CH7. 비정상 ARMA

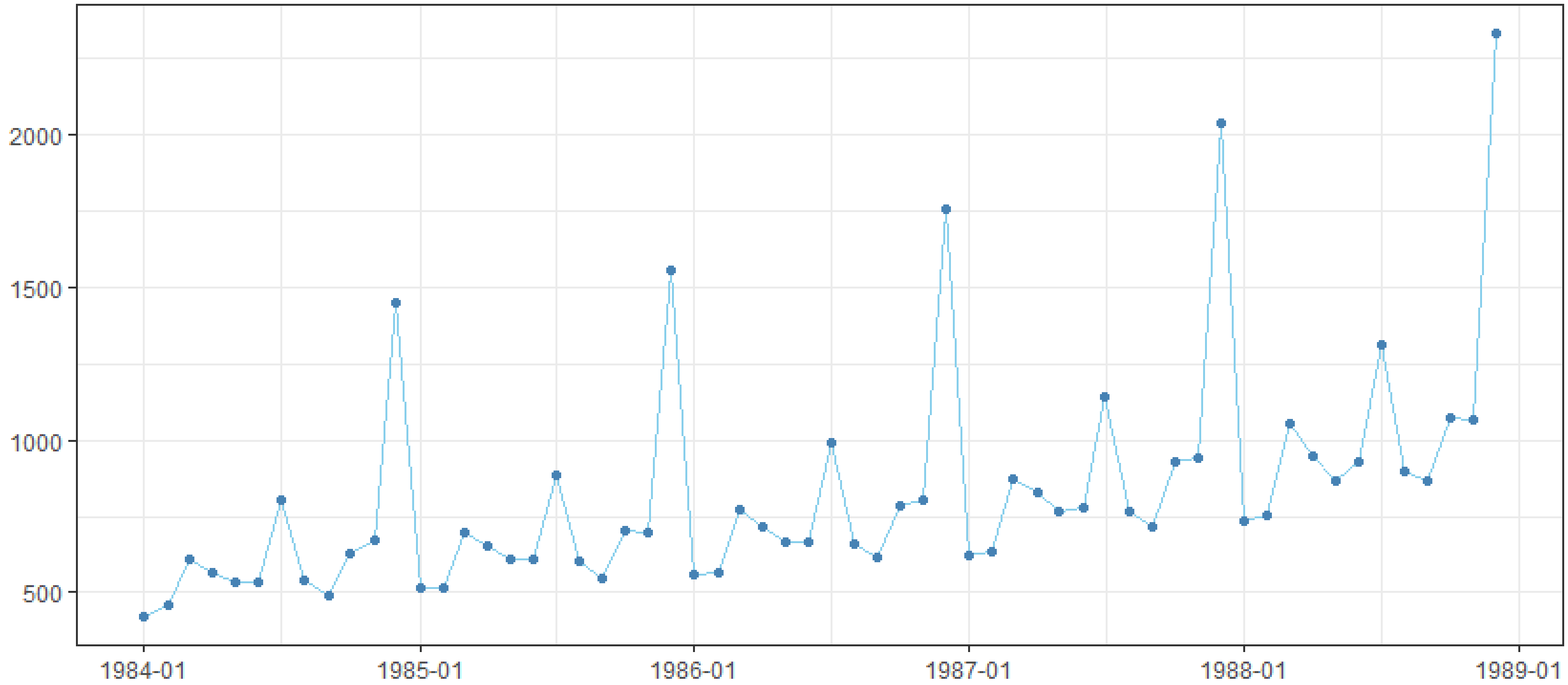
비정상 시계열의 정상화

- **정상시계열이란**
  - 평균이 관측 시점에 관계없이 일정
  - 분산이 존재하며, 관측 시점에 관계없이 일정.
  - 자기 공분산이 관측 시점과는 무관하고, 시차에만 의존.

- **비정상 시계열(nonstationary time series)의 시계열 그림 특징**
  - 시계열의 수준이 시간대에 따라 다르다.
  - 시계열이 추세를 가진다.
  - 시계열이 계절성을 보인다.
  - 시계열의 분산이 시간대에 따라 변한다.

