beta<1$ 인 경우 이 시계열은 정상 시계열임

   a. $E(y_t)= E(e_t+\beta e_{t-1}+ \beta^2 e_{t-2}+ \dots)=0$ (가정)

   b. $V(y_t)= V(e_t+\beta e_{t-1}+ \beta^2 e_{t-2}+ \dots)= \dfrac{\sigma^2}{1-\beta}$

   c. $Cov(y_t,y_{t-s})= E[(e_t+\beta e_{t-1}+ \beta^2 e_{t-2}+ \dots)(e_{t-s}+\beta e_{t-s-1}+ \beta^2 e_{t-s-2}+ \dots)]$

$$= E(\beta^s e_{t-s}\, e_{t-s}+(\beta^{s+2} e_{t-s-1}\, e_{t-s-1}+ \dots) = \frac{\beta^s\sigma^2}{1-\beta^2}$$

C. 실제 경제 시계열의 분석에 있어서 해당 시계열이 단위근을 가지고 있는가를 판별하는 것이 중요함 = 다양한 단위근 검정 방법들 (추후에 설명)

III. 추세 정상(trend stationary) 확률과정과 차분 정상(difference stationary) 시계열

B. 차분 정상 시계열

i. 순수 임의보행: **$\Delta y_t = y_t - y_{t-1} = e_t$,**

ii. 상수항을 갖는 임의보행: **$\Delta y_t = y_t - y_{t-1} = a + e_t$,**

iii. ⇒ 이처럼 차분을 통해 정상 시계열이 되는 경우 이를 차분 정상 시계열이라고 함

C. 추세 정상 시계열

i. $y_t = \beta_1 + \beta_2 t + e_t$: 비확률추세(deterministic trend)를 가지고 있는 시계열임

1. $E(y_t) = \beta_1 + \beta_2 t$ **:** 평균이 시간의 함수이며 따라서 비정상 시계열임

2. $V(y_t) = \sigma^2$ **:** 분산은 상수

3. $\Rightarrow\ y_t - E(y_t) = y_t - (\beta_1 + \beta_2 t) = e_t$: 정상 시계열

ii. 이처럼 비확률 추세를 제거(detrend)할 경우 정상 시계열이 되는 비정상 시계열을 추세 정상 시계열이라고 함

D. $y_t = \beta_1 + \beta_2 t + y_{t-1} + e_t$ **:** 상수항을 갖는 임의보행에 비확률 추세가 포함됨

i. 즉 확률적 추세와 비확률 추세가 모두 포함됨

ii. $\Rightarrow\ \Delta y_t = \beta_1 + \beta_2 t + e_t$: 차분을 해도 여전히 비정상 시계열임

iii. $\Rightarrow\ \Delta y_t - E(\Delta y_t) = e_t$ **:** 차분을 통해 얻은 시계열을 추세제거함으로써 정상시계열을 얻을 수 있음

IV. 누적확률과정(integrated stochastic process)

B. 1 차 차분을 통해 정상 시계열이 되는 비정상 시계열을 1 차 누적되었다고 함 $\Rightarrow\ y_t \sim I(1)\ \Rightarrow\ \Delta y_t \sim I(0)$

C. d 차 차분을 해야 정상 시계열이 되는 비정상 시계열을 d 차 누적되었다고 함⇒ $y_t \sim I(d)$

D. 누적 시계열의 성질

i. $y_t \sim I(d) \Rightarrow z_t \equiv (a + b y_t) \sim I(d)$, **b≠0**

ii. $y_t \sim I(d_1),\ x_t \sim I(d_2)\ \Rightarrow\ z_t \equiv (a x_t + b y_t + c) \sim I(\max(d_1, d_2))$

1. $y_t \sim I(1),\ x_t \sim I(0)\ \Rightarrow\ z_t \equiv (a x_t + b y_t + c) \sim I(1)$

iii. $y_t \sim I(d)$, $x_t \sim I(d) \Rightarrow z_t \equiv (ax_t + by_t + c) \sim I(d^*),\ d^* \le d$

