

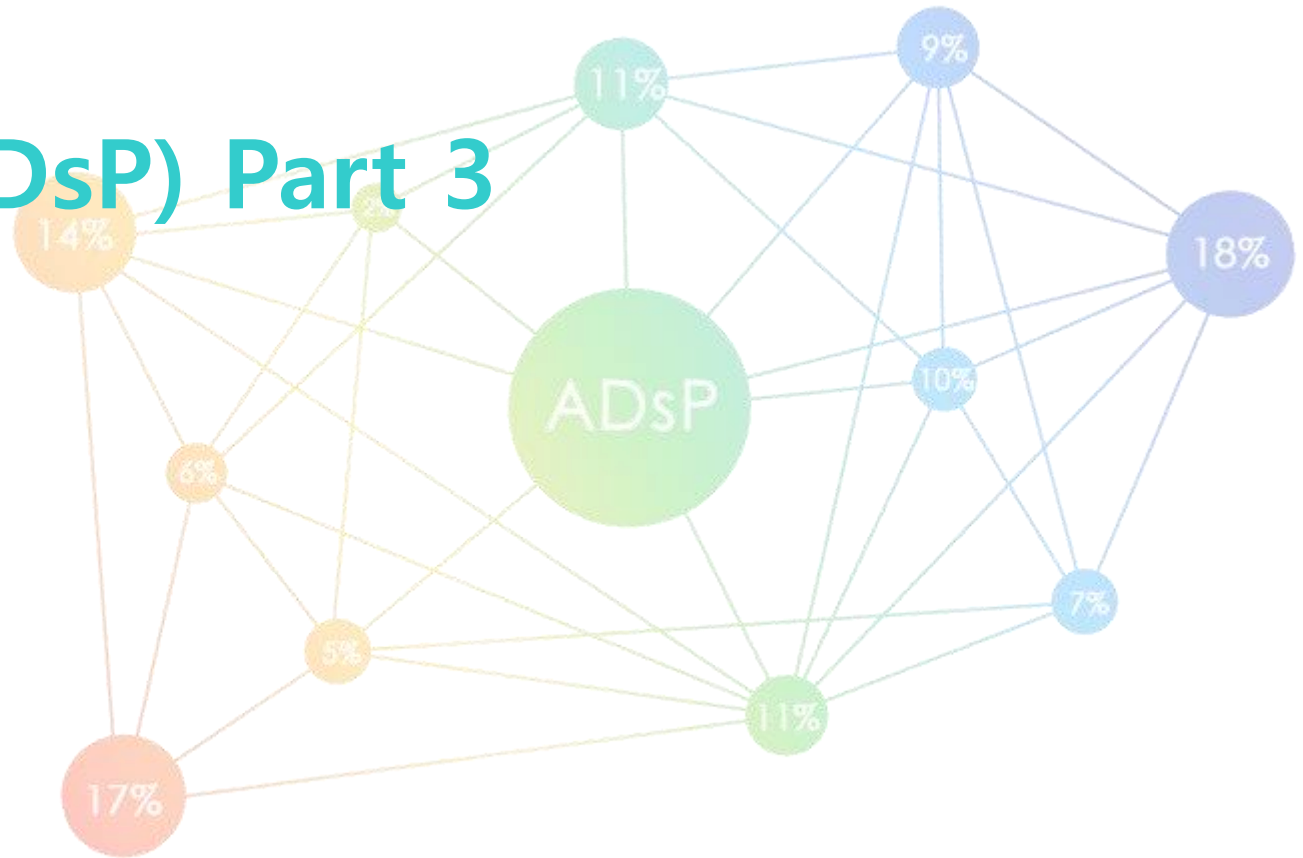
## 데이터분석준전문가(ADsP) Part 3

# 데이터분석

02

## 제2장 통계분석 제3절 다변량분석

1. 상관분석
2. 다차원척도법
3. 주성분분석



# 1. 상관분석



상관분석은 데이터 안의 두 변수 간의 관련성을 파악하는 방법이다.  
상관계수는 두 변수 선형성의 정도를 의미하며, 이를 측정하는 방법에는 피어슨 상관계수, 스피어만 상관 계수, 켄달의 순위상관계수 등이 있다.  
그러한 흔히 상관계수라고 하면 피어슨 선형 상관계수를 뜻한다.

Ex) 상관분석은 두 변수의 선형관계만을 의미,  
인과관계를 의미하는 것은 아니다.

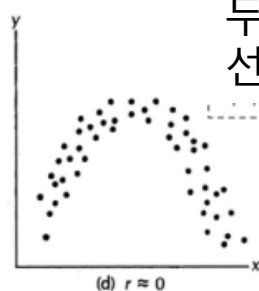
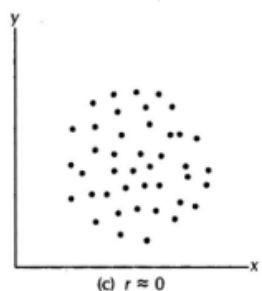
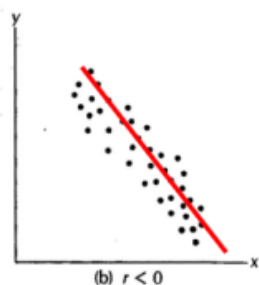
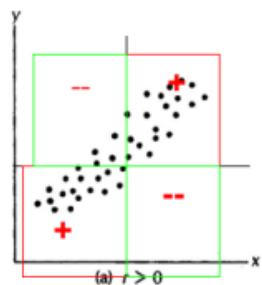
# 1. 상관분석



## (1) 상관계수(Correlation)

중심의 위치나 흩어짐의 정도에 평균이나 표준편차를 이용하듯이 두 변수의 상관의 정도를 수치로 표현한 것이 “상관계수”

$$\text{상관계수}(r) = \text{Cov}_{(x,y)} / S_x S_y = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X)} \sqrt{\text{var}(Y)}}$$



두 변수간의 관계는 2차식 이나  
선형관계 정도를 측정하는 상관계수는 0이다.

# 1. 상관분석



## (2) 상관계수 범위

상관계수  $r$ 의 범위는  $-1 \leq r \leq +1$ 이다.

상관계수가 1에 가까울수록 상관이 높다 하고, 0에 접근할수록 상관이 낮다고 말한다.

$r=0$ 은 두 변수 사이에 직선적 상관관계가 없음을 의미한 가지

유의할 점은  $r=0$ 이 두 변수 간에 직선적(선형) 관계가 없음을 의미하며,

→ 곡선관계에서는 (상관계수가= 0)

두 변수 사이에 어떤 관계도 존재하지 않는다는 뜻이 아니라는 것이다.

따라서 상관분석을 할 때는 먼저 산점도를 그려서 두 변수 간의 관계를 미리 알아보는 것이 중요하다.

# 1. 상관분석



## (3) 공분산(Covariance)

$$\text{Cov}(x, y) = E[(x - \mu_1)(y - \mu_2)]$$

$$= E(XY) - E(X)E(Y)$$

\* X, Y가 독립이라면 공분산은 0

$$E(XY) = E(X)E(Y)$$

두 확률변수가 얼마나 같이 변하는지를 측정한다.