비정상 시계열의 정상화

- **백화점 월별 매출액**



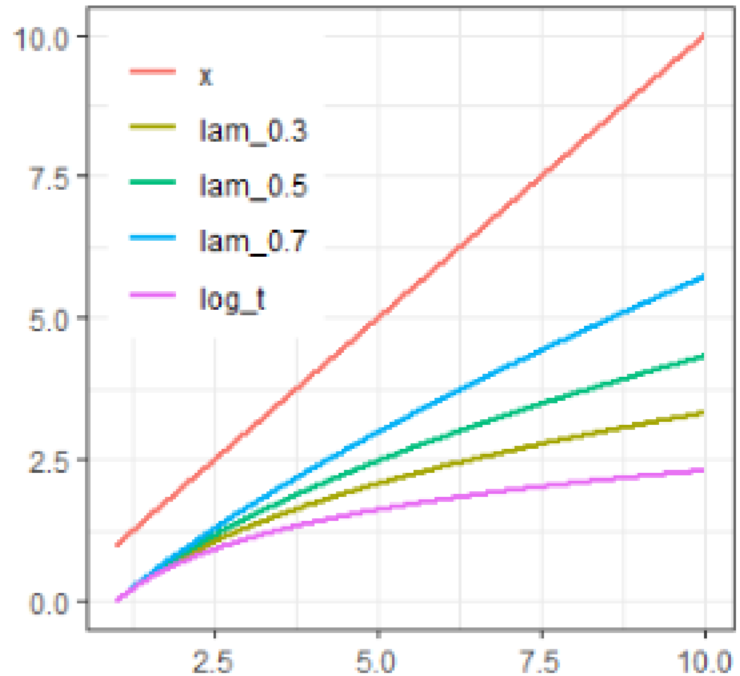
비정상 시계열의 정상화

- **분산이 일정하지 않은 경우**
  - 시계열 그림을 통해 확인 가능
  - 시간의 흐름에 따라 시계열의 값이 커지면서 변동 폭이 커지는 경향
  - 대부분의 비정상 요인 중에서 가장 먼저 시정해야 함
  - 분산 안정화 변환(variance stabilizing transformation) 시행
  - Box-Cox 변환, 로그 변환이나 제곱근 변환 등의 변환을 많이 시도
  - 앞부분의 작은 분산은 늘려주고 뒷부분의 큰 분산은 줄여주는 효과
  - 모두 양수여야 한다는 제약 조건이 필요함

비정상 시계열의 정상화

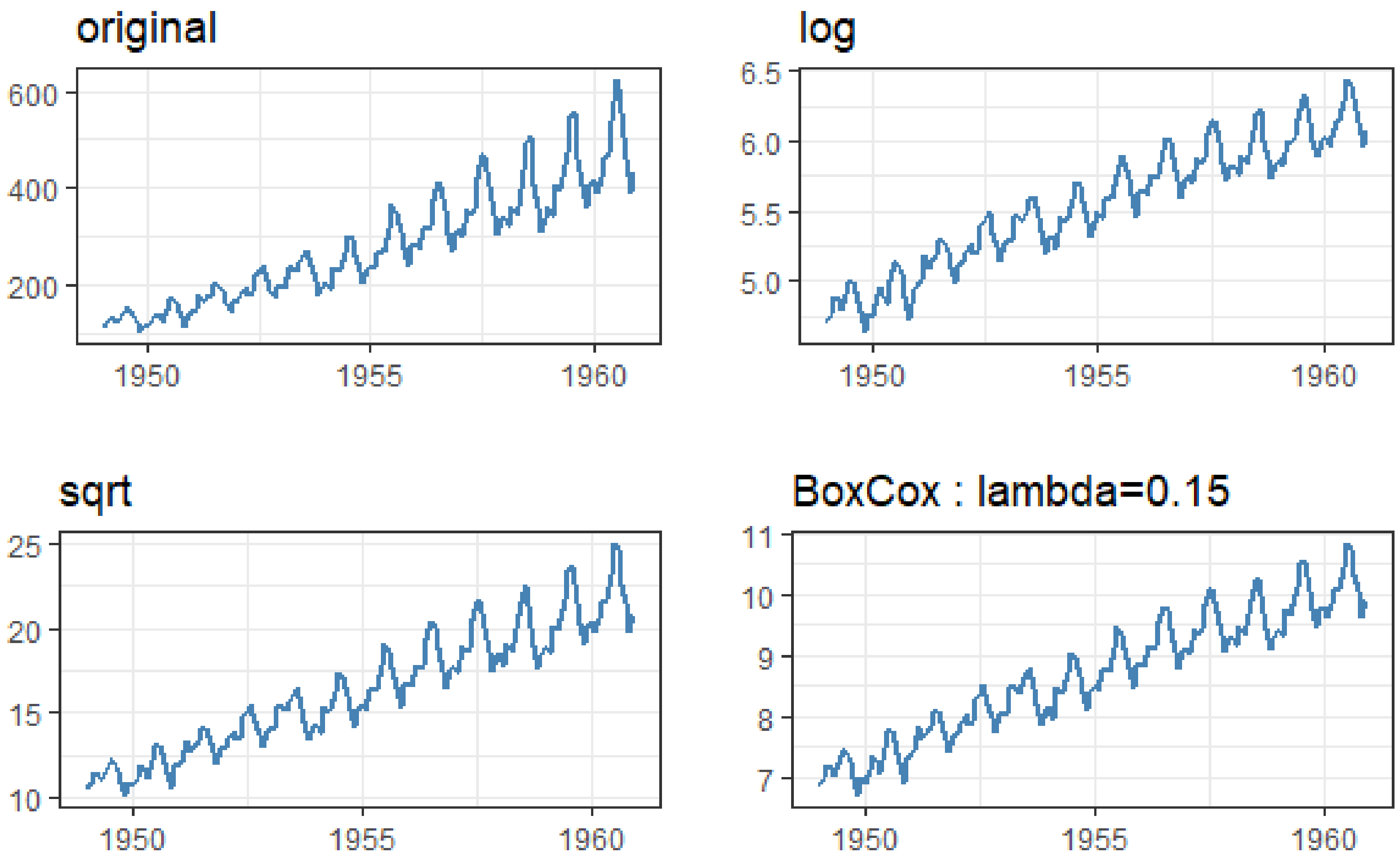
- **Box-Cox Transformation**

  - $$f_\lambda(Z_t) = \begin{cases} \frac{Z_t^\lambda - 1}{\lambda}, & Z_t \geq 0, \lambda > 0 \\ \log(Z_t), & \lambda = 0 \end{cases}$$