V. 허구적 회귀(spurious regression)

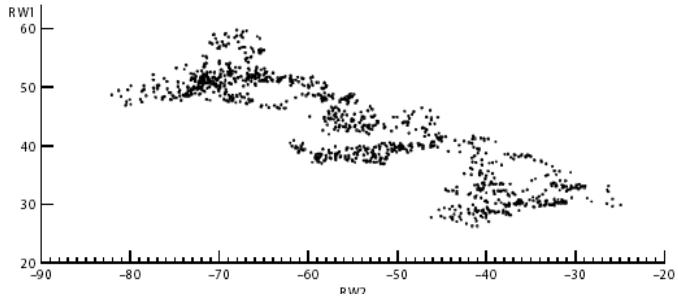
B. 회귀분석에 있어서 비정상 시계열 자료를 사용할 경우 아무런 관련 없는 변수들간에 매우 유의한 것처럼 보이는 결과를 얻을 수 있으며, 이러한 회귀분석을 허구적 회귀라고 함

C. 실례

i. $\mathbf{y_t = y_{t-1} + 0.5e_t}$, $\mathbf{e_t \sim iid\ N(0,1)}$ $\mathbf{x_t = x_{t-1} + u_t}$, $\mathbf{u_t \sim iid\ N(0,1)}$

1. 이들 시계열들은 독립적인 확률과정으로부터 발생된 것이며 서로 아무런 관련이 없음

ii. 이 두 임의보행 시계열을 가지고 산포도를 그려보면 다음과 같은 역의 관계를 볼 수 있음



iii. 실제로 $\mathbf{y_t}$(rw1)를 $\mathbf{x_t}$(rw2)에 회귀할 경우 다음과 같은 결과를 얻을 수 있음

R^2 0.7495 Durbin-Watson 0.0305

Variable	DF	B Value	Std Error	t Ratio	Approx Prob
Intercept	1	14.204040	0.5429	26.162	0.0001
rw2	1	-0.526263	0.00963	-54.667	0.0001

iv. 이들 시계열들은 서로 아무런 관련이 없음에도 불구하고 회귀분석 결과 높은 유의성을 갖는 것으로 나타남.

1. 이러한 결과는 아무런 의미가 없으며, 허구적(spurious)임 (Yule, 1974)
2. $\mathbf{\Delta y_t}$(rw1)를 $\mathbf{\Delta x_t}$(rw2)에 회귀할 경우 아무런 유의성이 없는 것으로 나타날 것임

v. 위 회귀분석 결과는 극도로 낮은 DW 값을 통해 무언가 문제가 있음을 알리고 있음

1. Granger and Newbold(1974)에 따르면, 경험적으로 R^2 값이 DW 값보다 큰 경우 이러한 허구적 회귀 가능성을 의심해보아야 한다고 함

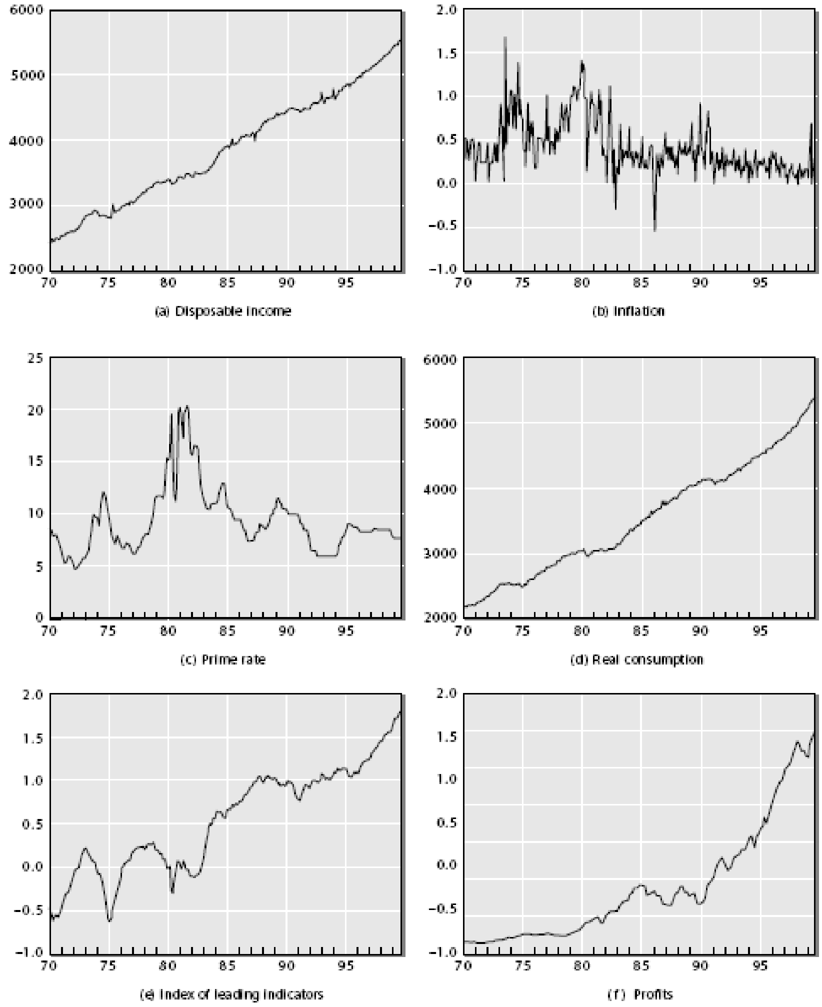
vi. 이러한 예는 시계열을 기반으로 한 회귀분석을 수행함에 있어서 극도의 주의를 필요로 함을 시사함