한 변수가 커질 때 다른 변수가 함께 증가하거나 한 변수가 작아질 때 다른 변수도 함께 작아지는 것과 같이 크기 변화의 방향이 같다면 공분산은 양의 값을 가진다.

→ 공분산 측정단위에 따라 영향을 받기 쉽다.

→ 상관계수 사용

# 1. 상관분석



## (4)상관계수 결측치 처리하기

`cor(x,y,①use="complete.obs",②use="pairwise.complete.obs",③use="everything")`

- ① NA값이 있으면 해당 관측치 제외하고 계산
- ② NA값이 포함된 벡터에서만 제외하고 계산
- ③ NA값 포함해서 값을 구한다.

# 1.상관분석



## (5)변수와 상관계수와의 관계

두 변수가 독립이면 상관계수가 0이지만

상관계수가 0이라고 해서 두 변수가 독립은 아니다.

```
> x<-seq(-1,1,by=0.05) #x -1~1
```

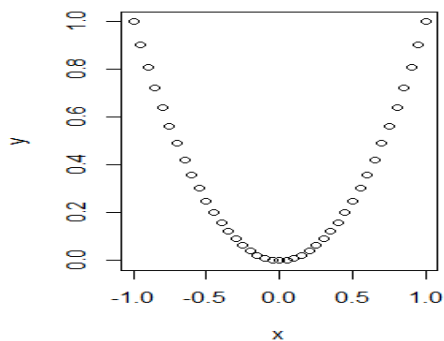
```
> y<-x^2 #y=x2
```

```
> options(scipen = 999)
```

```
> cor(x,y)
```

```
[1] 0.000000000000000001775033 # x와 y사이의 상관계수는 0에 근사함
```

```
> plot(x,y)
```



# 1. 상관분석



## (6)상관계수의 유의성 검정

`cor.test()` 함수를 사용해 상관계수 검정을 수행하여 상관계수의 유의성 검정을 판단할 수 있다.

이 때 귀무가설은 "상관계수가 0이다",  
대립가설은 "상관계수가 0이 아니다"이다.



# 1. 상관분석



## (8)스피어만 상관계수

스피어만 상관계수는 상관계수를 계산할 때 두 데이터의 실제값 대신 두 값의 순위 사용해 상관계수 계산한다.

또한 비선형관계의 연관성 파악이 가능, 이산형, 순서형 데이터에 적용이 가능하다.

Q03

스피어만의 상관계수에 대한 설명 중 옳바르지 않은 것은?

- ① 두 변수 간의 비선형적인 관계는 나타내지 못한다.
- ② 연속형외에 이산형도 가능하다.
- ③ 관계가 랜덤이거나 존재하지 않을 경우 상관 계수 모두 0에 가깝다.
- ④ 스피어만 상관 계수는 원시 데이터가 아니라 각 변수에 대해 순위를 매긴 값을 기반으로 한다.

Q03

스피어만의 상관계수에 대한 설명 중 옳바르지 않은 것은?

정답

- ① 두 변수 간의 비선형적인 관계는 나타내지 못한다.
- ② 연속형외에 이산형도 가능하다.
- ③ 관계가 랜덤이거나 존재하지 않을 경우 상관 계수 모두 0에 가깝다.
- ④ 스피어만 상관 계수는 원시 데이터가 아니라 각 변수에 대해 순위를 매긴 값을 기반으로 한다.

## 2. 다차원 척도법 (MDS, Multidimensional Scaling)



다차원 척도법은 개체들 사이의 유사성/비유사성을 측정하여 2차원 또는 3차원 공간상에 표현하는 분석 방법으로 개체들 간의 근접성을 시각화하여 데이터 속에 잠재해 있는 패턴이나 구조를 찾아내는 통계기법이다.

→ 차원의 수가 많을수록 추정의 적합도가 높아지지만 해석이 어려워지므로 일반적으로 두 개 차원으로 시각화  
다차원척도법에서의 유사성 계산은 Euclidean 거리

자료들의 상대적 관계를 직관적으로 이해하는 방법

# 3. 주성분 분석



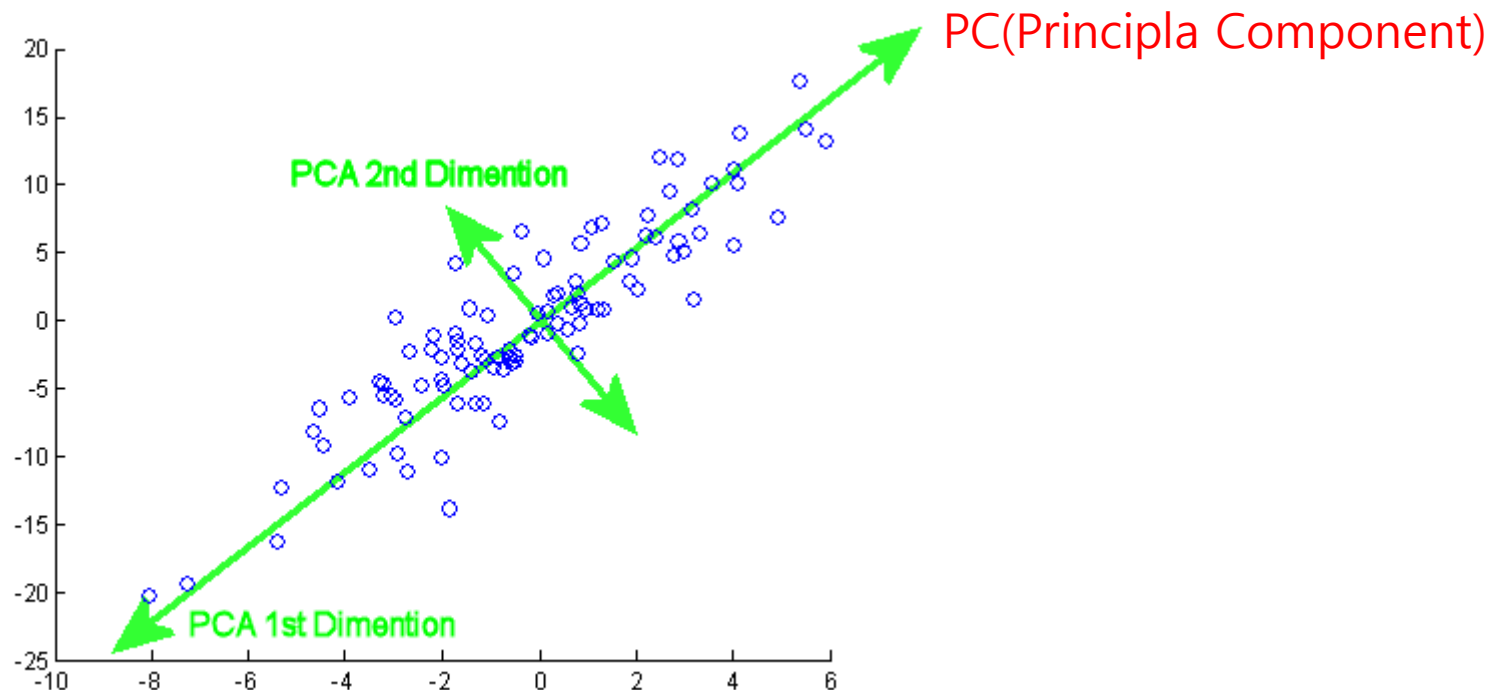
## (1) 주성분 분석

주성분 분석은 상관관계가 있는 고차원의 자료를 자료의 변동을 최대한 보존하는 저차원 자료로 변환하는 차원 감소(Dimensionality Reduction) 기법