비정상 시계열의 정상화

- 분산이 일정하지 않은 경우



비정상 시계열의 정상화

- **수준이 일정하지 않은 경우 – 결정적 추세**

  - deterministic trend
  - 추세가 결정적이고 동시에 영원히 지속되는 경우
  - 시계열 그림을 통해 확인 가능
  - 분해법을 이용하여 추세 제거 : $Z_t = T_t + S_t + I_t$
    - $\hat{I}_t = Z_t - \hat{T}_t - \hat{S}_t$
    - $\hat{T}_t = \hat{\beta}_0 + \hat{\beta}_1 t + \cdots + \hat{\beta}_k t^k$
    - $\hat{S}_t = \sum_{i=1}^{s} \hat{\delta}_i \times IND_{ti}$

비정상 시계열의 정상화

- **수준이 일정하지 않은 경우 – 확률적 추세**

  - stochastic trend
  - 인접자료들 간에 존재하는 강한 양의 상관관계 때문에 어떤 추세가 있는 것처럼 보이거나 수준이 일정하지 않은 것처럼 보이는 경우
  - 시계열 그림만으로 확인하기 어려우며 SACF(매우 느리게 감소)를 함께 그려 확인
  - 차분법을 이용하여 정상화

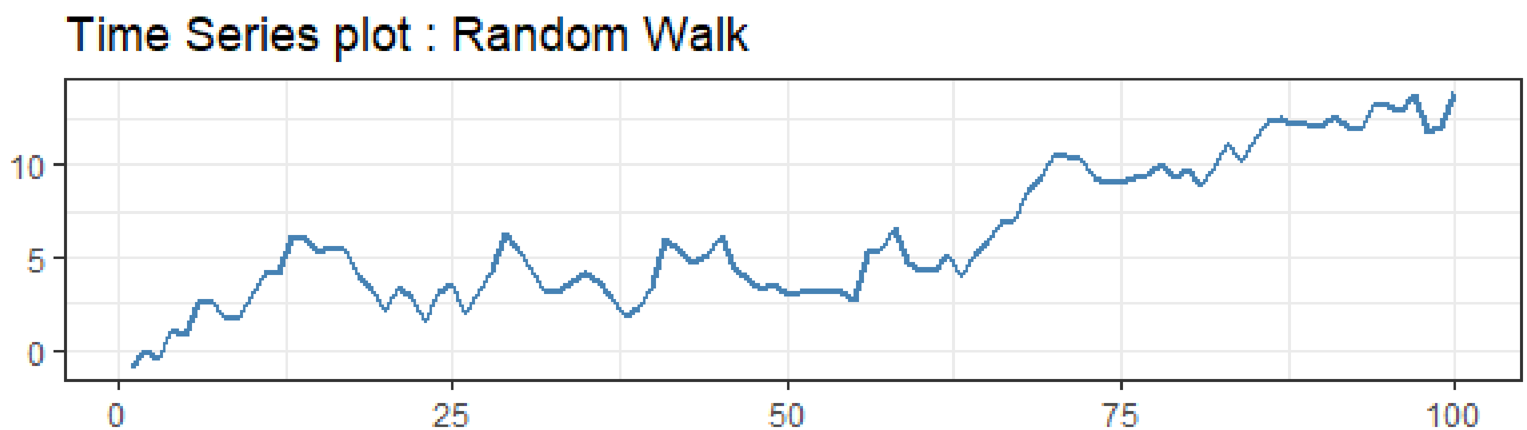
비정상 시계열의 정상화

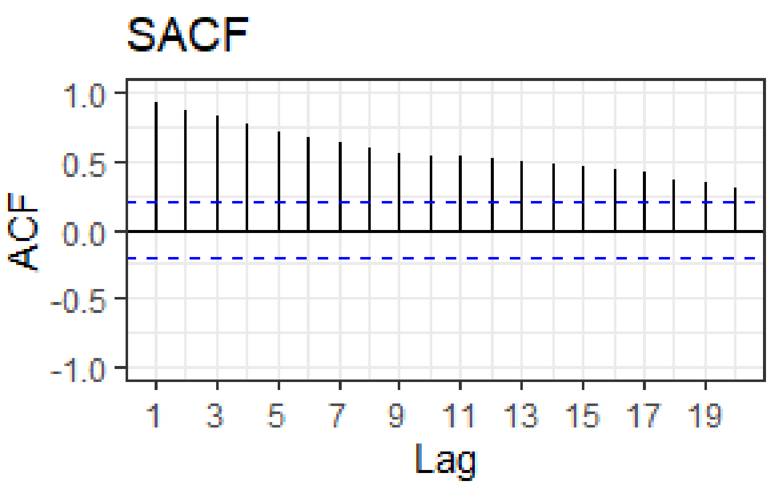
- **차분 연산자 (differencing operator)**

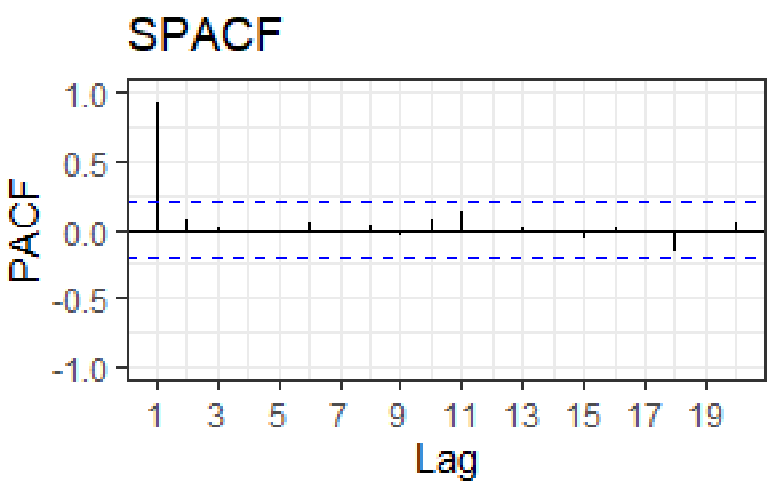
  - $\nabla = 1 - B$
    - $\nabla Z_t = (1-B)Z_t = Z_t - Z_{t-1}$
    - $\nabla^2 Z_t = (1-B)^2 Z_t = (1 - 2B + B^2)Z_t = Z_t - 2Z_{t-1} + Z_{t-2}$
  - 확률보행과정 정상화
    - $\{Z_t\}$: random walk => $Z_t = Z_{t-1} + \varepsilon_t,\ \varepsilon_t \sim_{iid} (0, \sigma^2)$
    - $\nabla Z_t = Z_t - Z_{t-1} = \varepsilon_t$ 이므로 $\nabla Z_t \sim WN(0, \sigma^2)$