VI.　 단위근 검정

B. 통상적인 시계열의 정상성에 대한 판단

i.　 시간 흐름에 따른 변화를 그래프로 나타내어 판단

1. 주요 거시 경제 변수들의 그래프



ii.　 표본상관도표(Correlogram)을 이용하여 판단

1. 표본 자기상관함수(SACF, sample autocorrelation function)

a. 시차 k 의 자기상관함수(ACF)는 다음과 같이 정의됨

i.　 $$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\mathrm{cov}(y_t, y_{t+k})}{\mathrm{var}(y_t)}$$

ii.　 ρ_k 는 보통의 상관과 마찬가지로 -1 과 1 사이의 값을 가지며, k=0 일 때 1 의 값을 가짐

b. 표본 자기 상관 함수는 다음과 같이 계산됨

i. $\hat{\gamma}_k = \dfrac{\sum(y_t - \overline{y})(y_{t+k} - \overline{y})}{T}$

ii. $\hat{\gamma}_0 = \dfrac{\sum(y_t - \overline{y})^2}{T}$

iii. $\Rightarrow \hat{\rho}_k = \dfrac{\hat{\gamma}_k}{\hat{\gamma}_0}$

2. 1 부터 시작하는 시차에 대해 $\hat{\rho}_k$ 를 그려준 것을 표본상관도표라함
   a. 백색잡음의 표본상관도표
      1. 0 주위에서 움직이면서 통계적으로 0 이라는 것을 기각할 수 없는 값을 가짐
   b. 상수항 없는 임의보행의 표본상관도표
      1. 통계적으로 0 이라는 것을 기각하는 매우 높은 값을 가지며 시차가 증가함에 따라 매우 완만하게 값이 감소함
   c. 시차의 결정
      i. 자기상관계수를 계산하기 위한 시차의 길이를 결정하는 문제는 경험적 문제로서 통상 전체 시계열의 3 분지 1 이나 4 분지 1 로 함
3. 자기상관계수의 통계적 유의성에 대한 판단
   a. Bartlett
      i. Bartlett 는 근사적으로 $\hat{\rho}_k \sim N\left(0, \frac{1}{k}\right)$ 임을 보임
      ii. $\dfrac{\hat{\rho}_k}{\sqrt{1/k}} > z_\alpha$, $\dfrac{\hat{\rho}_k}{\sqrt{1/k}} < -z_\alpha$ 일 경우 $\hat{\rho}_k$ 는 통계적으로 0 이라는 가설이 기각됨
   b. Box-Pierce Q-Statistic
      i. 개별적 자기상관계수의 유의성 검정이 아니라 특정 시차까지의 자기상관계수가 모두 0 이라는 복합가설을 검정할 경우 사용되는 검정통계량
         1. $Q = T\sum_{k=1}^{n} \hat{\rho}_k^2 \sim \chi^2(n)$ (근사적으로 카이제곱 분포를

함)