첫 번째 주성분으로 전체 변동을 가장 많이 설명할 수 있도록 하고,  
두 번째 주성분으로는 첫 번째 주성분과는 상관성이 낮아서 첫 번째 주성분이 설명하지 못하는 나머지 변동을 정보의 손실 없이 가장 많이 설명할 수 있도록 변수들의 선형조합을 만들게 됩니다.

→ 주성분 간에는 상관계수 0이면서 각 주성분은 분산을 최대화

# 3.주성분분석



회귀분석이 독립변수에 따른 종속변수와의 거리인 잔차 최소화,  
주성분분석 독립변수들과 주성분과의 거리인 정보손실량 최소화, 분산최대화

# 3. 주성분분석



## (2) 주성분분석을 하는 이유

- ① 속성이 많으므로 2~3개로 속성을 줄여서  
주성분으로 만들어 포괄적으로 봄으로써  
데이터를 쉽게 이해하기 원하기 때문
- ② 다중공선성이 존재할 경우 해결 방법 중의 하나가 바로 상관도가  
높은 변수들을 하나의 주성분 혹은 요인으로 축소하여  
모형 개발에 활용
- ③ 회귀분석에서 설명변수의 개수 결정,  
군집 분석의 사전 분석에 활용

# 3. 주성분분석



## (3) 공분산 행렬 VS 상관행렬

주성분분석의 문제는 척도에 영향을 받는다는 점이다.  
변수들의 선형결합을 유도할 때 분산을 이용하기 때문

- 공분산행렬(covariance matrix)과 상관계수행렬(correlation matrix)의  
고유값과 고유벡터가 일반적으로 동일하지 않을 수 있으므로  
주성분이 달라질 수 있음



# 3.주성분분석



## (4)주성분 분석 해석하기

```
fit<-prcomp(USArrests,scale=TRUE)# scale=TRUE 표준화
```

```
# princomp() 상관계수 행렬 옵션 cor=TRUE
```

```
summary(fit)
```

Importance of components:

	PC1	PC2	PC3	PC4
Standard deviation	1.5749	0.9949	0.59713	0.41645
Proportion of Variance	0.6201	0.2474	0.08914	0.04336
Cumulative Proportion	0.6201	0.8675	0.95664	1.00000

# 3.주성분분석



# Standard deviation : 표준편차

# Proportion of Variance : 분산비율, 각 주성분의 차지하는 비율을 말하며  
클수록 영향도가 그만큼 높다는 의미

# Cumulative Proportion : 분산의 누적 합계.

첫 번째 주성분분석 하나가 전체 분산의 62%를 설명하고 있다.

두 번째는 24.7%를 설명하고 있다. 반대로 얘기 하면 첫 번째 주성분분석만  
수용 했을 때 정보 손실은(100-62)=38%가 된다.

고유값(eigenvalue) = (standard deviation)^2

# 3.주성분분석



fit\$rotation -> 주성분 분석 함수 계수

	PC1	PC2	PC3	PC4
Murder	-0.5358995	0.4181809	-0.3412327	0.64922780
Assault	-0.5831836	0.1879856	-0.2681484	-0.74340748
UrbanPop	-0.2781909	-0.8728062	-0.3780158	0.13387773
Rape	-0.5434321	-0.1673186	0.8177779	0.08902432

첫 번째 주성분 함수

$Y1 = -0.536\text{Murder} - 0.583\text{Assault} - 0.278\text{UrbanPop} - 0.543\text{Rape}$

# 3.주성분분석



## (5)주성분수 결정

정보손실을 최소화하는 방법과 주성분의 분산을 최대화하는 방법이며, 누적기여율과 고유치 값은 주성분의 수를 결정하는 기준.

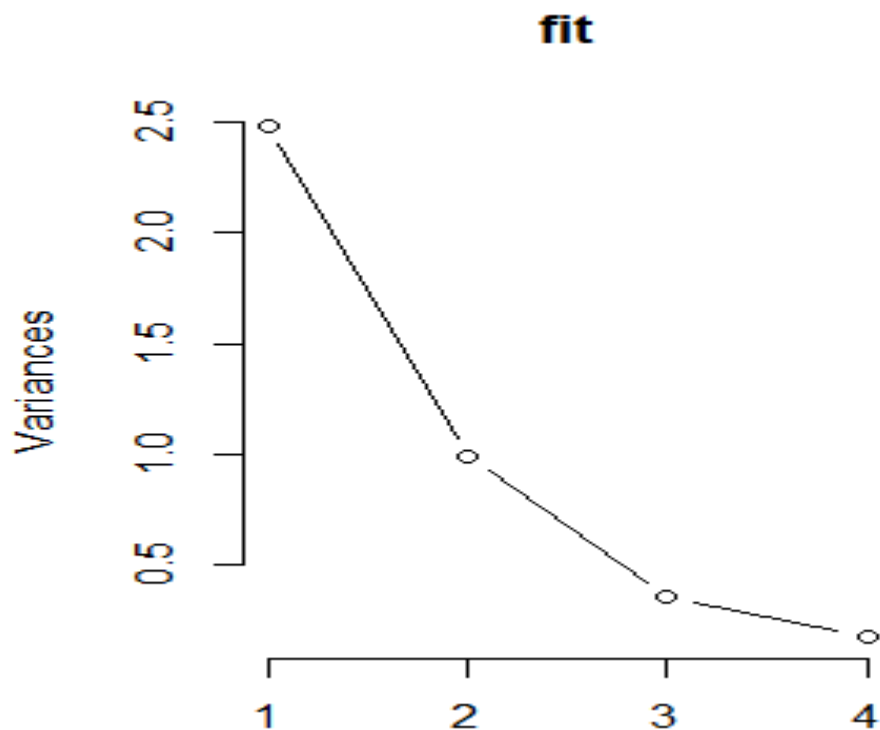
일반적으로 고유치 값이 1.0 이상, 누적기여율이 80% 되는 주성분을 기준으로 주성분 수를 결정한다.

# 3.주성분분석



## (6)스크리 플롯

이 그래프는 PCA에서 사용하는 스크리 플롯(scree plot)이라고 함  
그래프가 완만해 지는 부분 이전까지만 활용하는 것이 바람직함



# Quiz



Q04

주성분분석 해석으로 옳지 않은 것은?