비정상 시계열의 정상화

■ **확률보행 과정**

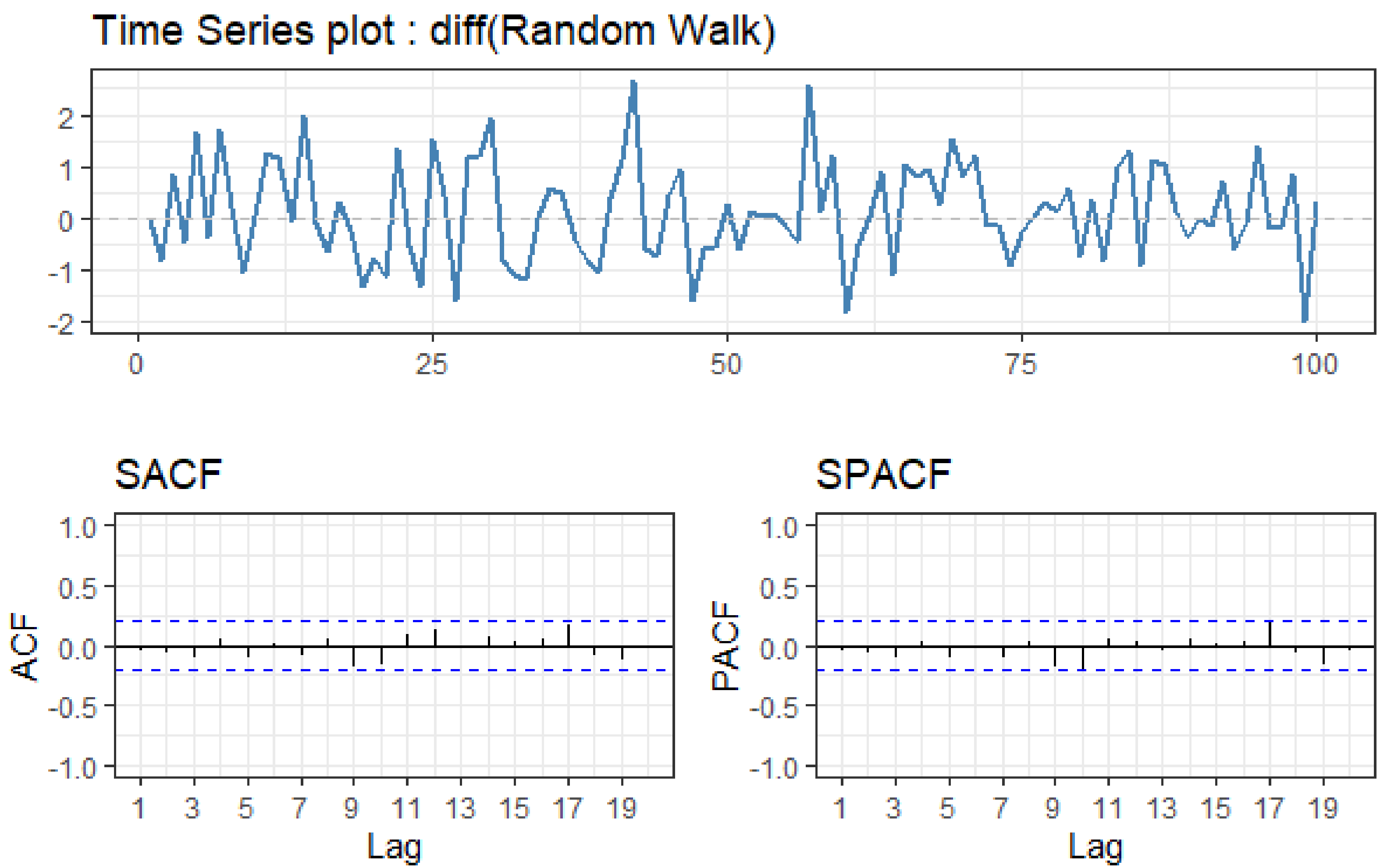






비정상 시계열의 정상화

- 확률보행 과정



비정상 시계열의 정상화

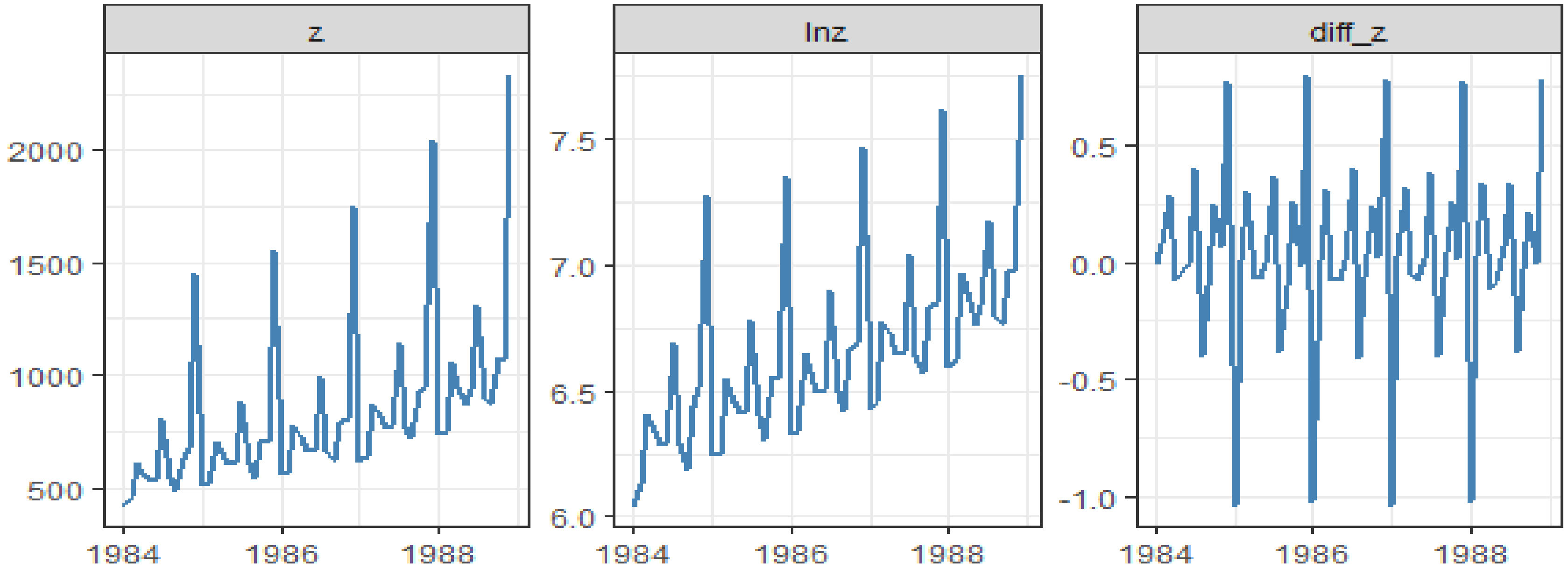
- **차분을 통한 결정적 추세 제거**
  - $Z_t = \beta_0 + \beta_1 t + \varepsilon_t,\ \varepsilon_t \sim WN(0, \sigma^2)$
  - $\nabla Z_t = Z_t - Z_{t-1} = \beta_0 + \beta_1 t + \varepsilon_t - \{\beta_0 + \beta_1(t-1) + \varepsilon_{t-1}\} = \beta_1 + \varepsilon_t - \varepsilon_{t-1}$

    $$\Longrightarrow \{\nabla Z_t\} \sim MA(1)$$
  - 일반적으로 p차 다항 추세를 갖는 비정상 시계열의 경우 p차 차분 을 시행
  - 그러나 결정적 다항 추세의 경우 차분을 통해 정상화한 시계열이 분해법을 통한 것보다 모형이 더 복잡해지므로 차분법보다는 분해법을 선호함

비정상 시계열의 정상화

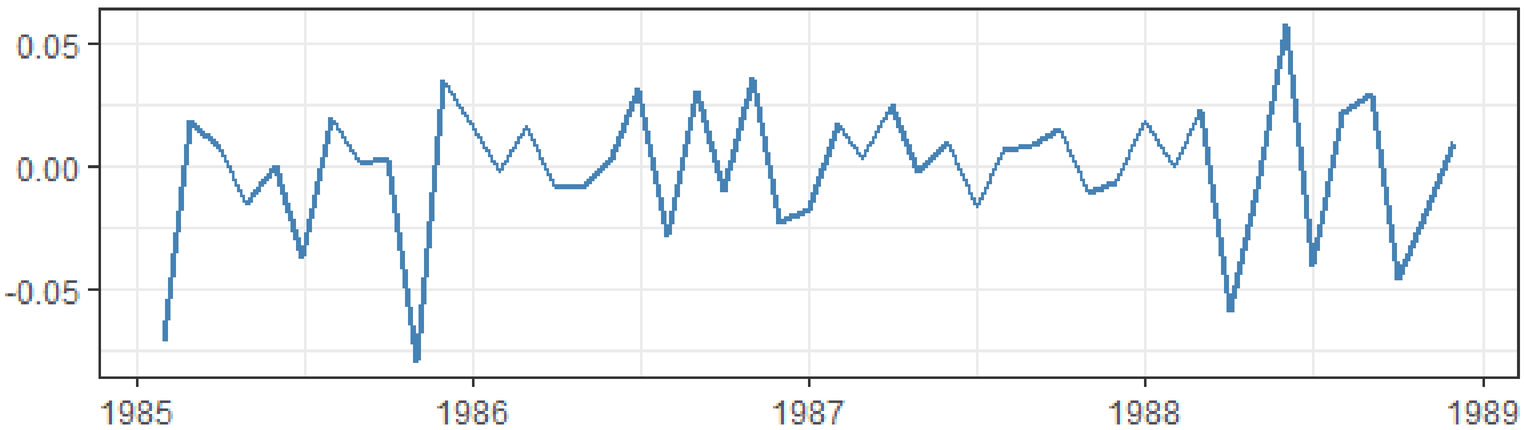
- **계절성분**
  - 결정적 계절 추세의 경우 분해법 이용
  - 확률적 계절 추세의 경우 계절 차분이용

