

1주차 2차시 : 회귀분석 개요

1. 계량경제의 분류
2. 회귀분석의 개요

1. 계량경제의 분류

구분	회귀분석		시계열분석		패널분석	
	단일방정식	연립방정식	일변량 (univariate)	다변량 (multivariate)	회귀분석	시계열분석
데이터	횡단면/시계열	횡단면/시계열	시계열	시계열	pool/패널	pool/패널
모형	①단순모형	-	④AR(p)	⑧VAR(p) / VECM	⑨패널모형	⑩패널VAR(p) / 패널 VECM
	②다중모형	③연립방정식모형	⑤MA(q) ⑥ARMA(p,q) ⑦ARIMA(p,d,q)			
	정태모형/동태모형	정태모형/동태모형	동태모형	동태모형		
추정방법	LS(OLS, WLS, ILS), MLE	LS(OLS, 2SLS, 3SLS)	OLS, NLS, MLE	OLS	OLS, LSDV, DID	OLS, GMM
software	Excel, Stata, R, Python					

(참고 1) 자료의 형태

- ① 횡단면 자료(Cross-Sectional data) : 일정시점에서 하나 이상의 변수에 대해 수집된 자료
- ② 시계열 자료(Time Series data) : 일별, 주별, 월별, 분기별, 연도별 등 시간에 걸쳐 수집한 자료
- ③ 합동 횡단면 자료(Pooled Cross Sections data) : 횡단면 자료와 시계열자료가 결합된 자료
- ④ 패널자료(Panel data) : 동일한 횡단면 단위에 기준을 두고 시간의 흐름에 따라 수집한 자료

(참고 2) 활용 프로그램

- OA : MS Excel
- 소프트웨어 / 통계 패키지 : Stata
- 프로그래밍 언어 : R, Python

①단순모형 : $y_i = \alpha + \beta x_i + \epsilon_i$ $y_t = \alpha + \beta x_t + \epsilon_t$
 ($i = 1, 2, \dots, N$) ($t = 1, 2, \dots, T$)

②다중모형 : $y_i = \alpha + \beta X_i + \epsilon_i$
 $X_{ki} = (x_{1i}, x_{2i}, \dots, x_{k-1i})$
 ($i = 1, 2, \dots, N$)

③연립방정식모형(예시) : $y_{1i} = \beta_{10} + \beta_{11}y_{2i} + \gamma_{11}x_{1i} + u_{1i}$
 $y_{2i} = \beta_{20} + \beta_{21}y_{1i} + \gamma_{21}x_{1i} + u_{2i}$

④AR(p) : $y_t = \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \epsilon_t$
 ($p = 1, 2, \dots, p$)

⑤MA(q) : $y_t = e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q}$
 ($q = 1, 2, \dots, q$)

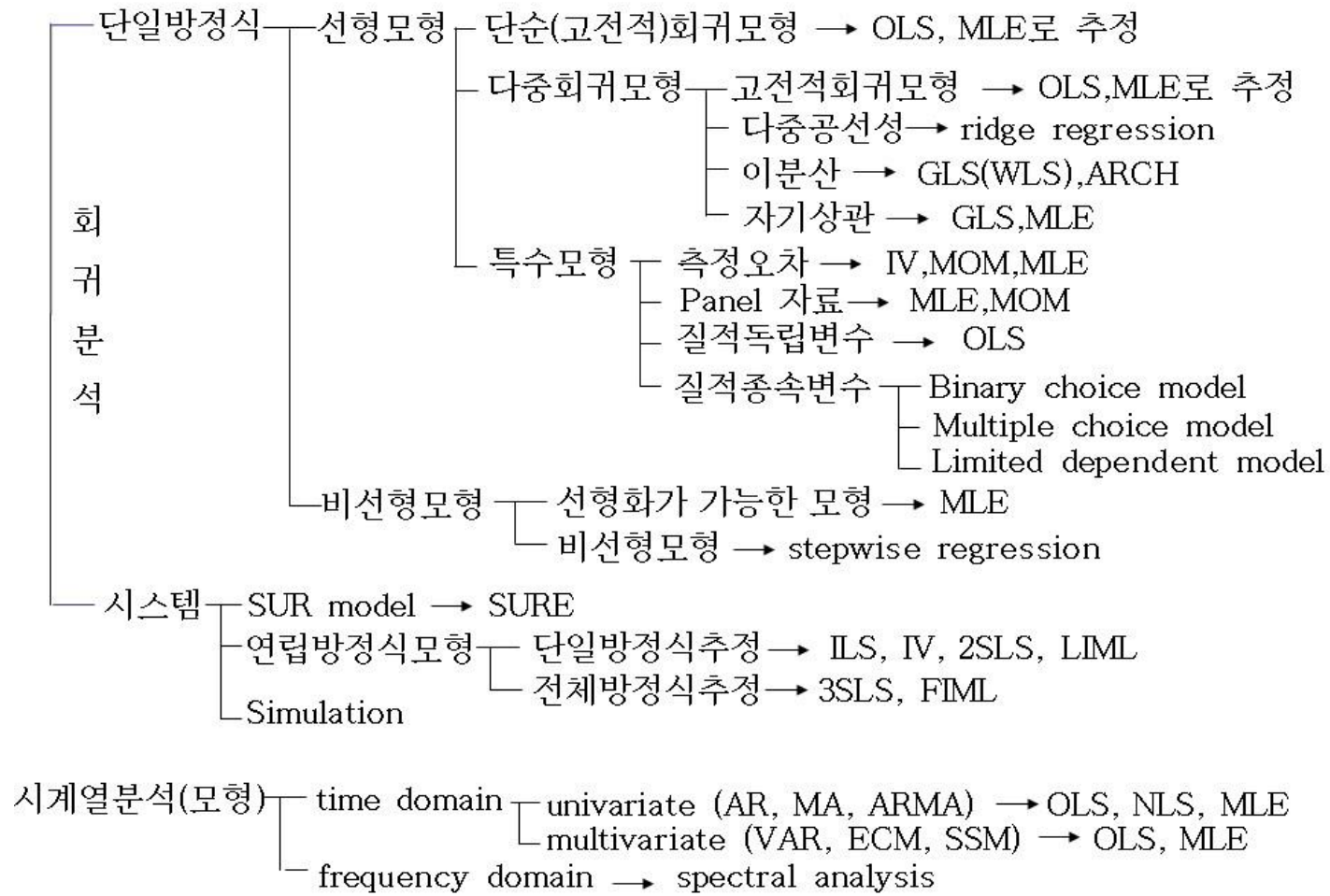
⑥ARMA(p,q) : $y_t - \phi_1 y_{t-1} - \dots - \phi_p y_{t-p} = e_t + \theta_1 e_{t-1} + \dots + \theta_q e_{t-q}$
 ($p = 1, 2, \dots, p, q = 1, 2, \dots, q$)