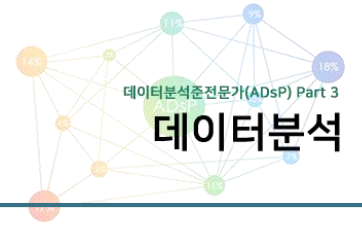
```
> data3<-princomp(data1,cor=TRUE) # ISLR 패키지 data(Hitters)
> data3
Call:
princomp(x = data1, cor = TRUE)

Standard deviations:
Comp.1      Comp.2      Comp.3      Comp.4      Comp.5      Comp.6      Comp.7
2.77339670  2.03026013  1.31485574  0.95754099  0.84109683  0.72374220  0.69841796
생략
17 variables and 263 observations.
> summary(data3)
Importance of components:

               Comp.1      Comp.2      Comp.3      Comp.4      Comp.5
Standard deviation      2.7733967  2.0302601  1.3148557  0.9575410
0.84109683
Proportion of Variance  0.4524547  0.2424680  0.1016968  0.0539344
0.04161435
Cumulative Proportion  0.4524547  0.6949227  0.7966195  0.8505539
0.89216822
생략
```

- ① 80% 이상 설명하려면 주성분 4개이상 선택하면 된다
- ② 제1성분의 설명력은 45%이다
- ③ 공분산행렬을 활용한 결과이다.
- ④ 17차원에서 2차원으로 줄이면 데이터 손실율은 약 30.51%이다

# Quiz



Q04

주성분분석 해석으로 옳지 않은 것은?

```
> data3<-princomp(data1,cor=TRUE) # ISLR 패키지 data(Hitters)
> data3
Call:
princomp(x = data1, cor = TRUE)

Standard deviations:
Comp.1      Comp.2      Comp.3      Comp.4      Comp.5      Comp.6      Comp.7
2.77339670 2.03026013 1.31485574 0.95754099 0.84109683 0.72374220
0.69841796
생략
17 variables and 263 observations.
> summary(data3)
Importance of components:

               Comp.1      Comp.2      Comp.3      Comp.4      Comp.5
Standard deviation      2.7733967 2.0302601 1.3148557 0.9575410
0.84109683
Proportion of Variance 0.4524547 0.2424680 0.1016968 0.0539344
0.04161435
Cumulative Proportion 0.4524547 0.6949227 0.7966195 0.8505539
0.89216822
생략
```