2. 특정한 시계열이 백색잡음인가를 검정할 경우 사용됨

ii. Ljung-Box(LB) statistic

1. $LB = T(T+2)\sum_{k=1}^{n}\left(\frac{\hat{\rho}_k^2}{T-k}\right) \sim \chi^2(n)$ (근사적으로)

2. 대표본에서는 Q 나 LB 모두 카이제곱 분포로 수렴하나, LB 의 소표본 성질이 보다 나음이 알려져 있음(즉 검정력이 높음)

iii. 단위근 검정(Unit root test)

1. 지난 몇 년간 시계열의 정상성(혹은 비정상성)을 검정하는데 있어서 광범위하게 사용되어온 방법

a. $y_t = \rho y_{t-1} + v_t$, v_t는 분산이 σ_v^2인 백색잡음

b. 여기서 ρ=1 이면, 상기의 AR(1)과정은 단위근을 가지는데, 이 때 y_t는 상수항이 없는 임의보행 $y_t = y_{t-1} + v_t$ 이며, 이는 비정상 임.

c. 반면에 $|\rho| < 1$이면, AR(1)과정은 안정적임

d. 시계열 y_t 가 비정상인가의 여부를 ρ = 1 이라는 귀무가설과 $|\rho| < 1$이라는 대립가설 (또는 단순히 ρ < 1)에 대한 검정을 통해 판단할 수 있음 (⇒ 단위근 검정)

2. 이러한 검정을 위한 검정 통계량은 다음과 같은 약간의 변형된 식으로부터 얻어짐

$$\begin{aligned} y_t - y_{t-1} &= \rho y_{t-1} - y_{t-1} + v_t \\ \Delta y_t &= (\rho - 1) y_{t-1} + v_t \quad , \\ &= \gamma y_{t-1} + v_t \end{aligned}$$

단, $\gamma \equiv \rho - 1$

$$\Rightarrow \begin{matrix} H_0 : \rho = 1 & \leftrightarrow & H_0 : \gamma = 0 \\ H_1 : \rho < 1 & \leftrightarrow & H_1 : \gamma < 0 \end{matrix}$$

a. y_t 가 임의보행을 따른 경우 γ = 0 이고 $\Delta y_t = y_t - y_{t-1} = v_t$ 이며, 이 일계차분값 $\Delta y_t = y_t - y_{t-1}$ 은 백색잡음이 됨.(즉 안정적 시계열임)

3. Dickey-Fuller 검정

a. 주어진 시계열 y_t 의 비정상성 여부는 주어진 시계열 y_t 를 차분하여 그를 다시 y_{t-1} 에 회귀하여 얻는 계수의 추정치가 0 인지 0 보다 작은지를 검정하는 문제로 귀결됨

i. 불행히도, 통상적인 t 검정통계량은 귀무가설하에서(즉 $\gamma = 0$) 대표본하에서조차 t 분포를 하지 않으며 따라서 근사적으로 정규분포를 하지도 않음

ii. γ에 대한 t 값을 디키-풀러(DF) 검정통계량 또는 τ(tau) 검정통계량이라 하며, Dickey-Fuller 가 이 타우통계량에 대한 임계값(critical value)을 Monte Carlo 실험을 통해 계산하여 표로 제시함

1. 후에 Mackinnon 이 보다 확장된 표를 만들었으며, 일부 통계패키지에 포함됨

2. γ =0 이 기각될 경우, 즉 정상 시계열의 경우 통상적인 t 검정을 적용할 수 있음

b. DF 검정은 임의보행 과정이 상수항(drift)를 가질 경우, 그리고 비확률 추세를 포함할 경우 등을 고려하여 다음의 세 가지 경우에 대해 각각의 귀무가설을 검정할 수 있음

i. $\Delta y_t = \gamma y_{t-1} + v_t$

ii. $\Delta y_t = \alpha_0 + \gamma y_{t-1} + v_t$

iii. $\Delta y_t = \alpha_0 + \alpha_1 t + \gamma y_{t-1} + v_t$

iv. 위 세 가지 경우에 있어서 귀무가설은 모두 γ =0, 즉 해당 시계열이 비정상이라는 것이고 대립가설은 γ <0, 즉 해당 시계열이 평균 0 인 정상 시계열(i)이거나, 평균이 0 이 아닌($\alpha_0 / (1-\rho)$) 정상시계열(ii)이거나, 비확률추세 주변에서 정상적인 시계열이라는 것임

v. 중요한 것은 각각의 DF 검정에 있어서 γ =0 을 검정하기 위한 타우 검정통계량의 임계치가 각각 다르다는 것임

1. 사전에 어떤 모형설정이 옳은가를 알 수는 없음 (시행착오가 불가피)