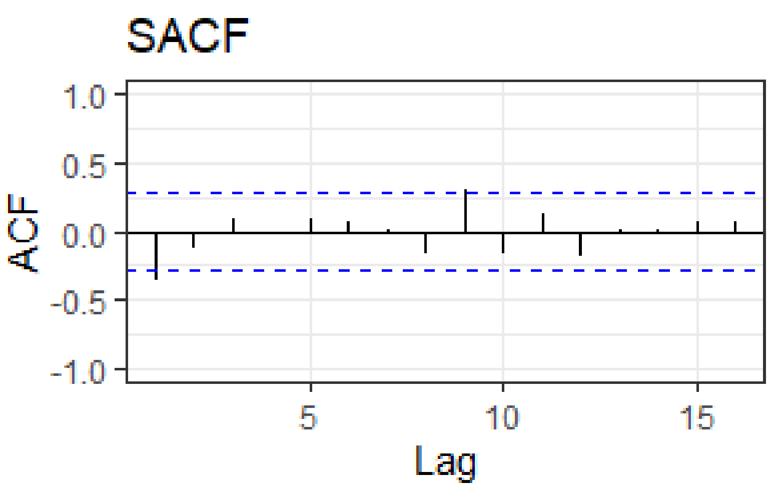
$$\nabla_s = 1 - B^s$$

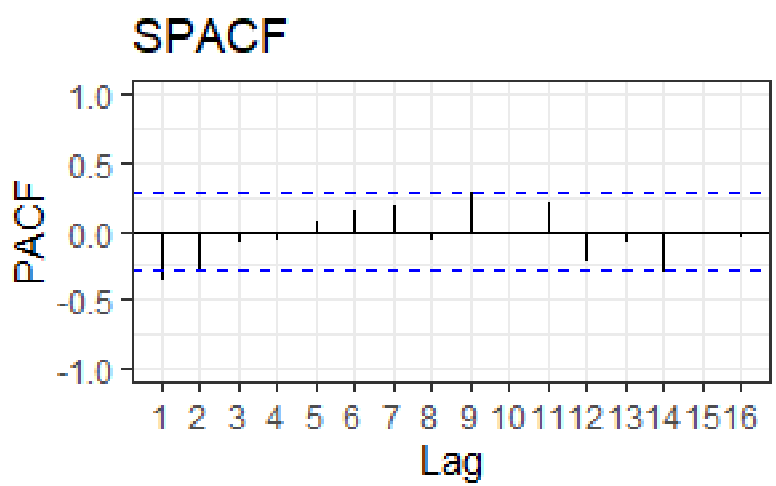


비정상 시계열의 정상화

- **계절성분 : 로그변환 -> 차분 -> 계절차분**







비정상 자기회귀-이동평균 과정

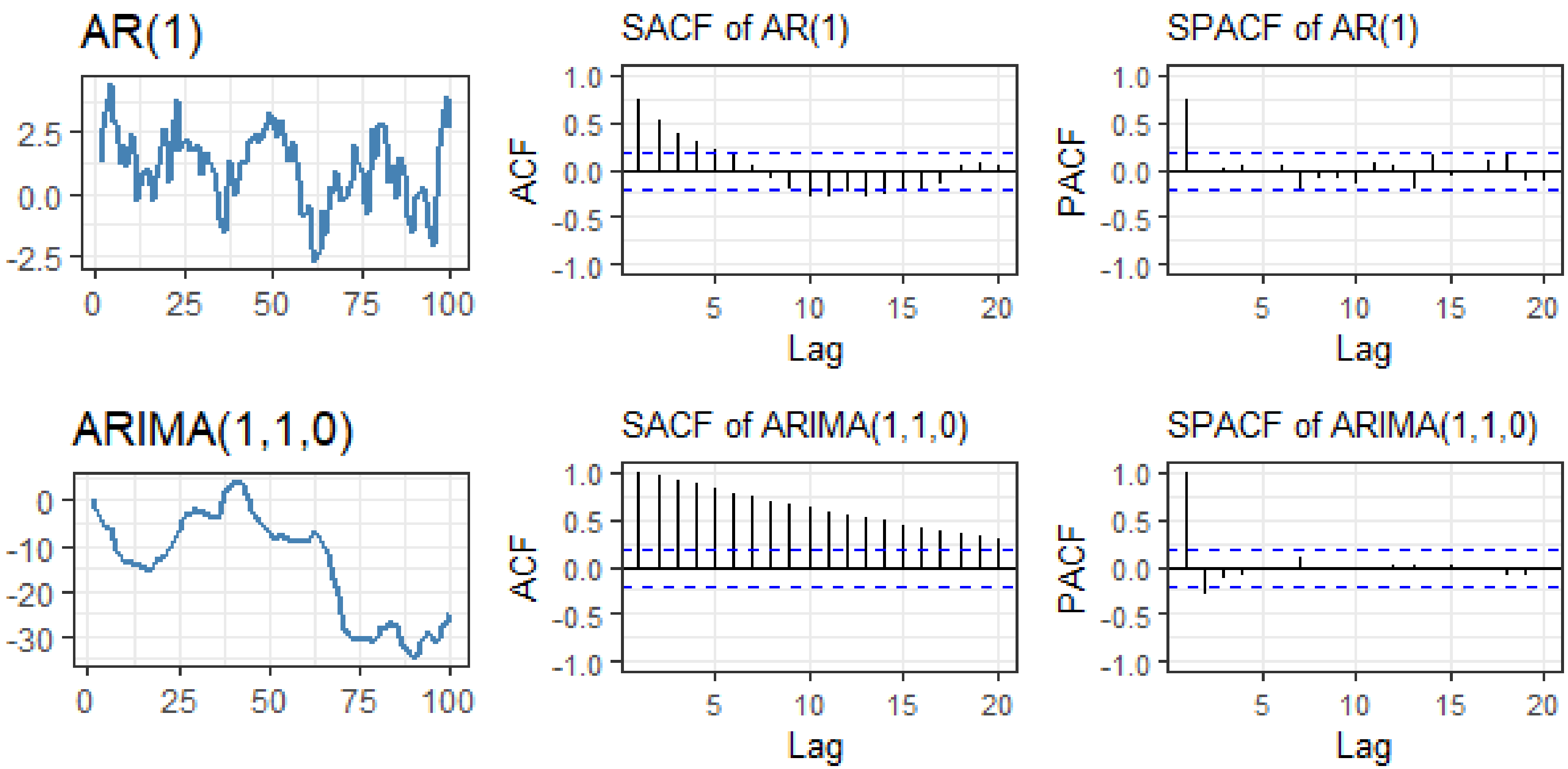
- **ARIMA (Autoregressive Integrated Moving Average) model**
  - $\{Z_t\} \sim ARIMA(p,d,q) \Leftrightarrow$

    $\{\nabla^d Z_t\} \sim ARMA(p,q)$: stationary ARMA model