- ① 80% 이상 설명하려면 주성분 4개이상 선택하면 된다
- ② 제1성분의 설명력은 45%이다

정답

- ③ 공분산행렬을 활용한 결과이다.
- ④ 17차원에서 2차원으로 줄이면 데이터 손실율은 약 30.51%이다

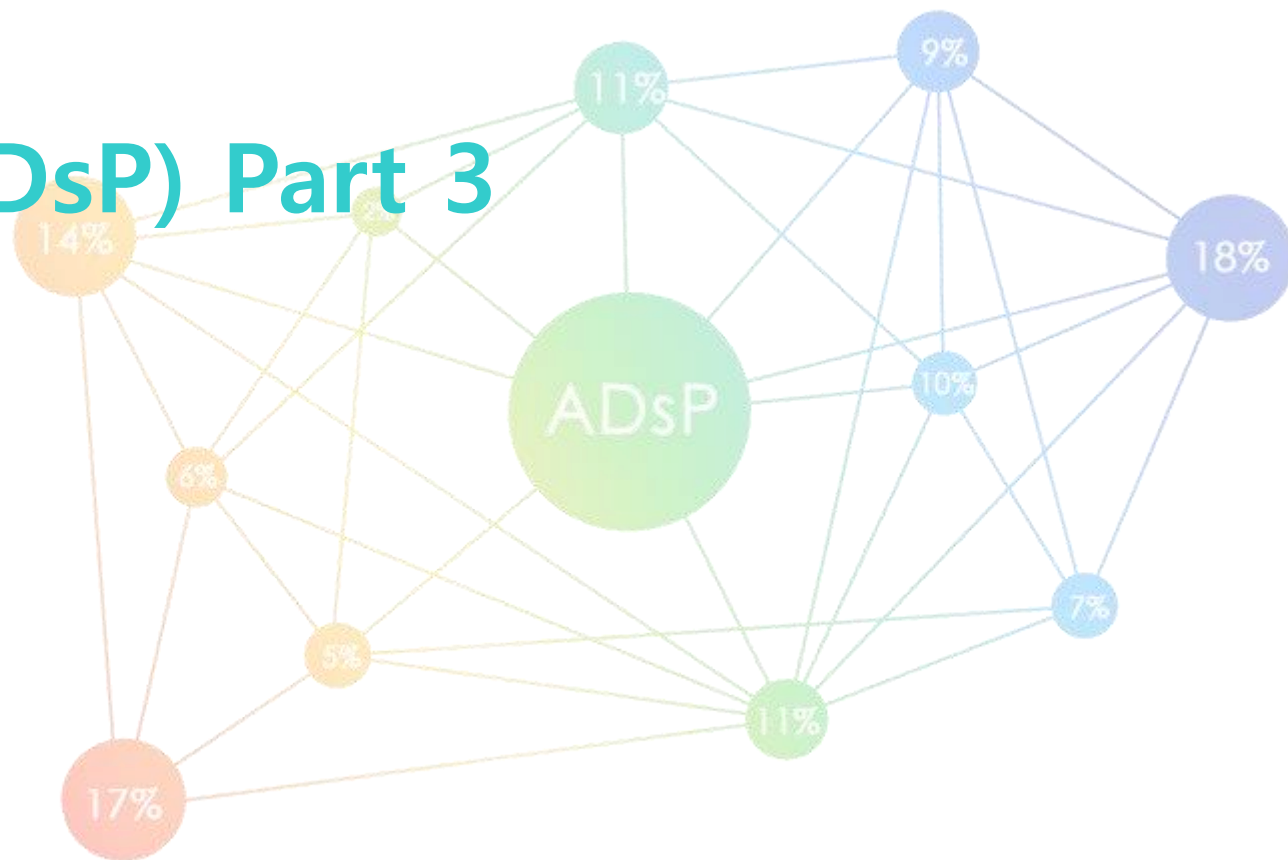
## 데이터분석준전문가(ADsP) Part 3

# 데이터분석

02

## 제2장 통계분석 제4절 시계열분석

1. 시계열이란?
2. 시계열 기본 개념 용어
3. 시계열 모양
4. 분해시계열



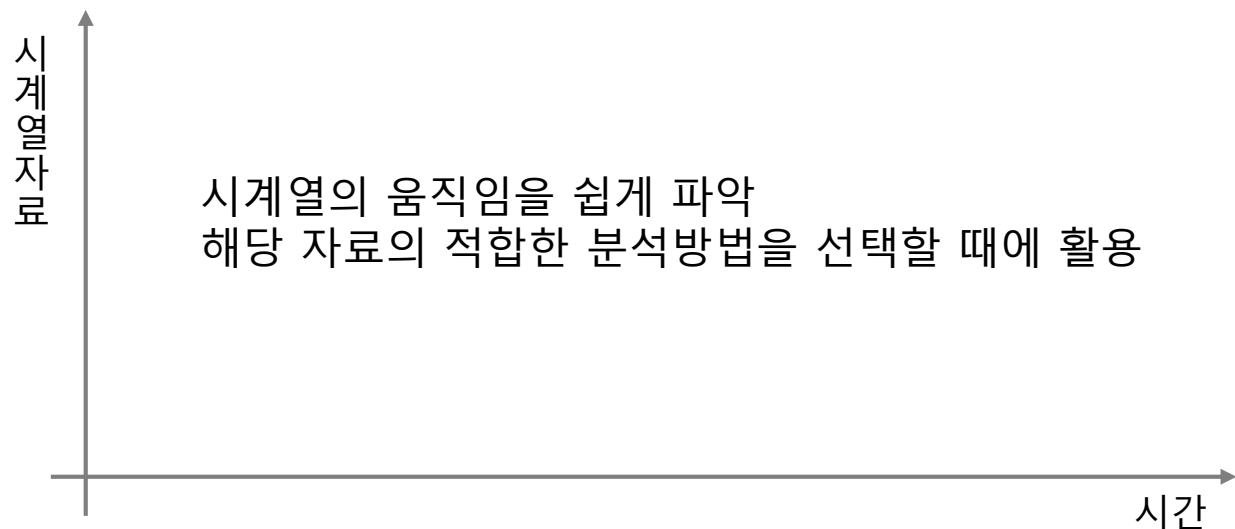


# 1.시계열이란?



## 시계열

시간의 흐름에 따라 자료의 변화



관측치의 시간적 순서를 나타내는 아래첨자  $t$ 를 사용하여 나타냄

$z_t$ (금월)의 관측치 의미  $z_{t-1}$ 은 1시점전(전월)의 관측치  $z_{t+1}$

1 시점후(다음월)의 관측치를 의미함

$$\{z_t\} = \{z_1, z_2, \dots, z_t\}$$

## 2.시계열 기본 개념 용어

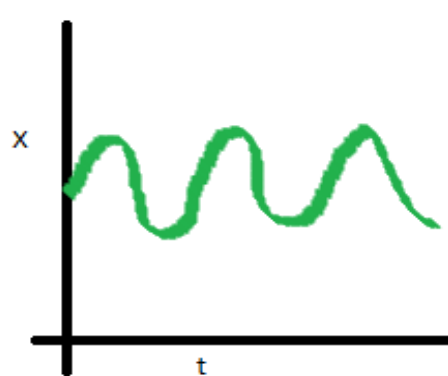


### 가.정상성(Stationarity=안정성)

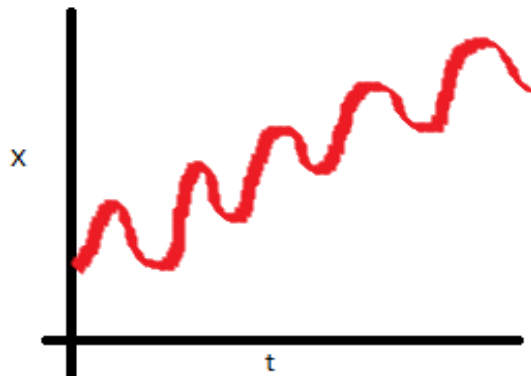
직관적으로 정상성의 의미는 시계열의 수준과 분산에 체계적인 변화가 없고  
엄밀하게 주기적 변동이 없다는 것으로 미래는 확률적으로 과거와 동일 의미

아래 3가지 조건을 만족하는 시계열을 정상시계열이라 한다.

1) 평균값은 시간  $t$ 에 관계없이 일정하다.  $E(Z_t)=\mu$



Stationary series

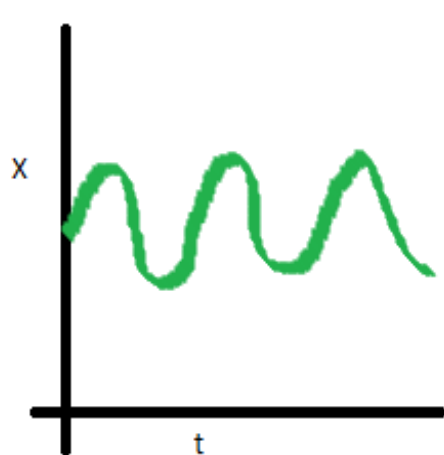


Non-Stationary series

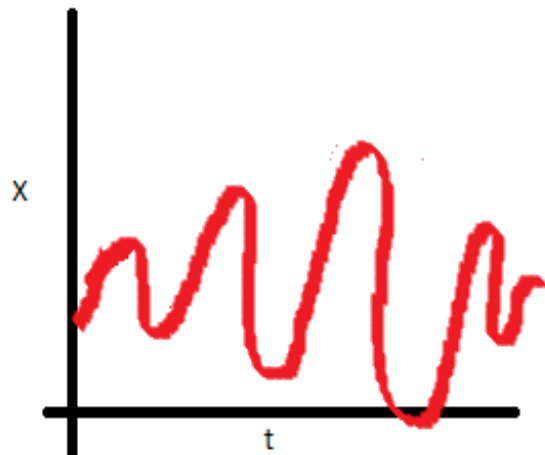
## 2.시계열 기본 개념 용어



2)분산값은 시간  $t$ 에 관계없이 일정하다.  $\text{Var}(Z_t)=\delta^2$



Stationary series



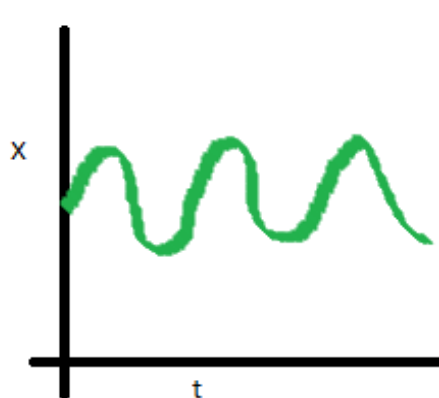
Non-Stationary series

## 2. 시계열 기본 개념 용어

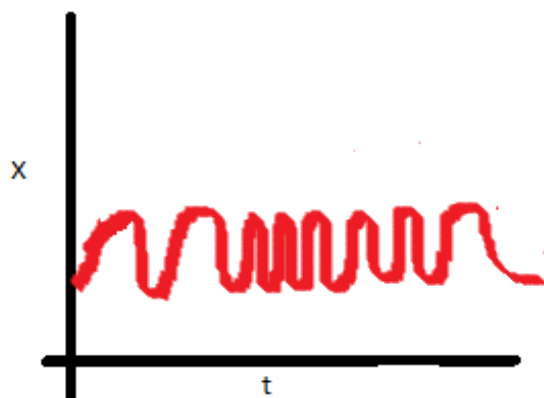


3) 공분산은 시간  $t$ 에 의존하지 않고 오직 시차에만 의존한다.

$$k = \text{시차} \quad \text{Cov}(Z_t, Z_{t+k}) = \gamma_k$$



Stationary series



Non-Stationary series

→ 실측값에서 시계열모형에 관한 통계적 추정 및 검정을 하기 위해서는  
정상성 가정 3가지 조건 중 하나라도 성립하지 않은 경우

**비정상 시계열**이라 함

평균이 일정하지 않으면 원계열을 차분해서 정상시계열로

분산이 일정하지 않으면 원계열에 자연로그 취하면 정상시계열로

## 2.시계열 기본 개념 용어



### 나.백색잡음=회귀분석의 오차항 가정과 유사

시계열이 상호독립적이고 동등하게 분포하는 확률변수의 관측치일 때 백색잡음 과정이라 한다.