2. 이들 임계치들은 통상적인 t 검정통계량의 임계치들보다 더 작은 값을 보임

A. 즉 통상적인 t 검정통계량을 적용할 경우 귀무가설을 과도하게 기각하는 쪽으로 편향된 검정을 하게 됨(시계열이 안정적이라고 판단하는 쪽으로 편향된 검정을 하게 됨)

3. 또한 $\alpha_0 = 0, \alpha_1 = 0$ 에 대한 결합 검정을 위해 통상적인 F 값을 구축할 경우, 이 또한 귀무가설하에서 F 분포를 하지 않으며, Dickey-Fuller 는 이 경우에 대한 임계값 역시 계산해 놓고 있음

Critical Values for the Dickey-Fuller Test

Model	1%	5%	10%
$\Delta y_t = \gamma y_{t-1} + v_t$	–2.56	–1.94	–1.62
$\Delta y_t = \alpha_0 + \gamma y_{t-1} + v_t$	–3.43	–2.86	–2.57
$\Delta y_t = \alpha_0 + \alpha_1 t + \gamma y_{t-1} + v_t$	–3.96	–3.41	–3.13
Standard critical values	–2.33	–1.65	–1.28

4. 증대된(Augmented) Dickey-Fuller (ADF) 검정
   a. DF 검정을 위한 세가지 모형 설정 모두 오차항이 계열상관 되어 있지 않다는 가정이 전제됨
   b. 오차항의 계열상관 되어 있을 경우 이를 고려하기 위해 고안된 다음과 같은 모형 설정으로부터 이루어지는 단위근 검정을 ADF 검정이라고 함
      i. $\Delta y_t = \alpha_0 + \alpha_1 t + \gamma y_{t-1} + \sum_{i=1}^{m} a_i \Delta y_{t-i} + v_t$
      ii. 증대된 항의 개수(m)은 오차항의 계열 상관이 없어지기 충분한 정도로 결정함
      iii. γ =0 에 대한 ADF 검정은 DF 검정과 동일한 극한 분포를 가지며, 따라서 동일한 임계값을 사용함
5. DF 검정: 사례
   a. 실질개인소비지출 (y_t) (앞서의 거시경제변수 그래프들 중 (d))
      i. 그래프
         1. 강한 추세를 보이고 있으며, 비정상일 것이라고 의심됨
      ii. 상관도표
         1. 매우 완만한 감소를 나타냄
      iii. DF 검정(확률적 추세만 있는 경우, 확률적 추세와 비확률적 추세가 모두 있는 경우)

$$\Delta \hat{PCE}_t = -1.5144 + .0030 PCE_{t-1}$$

(tau)　　(-0.349) (2.557)

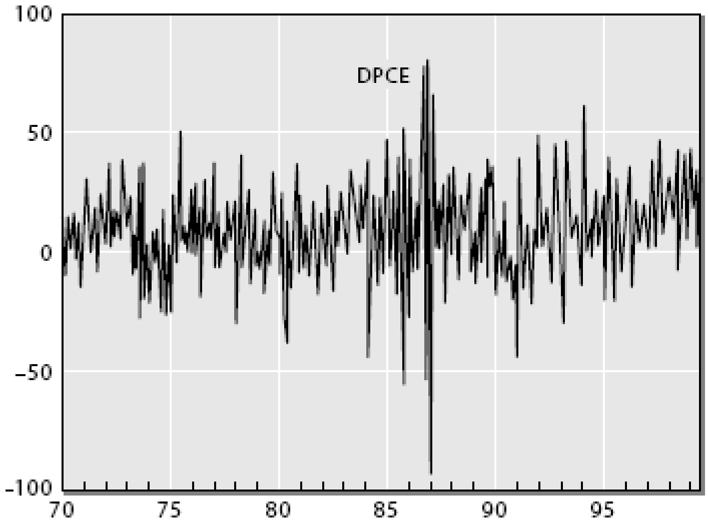
$$\Delta P\hat{C}E_t = 2.0239 + 0.0152t + 0.0013PCE_{t-1}$$
$$(\text{ tau}) \quad (0.1068) \ (0.1917) \ (0.1377)$$