  - $d = 0 \Rightarrow \{Z_t\} \sim ARMA(p,q)$
  - $d \geq 1 \Rightarrow d$차 polynomial trend 제거
  - $ARIMA(p,d,q)$: 비정상 시계열
  - random walk와 같은 비정상 AR(1) model은 AR(1)으로 표현하기보다 주로 ARIMA(0,1,0)로 표현
  - ARIMA(0,d,0), ARI(p,d), IMA(d,q) models

비정상 자기회귀-이동평균 과정

- **ARIMA(1,1,0)**

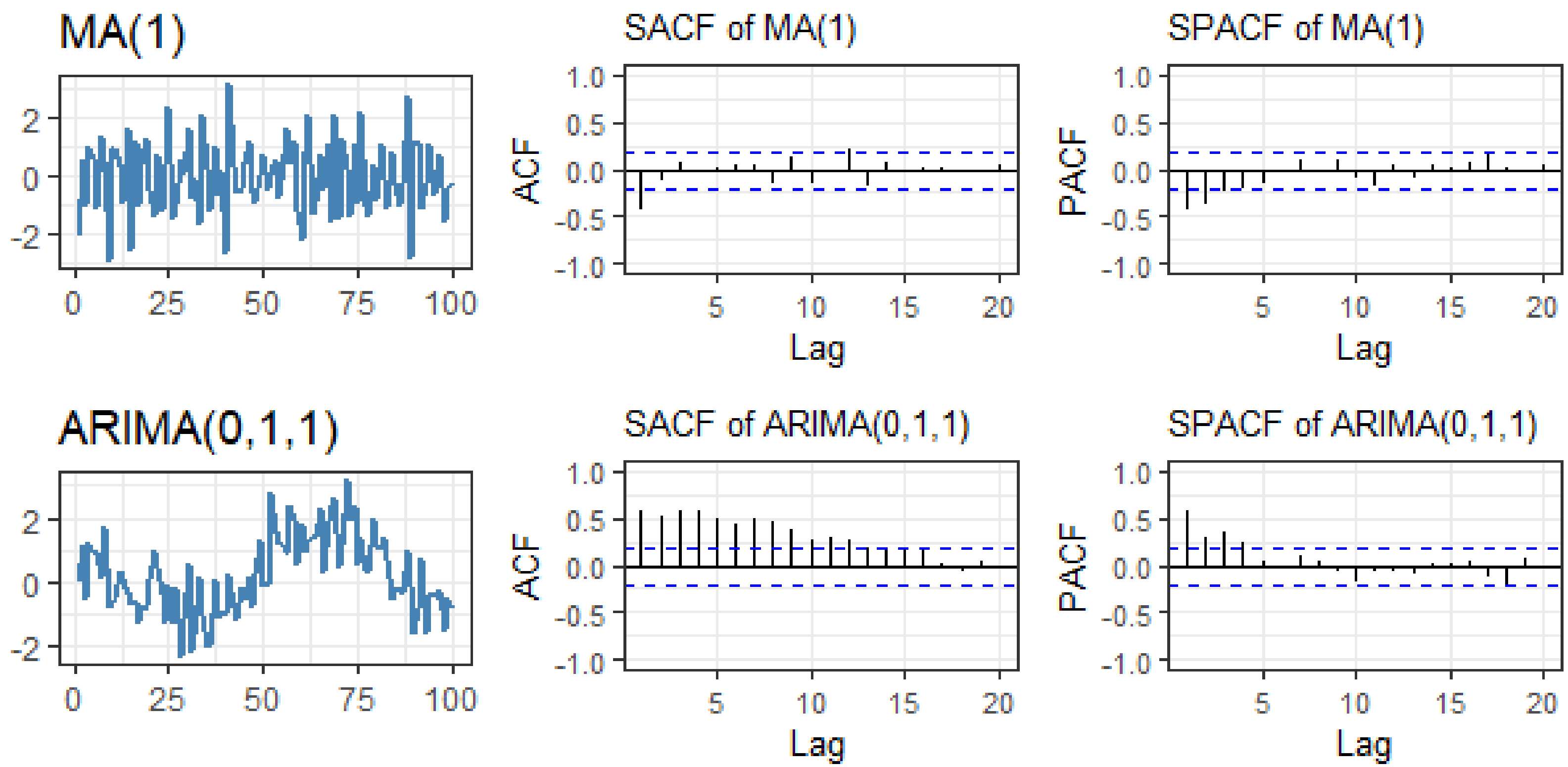
  - $(1-0.8B)Z_t = \varepsilon_t, \;\; (1-0.8B)(1-B)Z_t = \varepsilon_t, \quad \varepsilon_t \sim N(0,1)$