$Cov(u_t, u_{t-k})=0 \rightarrow$  자기상관성이 없다=독립적이다

$E(u_t)=0 \rightarrow$  t에 관계없이 평균이 0이다.

$Var(u_t)=\delta^2 \rightarrow$  t에 상관없이 동분산

$U_t$ 을 백색잡음과정이다.

### 다.자기상관계수

자기상관계수는  $u_t$ 와 시차 k만큼 떨어진  $u_{t-k}$  간의 '선형관계의 정도'를 나타내는 것으로 일반적으로 통계학에서 다루는 두 변수간의 관계 정도를 나타내는 상관계수이다.

## 2.시계열 기본 개념 용어



$$\text{자기상관계수} = \text{Cov}(y_t, y_{t+k}) / \sigma^2$$

자기상관계수는 다음과 같은 성질이 있다.

- ①시차 0에서의 자기상관계수는 1이다.
- ②사차 1이후 상관계수처럼 -1~1 사이의값을 갖는다

→ 시차 k에 대응하여 자기상관계수가 산정되어 시차와 자기상관계수들은 함수 관계로 파악  
"ACF(AutoCorrelation Function)"이라 한다.

## 2. 시계열 기본 개념 용어



### 라. 부분자기상관계수

시계열 분석에는 자기상관 이외에 부분자기상관이라는 것이 있는데 이는 모형의 형태(특히 AR모형의 차수)을 찾는데 중요한 역할을 한다.

**현재**  $u_t$ 와 시차  $k$ 만큼 떨어진  $u_{t-k}$ 간의 부분자기 상관이란 중간에 있는 자기계열들  $u_{t-1}, u_{t-2}, \dots, u_{t-k}$ 의 영향을 제거하고  $u_t$ 와  $u_{t-k}$ 간의 순수 선형관계의 정도를 나타내는 지표

**“PACF(Partial AutoCorrelation Function)”이라 한다.**

# 3. 시계열모형



## 가. 자기회귀모형(AR)이란?

시계열자료에서 현재 종속변수( $u_t$ )를 설명하기 위해 자신의 과거 독립변수와 오차의 구성요소들간에 선형관계 나타낼 수 있다면 이모형을 자기회귀모형이라 한다.

이때 현재 종속변수( $u_t$ )를 설명할 수 있는 유의한 과거 시점들이  $p$ 개라면 이 모형의 차수는  $p$ 차 이고,  $AR(p)$ 로 나타낸다.

AR 모형에서의 **부분자기상관**은 차수보다 큰 시차에서 0으로 절단 식별의 단계에서 0으로 간주 될 수 있는 차수가 어디냐를 검정하여 AR의 차수를 결정하는데 **표본부분자기상관계수**를 이용한다.



# 3. 시계열 모형



## 나. 이동평균모형 MA모형 (정상성 가정이 필요없음)

현재자료  $u_t$ 를 현재오차( $e_t$ ) 과거오차( $e_{t-1} \dots e_{t-q}$ )로 설명하는 것을 이동평균모형  
설명변수인 과거시점의 오차들의 수  $q \rightarrow$  차수  $MA(q)$ 로 표시

AR모형은 차수보다  $K$ 이상 떨어진 시차에 대해서 부분자기상관계수가

0이 되어 소멸 MA모형은 차수보다  $K$ 이상 떨어진 시차에

자기상관계수가 0이 되어 소멸

### AR모형

부분자기상관이 어느 시차에서 소멸되는지를 기준으로 이전시차를  
AR모형의 차수  $p$ 로 결정하는데 이용

### MA 모형

순수한 MA모형은 자기상관이 어느 시차에서 소멸되는지를  
기준으로 MA모형의 차수  $q$ 를 결정하는데 이용

# 3. 시계열 모형



## 다. 자기회귀이동평균모형(ARMA)

AR의 차수  $p$ 와 MA의 차수  $q$

ARMA모형의 차수라 하고  $ARMA(p,q)$ 로 표현한다

AR+MA 성질을 갖고 있음

ARMA모형에서는 자기상관이나 부분자기상관의 어느 쪽도 차수보다 높은 시차에서 0의 값을 가지는 소멸되는 성질을 가지고 않고 0으로 접근해 감

ARMA 자기상관이나 부분자기상관에서 소멸되는 성질이 없음

그러나 자기상관계수와 이동평균계수의 부호에 따라 표본자료에서의 시차1 또는 2에서 어떤 규칙적인 모습의 성질을 이용한다.

# 3.시계열모형



## 라.ARIMA 모형의 구조

ARIMA(p d q) p=AR차수 d=차분수 q=MA차수-> 불안정시계열을 d차 차분한 후 안정적인 ARMA(p,q)모형이 된다는 것을 의미

ARIMA은 AR, MA, ARMA와 같은 다양한 시계열 모형을 표기

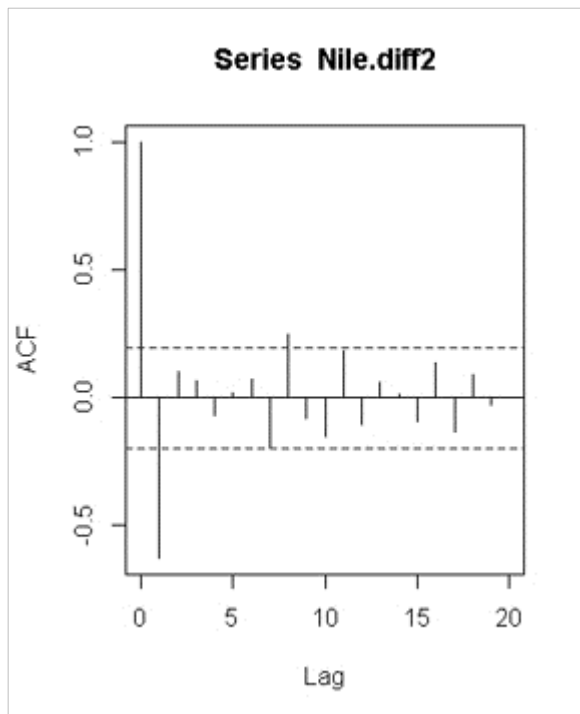
예) ARIMA(1 0 1)->ARMA(1 1), ARIMA(1 0 0)->AR(1),

ARIMA(0 0 1)->MA(1)