iv. ADF 검정(확률적 추세만 있는 경우) (lag=2 로 선택해준 경우임)

$$\Delta P\hat{C}E_t = -2.111 + 0.00397PCE_{t-1} - 0.2503\Delta PCE_{t-1} - 0.0412\Delta PCE_{t-2}$$
$$(\text{ tau}) \quad (-0.4951) \ (3.3068) \quad (-4.6594) \quad (-0.7679)$$

v. 타우 검정통계량의 값이 모두 양수로 나타나고 있으며, 이는 실질개인소비지출이 단위근을 갖느다는 귀무가설을 기각하지 못함을 의미함

vi. 이제 문제는 실질개인소비지출의 일차차분 ($\Delta PCE_t = PCE_t - PCE_{t-1}$)이 안정적인가를 검정해볼 필요가 있음



vii. 순수 임의보행(추세가 없음)에 대한 DF 검정 결과는 다음과 같음

$$\Delta D\widehat{PCE}_t = -0.9969DPCE_{t-1}$$
$$(\text{ tau}) \quad (-18.668)$$

1. 타우 검정통계량의 값이 매우 큰 음의 값을 가지며 임계치로부터 판단할 때 1%유의 수준에서도 ΔPCE_t 가 단위근을 갖는다는 귀무가설을 기각함.
2. 즉 ΔPCE_t 는 안정적 시계열이며, 이로부터 판단할 때,

PCE_t는 I(1)시계열이라고 판단할 수 있음

6. Phillips-Perron (PP)단위근 검정
   a. DF 검정의 중요한 가정은 오차항이 독립적이며 동일한 분포를 한다는 것임
   b. ADF 검정은 설명변수에 시차를 갖는 차분값을 포함시킴으로써 계열상관의 문제를 고려하고 있음
   c. PP 검정은 시차를 갖는 차분값의 포함 없이 계열상관을 고려하는 방법을 제시하였으며, 그 극한 분포는 ADF 와 동일함
7. 단위근 검정에 대한 비판
   a. 이외에도 많은 단위근 검정 방법이 제시되어 옴
   b. 이들 검정통계량들의 사이즈와 검정력(power)에 문제가 있음
      i. 즉 제 1 형 오류의 확률(귀무가설이 옳은데 기각할 확률)와 제 2 형 오류를 범하지 않을 확률(즉 귀무가설이 옳지 않을 때 기각할 확률)에 문제
   c. 검정의 사이즈
      i. 모형 설정의 오류로 인해 사이즈의 왜곡이 발생함 (i 이냐 ii 냐, 혹은 MA 요소가 제외되어 있음으로 인한 왜곡 등
   d. 검정력
      i. DF 유형의 검정들은 대부분 낮은 검정력을 지님
         1. 즉 귀무가설이 옳지 않음에도 기각하지 못하는 경우가 많다는 것이고, 이는 단위근이 존재한다는 귀무가설을 정당한 경우보다 더욱 빈번히 받아들이게 되는 경향이 있음을 의미함
            A. 단위근 검정의 검정력은 시간적 기간에 의존함(관측치의 수보다는)
               i. 즉 30 년간의 30 개의 관측치가 100 일동안의 100 개의 관측치보다 더욱 더 높은 검정력을 제공할 수 있음
         2. $\rho \approx 1$ 이지만 정확히 1 은 아닌 경우, 단위근 검정은 단위근이 존재하는 것으로 판단하게 됨
         3. 시계열상에 구조적 단절이 있을 경우 이들 단위근 검정은 이를 포착하지 못할 수 있음