비정상 자기회귀-이동평균 과정

- **ARIMA(0,1,1)**

  - $Z_t = (1 - 0.8B)\varepsilon_t, (1 - B)Z_t = (1 - 0.8B)\varepsilon_t, \quad \varepsilon_t \sim N(0,1)$



과대차분

- **과대차분 (Overdifferencing)**
  - 이미 정상화가 된 시계열을 차분하는 것을 말함
  - MA(1) 의 과대차분

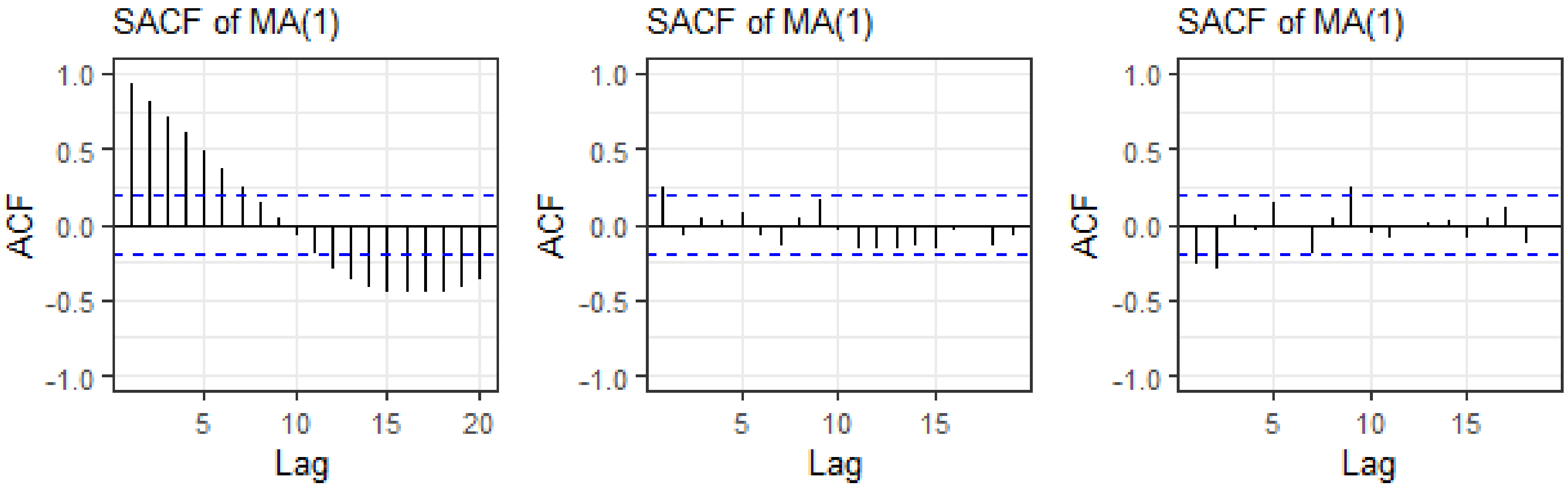
$$Z_t = \varepsilon_t - \theta\varepsilon_{t-1}, \varepsilon_t \sim WN(0, \sigma^2)$$

$$\nabla Z_t = \varepsilon_t - \theta\varepsilon_{t-1} - (\varepsilon_{t-1} - \theta\varepsilon_{t-2}) = \varepsilon_t - (1+\theta)\varepsilon_{t-1} + \theta\varepsilon_{t-2} \Rightarrow \text{MA(2)}$$

  - 과대차분을 하는 경우 정상성에는 큰 문제가 없으나 모형이 복잡해지거나 분산이 크게 만드는 수가 있다.

과대차분

- **과대차분 (Overdifferencing)**
  - $(1-B)Z_t = (1-0.8B)\varepsilon_t, \nabla Z_t, \nabla^2 Z_t$



End of Document