각각의 모형은 자기상관과 부분자기상관에서 각각 다른 모습을 가짐

모형	자기상관계수	부분자기상관계수
AR(p)	빠르게 0에 접근 (지수형태감소)	시차 p+1이후 절단
MA(q)	시차 q+1이후 절단	빠르게 0에 접근(지수형태감소)
ARMA(p,q)	빠르게 0에 접근	빠르게 0에 접근

# 3. 시계열 모형

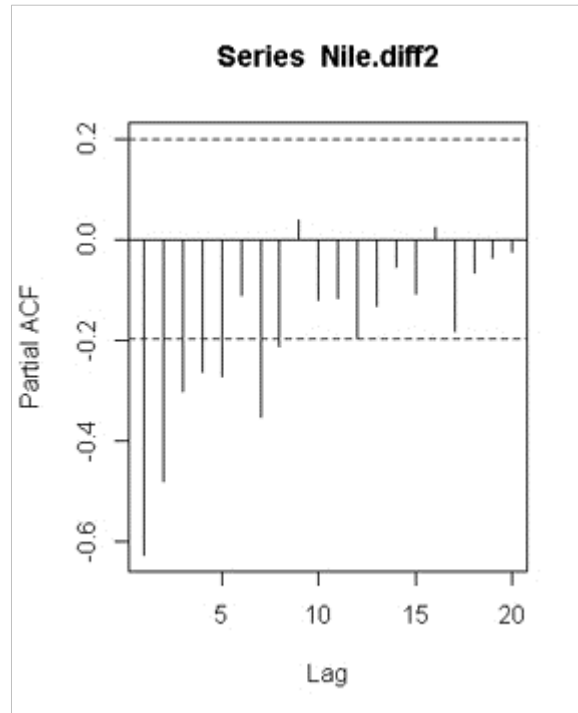


→ (점선으로 안으로 들어오면 절단이라 함)

Lag 1에서 -0.6 정도의 값을 갖고, Lag 2부터 점선 안의 경계로 들어와서 '절단' 된 상태입니다.

**MA(0,1)**

# 3. 시계열모형



→ Lag 6에서 '절단'되고, Lag 6,7 에서는 경계에서 벗어났다가 Lag 9에서는 다시 '절단' 되는 형태로 나타납니다.

**AR(8,0)**

# 3. 시계열 모형



## 마. 불안정 시계열의 안정화

1) 평균이 일정하지 않은 경우 → 차분

불안정을 가지는 대부분의 경제 시계열 자료는 1차 또는 2차 차분하면 안정화가 되고 불규칙 요인에 민감한 자료는 보다 높은 차분이 요구됨

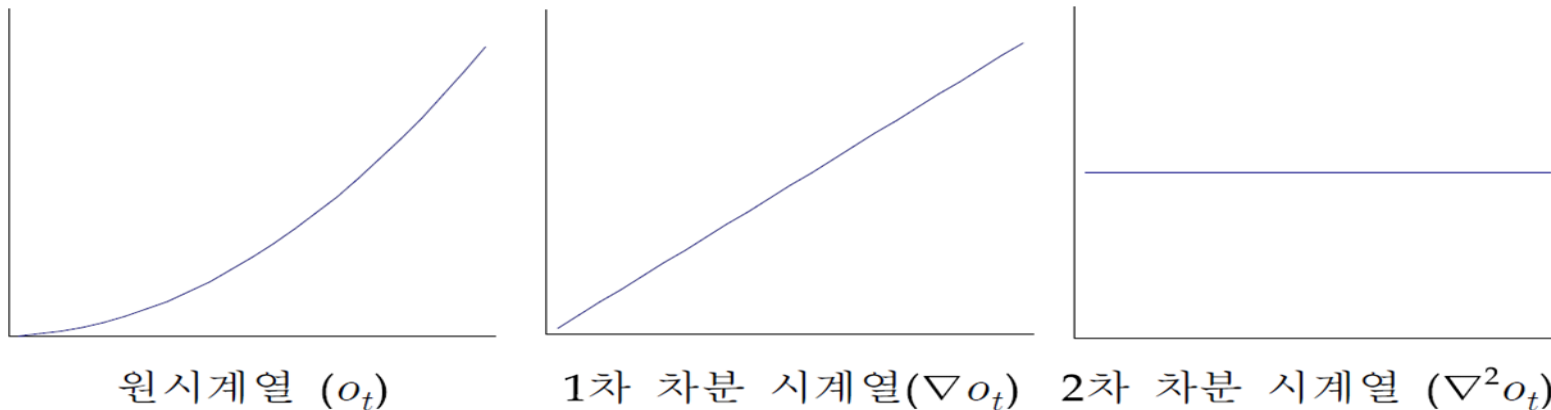
# 3. 시계열 모형



ex) 차분의 예

자료가 비선형 추세를 가진다고 하고  $\{1, 4, 9, 16, 25, 36, \dots\}$ 으로 변하는 자료라 하면 1차 차분하여  $\{3, 5, 7, 9, 11, \dots\}$ 가 되어 증가하는 선형 추세를 가지는 모습을 보인다.

여기서 2차 차분을 하면  $\{2, 2, 2, 2, \dots\}$ 가 되어 증가 추세가 없어지고, 일정한 값을 가지는 수평선이 됨을 알 수 있다.