iv. 비정상시계열의 변환

1. $\Delta y_t = \gamma y_{t-1} + v_t$
   a. γ=0 의 경우, 즉 단위근이 존재하는 경우 1 계 차분은 정상 시계열임 (평균이 0)
2. $\Delta y_t = \alpha_0 + \gamma y_{t-1} + v_t$
   a. γ=0 의 경우, 즉 단위근이 존재하는 경우 1 계 차분은 역시 정상 시계열임 (평균이 0 이 아님)
3. $\Delta y_t = \alpha_0 + \alpha_1 t + \gamma y_{t-1} + v_t$
   a. γ=0 의 경우, 즉 단위근이 존재하는 경우 1 계 차분은 비확률추세를 따라 정상임)
   b. $\hat{v}_t = \Delta y_t - \hat{\alpha}_0 - \hat{\alpha}_1 t$: 즉 비확률추세에 대해 회귀해준 후 얻는 잔차는 추세가 제거된 시계열임(detrended time series)
   c. 만일에 비확률 추세가 선형이 아니라 이차식이라면
      i. $\Delta y_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + v_t$
      ii. $\hat{v}_t = \Delta y_t - \hat{\alpha}_0 - \hat{\alpha}_1 t - \hat{\alpha}_2 t^2$

v. 공적분 : 하나의 단위근 시계열을 다른 단위근 시계열에 회귀함

1. $y_t = PCE_t$, $x_t = PDI_t$,
   a. PDI_t 실질 개인 가처분 소득
      i. 그림에서 판단해 볼 수 있듯이 이 역시 단위근을 가지고 있음
   b. PCE_t , PDI_t, 모두 I(1)시계열임
2. $PCE_t = \beta_1 + \beta_2 PDI_t + \varepsilon_t$
   a. 단위근 시계열간의 회귀 → 허구적 회귀의 가능성
   b. 하지만 $\varepsilon_t = PCE_t - \beta_1 - \beta_2 PDI_t$ 에 대한 단위근 검정을 해보았을 때, 그 결과 I(0)인 것으로 나타난다면 이는 흥미로운 경우임
      i. 즉 PCE_t , PDI_t 이 확률적 추세를 가지고 움직임에도 불구하고 그 선형 결합($PCE_t - \beta_1 - \beta_2 PDI_t$)은 두 시계열에 존재하는 확률적 추세를 서로 상쇄 시킴
   c. 이 경우 위 회귀식은 의미를 가지며, 이 경우 두 변수(시계열)이 공적분되어 있다고 함 (공적분 회귀 공적분 모수(벡터))
      i. 경제학적으로 볼 때, 두 변수 사이에 장기적 균형 관계가 존재할 경우 공적분 됨
         1. 화폐수량설 : 물가수준과 통화공급량 (MV=PT)
         2. PPP 이론 : 환율과 물가수준의 비
      ii. 공적분 된 변수들 간에는 전통적 회귀분석 방법의 적용이

가능함

1. t 검정, F 검정 등

iii. 비정상시계열들간의 회귀분석시 공적분 여부에 대한 검정은 허구적 회귀를 피하기 위한 사전 검정으로 볼 수 있음

3. 공적분 검정(Testing for Cointegration)

a. 회귀식으로부터의 잔차에 대해 DF 또는 ADF 단위근 검정을 적용하는 방법 : (Engle-Granger(EG) 혹은 Augmented Engle-Granger(AEG) 검정이라고도 함)

i. PCE_t , PDI_t, 가 공적분되어 있지 않다면 어떠한 선형결합도 I(1)시계열이 될 것임

ii. 이 때 문제는 DF 또는 ADF 단위근 검정이 적용되는 잔차는 공적분모수에 대한 추정치를 기반으로 계산되며 따라서 이 경우 DF 의 임계치는 부적절하게 됨

1. Engle and Granger(1987) 이 임계치를 계산하였음

2. $$P\hat{C}E_t = -171.4412 + 0.9672 PDI_t$$
$$\text{(t-stats) } (-7.4808) \qquad (119.8712) \qquad R^2 = 0.9940 \quad d = 0.5316$$

A. DW 값보다 결정계수 값이 더 크며, 이는 회귀결과가 허구적일 가능성이 있음을 시사함

B. $$\Delta\hat{\varepsilon}_t = -0.2753\hat{\varepsilon}_{t-1}$$
$$\text{(tau) } (-3.7791)$$

i. Engle and Granger 의 1% 임계치는 -2.5899 이며, 따라서 잔차가 I(0)라고 결론내릴 수 있음

C. 따라서 이 경우 위 회귀식은 공적분회귀이며 회귀결과는 허구적이지 않음

i. 이 식은 정태적 혹은 장기 소비 함수로 해석되며, 0.9672 는 장기적 혹은 균형 한계소비성향을 나타냄

b. 공적분회귀 DW (CRDW) 검정

i. 공적분회귀로부터의 DW 값을 이용함

1. 귀무가설인 d=0 임 (1 계 자기상관에 대한 검정시 귀무가설은 d=2 였음)