⑦ARIMA(p,d,q) : $\Delta y_t - \phi_1 \Delta y_{t-1} - \dots - \phi_p \Delta y_{t-p} = e_t + \theta_1 e_{t-1} + \dots + \theta_q e_{t-q}$
 ($p = 1, 2, \dots, p, q = 1, 2, \dots, q$)

⑧VAR(p) : $X_t = \mu + A_1 X_{t-1} + \dots + A_k X_{t-k} + u_t$ 또는 $X_t = \mu + A(L) X_{t-1} + u_t$
 ($t = 1, 2, \dots, T$) ($t = 1, 2, \dots, T$)

⑨패널모형 : $y_{it} = \alpha + \sum_{k=1}^K \beta_k X_{kit} + u_{it}$ 또는 $y_{it} = \alpha_i + \sum_{k=1}^K \beta_k X_{kit} + \epsilon_{it}$, (단 $\alpha_i = \alpha + \mu_i$)

⑩패널 VAR(p) : $x = \mu_i + A_i(L) X_{t-1} + u$
 ($i = 1, 2, \dots, N, t = 1, 2, \dots, T$)



2. 회귀분석 개요

(1) 회귀(regression)란?

① 역사적 원천: Galton의 “보편적 회귀의 법칙”

② 현대적 의미: 종속변수와 독립변수와의 의존관계를 분석하고, 이미 알려진 독립변수의 값으로 종속변수의 평균적인 값을 추정 또는 예측

(2) 확정적(함수적) 관계 vs. 확률적(통계적) 관계

① 확정적 관계: $Y = \beta_0 + \beta_1 X$

② 확률적 관계: $Y = \beta_0 + \beta_1 X + u$

(3) 회귀(regression)와 인과(causality)

- 통계적 관계(회귀)가 아무리 강하다고 하더라도 반드시 인과관계를 나타내지 않으며 인과관계는 경제이론이나 선행적인 것으로부터 도출

(예1) reg 수확량 on 강우량(o.k) / reg 강우량 on 수확량(no)

(예2) reg 소비 on 소득(o.k) / reg 소득 on 소비(no)

(4) 회귀분석(regression analysis)와 상관분석(correlation analysis)

회귀분석	상관분석
<ul style="list-style-type: none"> - 독립변수는 확정변수로 가정하고 종속변수는 확률변수로 가정 - 독립변수의 주어진 값으로 종속변수의 평균값을 추정·예측하는 것이 목적 - 독립변수와 종속변수로 구분 	<ul style="list-style-type: none"> - 두 변수 모두 확률변수를 가정 - 두 변수간의 선형성의 정도를 측정하는 것이 목적 - 독립변수와 종속변수의 구분이 없음

(5) 용어

Y	X
Dependent variable(종속변수)	Independent variable(독립변수)
Explained variable(설명된 변수)	Explanatory variable(설명변수)
Predictand(예측된 변수)	Predictor(예측변수)
Regressand(피회귀변수)	Regressor(회귀변수)
Response(반응변수)	Stimulus(자극변수)
Endogenous variable(내생변수)	Exogenous variable(외생변수)
Controlled variable(통제된 변수)	Control variable(통제변수)
Outcome(결과)	Covariate(공변변수)