# 3. 시계열모형



## 2) 분산이 이분산인 경우

정상시계열의 조건-> 동분산: 분산이 시간대에 관계없이 동일해야 하는 성질  
분산이 시간대에 따라서 일정하지 않은 경우-> 모수 추정의 왜곡

분산을 안정화하기 위한 자료의 변환 방법  
로그변환과 제곱근 변환

## 3) 과대차분(overdifferencing)

안정화 된 경우에도 차분하는 경우 자기상관계수 복잡, 분산을 크게

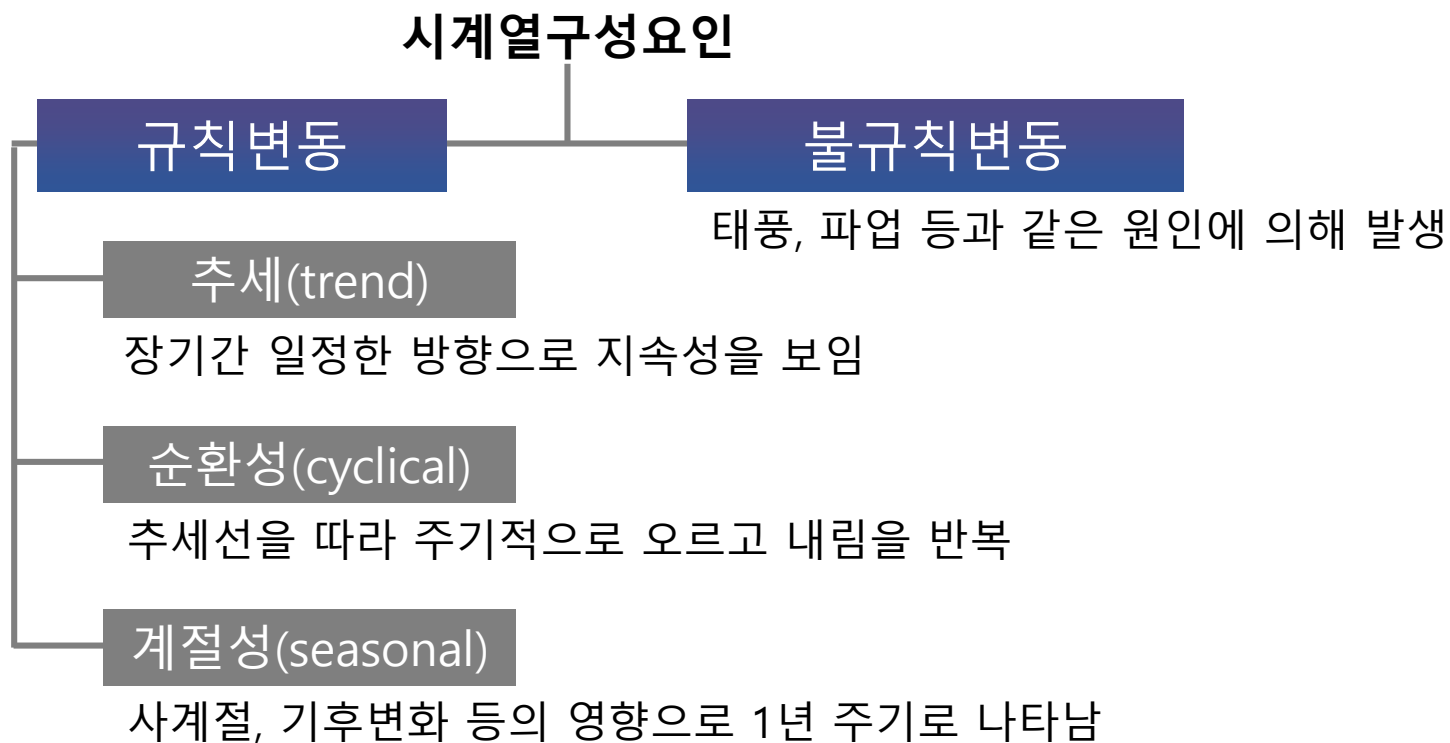


# 4. 분해시계열



## 분해시계열이란

시계열에 영향을 주는 요인을 시계열에서 분리해 분석하는 방법 시계열의 구성하는요소는 4가지로 분류



# 4. 분해시계열



## 1) 추세

시계열 기간이 짧은 자료-> 진정한 추세가 형성이 안됨

## 2) 순환요인

2-3년 주기로 순환. 시간의 흐름에 따라 추세를 중심으로 상하반복변화

## 3) 계절요인

계절성 1년 주기로 갖는다. 순환요인이 2-3년 주기

일정기간마다 주기적으로 나타나는 사회적 관습 요인도 계절요인에 해당

## 4) 불규칙변동

우연적으로 발생하는 변동. 추세, 순환, 계절요인을 제거하고 남은 변동

# Quiz



Q05

현재와 시차 (a)만큼 떨어진 과거간의 부분자기상관이란 중간에 있는 자기계열들의 과거들 (b),(c)의 영향을 제거하고 난 후의 현재와 시차 3만큼 떨어진 둘만의 순수한 선형관계의 정도를 나타내는 계수이다.

- ①(a)->3(b)->1(c)->2
- ②(a)->1(b)->2(c)->3
- ③(a)->1(b)->3(c)->2
- ④(a)->2(b)->3(c)->1

# Quiz



Q05

현재와 시차 (a)만큼 떨어진 과거간의 부분자기상관이란 중간에 있는 자기계열들의 과거들 (b),(c)의 영향을 제거하고 난 후의 현재와 시차 3만큼 떨어진 둘만의 순수한 선형관계의 정도를 나타내는 계수이다.

정답

①(a)->3(b)->1(c)->2

②(a)->1(b)->2(c)->3

③(a)->1(b)->3(c)->2

④(a)->2(b)->3(c)->1

Q06

안정시계열에 대한 설명으로 적절한 것은?

- ① 우상향하면 증가하는 시계열이다.
- ② 평균을 중심으로 일정한 폭으로 진동하는 시계열이다.
- ③ 지수형으로 감소하는 시계열이다.
- ④ 진동하면서 진폭이 커지는 시계열이다.

Q06

안정시계열에 대한 설명으로 적절한 것은?

① 우상향하면 증가하는 시계열이다.

정답

② 평균을 중심으로 일정한 폭으로 진동하는 시계열이다.

③ 지수형으로 감소하는 시계열이다.

④ 진동하면서 진폭이 커지는 시계열이다.

Q07

불안정시계열의 특징이 아닌 것은?

- ① 시간에 따라 평균이 달라진다
- ② 시간에 따라 분산이 변한다
- ③ 시간에 관계없이 분산이 일정하다
- ④ 선형추세를 갖고 있다.

Q07

불안정시계열의 특징이 아닌 것은?

- ① 시간에 따라 평균이 달라진다
- ② 시간에 따라 분산이 변한다
- ③ 시간에 관계없이 분산이 일정하다.
- ④ 선형추세를 갖고 있다.

정답



Q08

자기상관이 지수형태로 감소하는 모형 식별은?

- ① 자기회귀모형
- ② 이동평균모형
- ③ 일반회귀모형
- ④ 계량경제모형

## Q08 자기상관이 지수형태로 감소하는 모형 식별은?

정답

- ① 자기회귀모형
- ② 이동평균모형
- ③ 일반회귀모형
- ④ 계량경제모형

정답

자기상관이 지수형으로 감소하고 부분자기상관이  
P시차이후에 0과 유의한 값을 가지면 자기회귀모형이다.

# Quiz



Q09

시계열에 영향을 주는 일반적인 요인을  
시계열에서 분리해 분석하는 방법을 무엇이라 하는가?

# Quiz



Q09

시계열에 영향을 주는 일반적인 요인을  
시계열에서 분리해 분석하는 방법을 무엇이라 하는가?

정답

분해시계열

# Quiz



Q10

비정상 시계열을 정상 시계열로 전환하는 방법 중 현 시점의 자료값에서 전 시점의 자료값을 빼주는 것을 무엇이라 하는가?

# Quiz



Q10

비정상 시계열을 정상 시계열로 전환하는 방법 중 현 시점의 자료값에서 전 시점의 자료값을 빼주는 것을 무엇이라 하는가?

정답

차분

Q11

다음 설명하는 시계열 모형은 무엇인가?

현시점의 자료를 유한개의 백색잡음의 선형결합으로 표현되었기 때문에 항상 정상성을 만족한다. 자기상관함수  $p+1$ 시차 이후 절단된 형태를 취한다.

## Q11 다음 설명하는 시계열 모형은 무엇인가?

현시점의 자료를 유한개의 백색잡음의 선형결합으로 표현되었기 때문에 항상 정상성을 만족한다. 자기상관함수  $p+1$ 시차 이후 절단된 형태를 취한다.

정답

MA