A. $d \approx 2(1-\hat{\rho})$

B. 오차항에 단위근이 존재한다면 ρ=1 이며, 이는 d 가 0 에 가까운 값임을 의미함

2. Sargan and Bhargava(1983)이 이 경우의 임계치를 계산함
   A. 1%, 5%, 10%에 대한 임계치가 0.511, 0.386, 0.322 임
   B. 즉 이 임계치 보다 큰 경우 d=0 이라는 귀무가설을 기각하게 됨 (즉 공적분 되어 있지 않다는 귀무가설을 기각하게 됨)
   C. 위 식에서 d=0.5316 이며 1%에서도 공적분되어 있지 않다는 귀무가설은 기각됨(즉 공적분되어 있음)

vi. 오차수정모형 or 기제(Error Correction Model or Mechanism : ECM)
1. 위 예에서 PCE 와 PDI 는 공적분되어 있음을 보였고, 이는 PCE 와 PDI 가 장기적 균형관계에 있음을 의미함
2. 단기적으로 PCE 와 PDI 는 불균형 상태에 있을 수 있으며, 공적분회귀식에서의 오차항을 균형오차로 취급할 수 있음
   a. 이 오차항을 PCE 의 단기적 행태를 그 장기적 값과 연결하는 데 사용할 수 있음 → 오차수정모형
3. Granger representation theorem : 두변수 X 와 Y 가 공적분 관계에 있을 경우, 둘 사이의 관계를 ECM 으로 표현할 수 있음
4. ECM: $\Delta PCE_t = \alpha_1 + \alpha_2 \Delta PDI_t + \alpha_3 \varepsilon_{t-1} + \eta_t$
   a. $\varepsilon_{t-1} = PCE_{t-1} - \beta_1 - \beta_2 PDI_{t-1}$: 공적분 회귀로부터의 오차항의 1 기 과거 값
   b. t-1 기에서 t 기 사이의 PCE 의 변화가 같은 기간의 PDI 의 변화에 대한 즉시적인 조정과 t-1 기의 균형오차에 대한 조정을 포함하고 있음
      i. 균형오차항이 0 이 아닐 경우 모형은 균형을 벗어나 있음을 의미함
      ii. t-1 기의 균형오차가 +라면 t-1 기의 PCE 가 균형에 비해 너무 높은 수준임을 의미함
      iii. α_3(오차수정항)은 - 일 것으로 기대되며, 이 경우 ΔPDI 가 0 이라면 ΔPCE 는 - 가 됨
         1. 즉 전기의 PCE 가 균형보다 높은 수준이었다면, 다음기에는 그러한 균형오차를 보정하기 위해 하락함을 의미
      iv. α_3 의 절대값은 균형이 얼마나 빨리 회복되는가 하는 정도를 나타냄
   c. 실제로는 $\hat{\varepsilon}_{t-1} = PCE_{t-1} - \hat{\beta}_1 - \hat{\beta}_2 PDI_{t-1}$ 를 으로 사용함
5. 실제추정치:

$$\Delta P\hat{C}E_t = 11.6918 + 0.2906\Delta DPI_t - 0.0867\hat{\varepsilon}_{t-1}$$

a. (t-stats) (5.3249) (4.1717) (-1.6003)

$$R^2 = 0.1717 \quad d = 1.9233$$

b. 오차수정항의 통계적 유의성은 매우 취약함(10% 유의수준에서 0 이라는 것을 기각할 수 없는 수준)

i. 즉 PCE 는 PDI 의 변화에 대해 같은 기간 내에 적응함을 의미

ii. 개인가처분소득의 단기적 변화(ΔPCE)는 실질소비지출에 +의 통계적으로 유의한 영향을 미치고 있으며, 이를 단기 한계소비성향으로 해석할 수 있음 (장기 혹은 정태적 한계소비성향은 0.9672)