

Principles of Econometrics (3e)

Ch. 7 비선형 관계

2013년 1학기

윤성민

7.1 다항식

- 두 가지 유형의 비선형 모형

- ① 변수 간에는 비선형이지만, 모수 측면에서는 선형인 경우
⇒ 적절한 변형을 통해 OLS 이용 추정할 수 있음

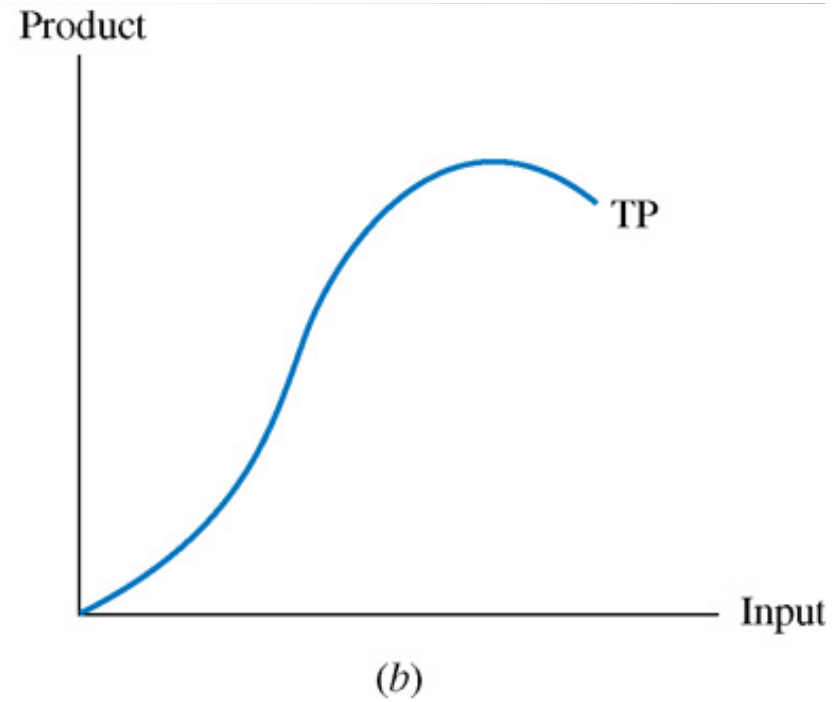
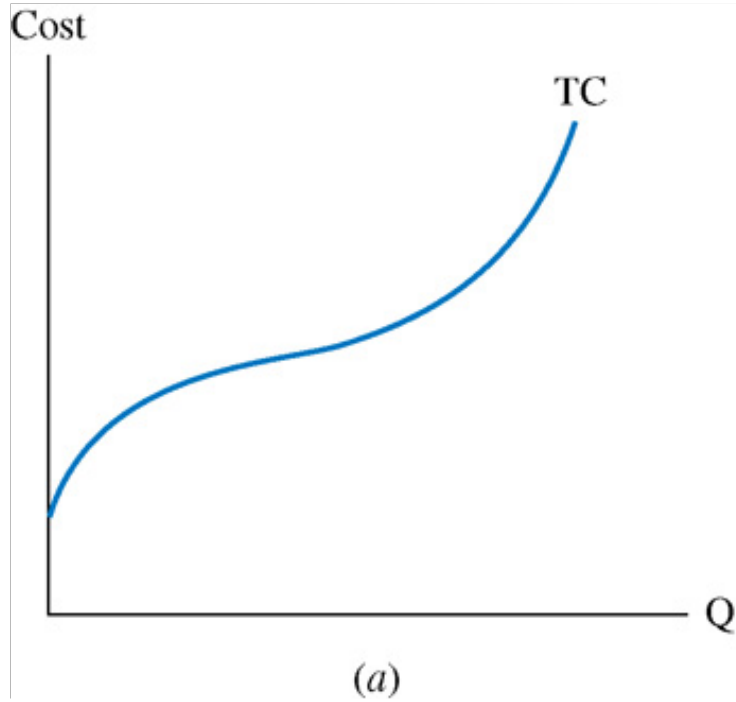
$$Y = \alpha L^{\beta} K^{\gamma} \quad \Rightarrow \quad \ln(Y) = \delta + \beta \ln(L) + \gamma \ln(K)$$

- ② 모수 측면에서 비선형, 변형을 하더라도 선형이 될 수 없음
⇒ 비선형 최소제곱법(nonlinear least squares: NLS)을
이용하여 추정

$$y_t = \beta x_{t1} + \beta^2 x_{t2} + e_t \qquad y_t = \frac{\alpha}{1 + \exp(-\beta - \delta t)} + e_t$$

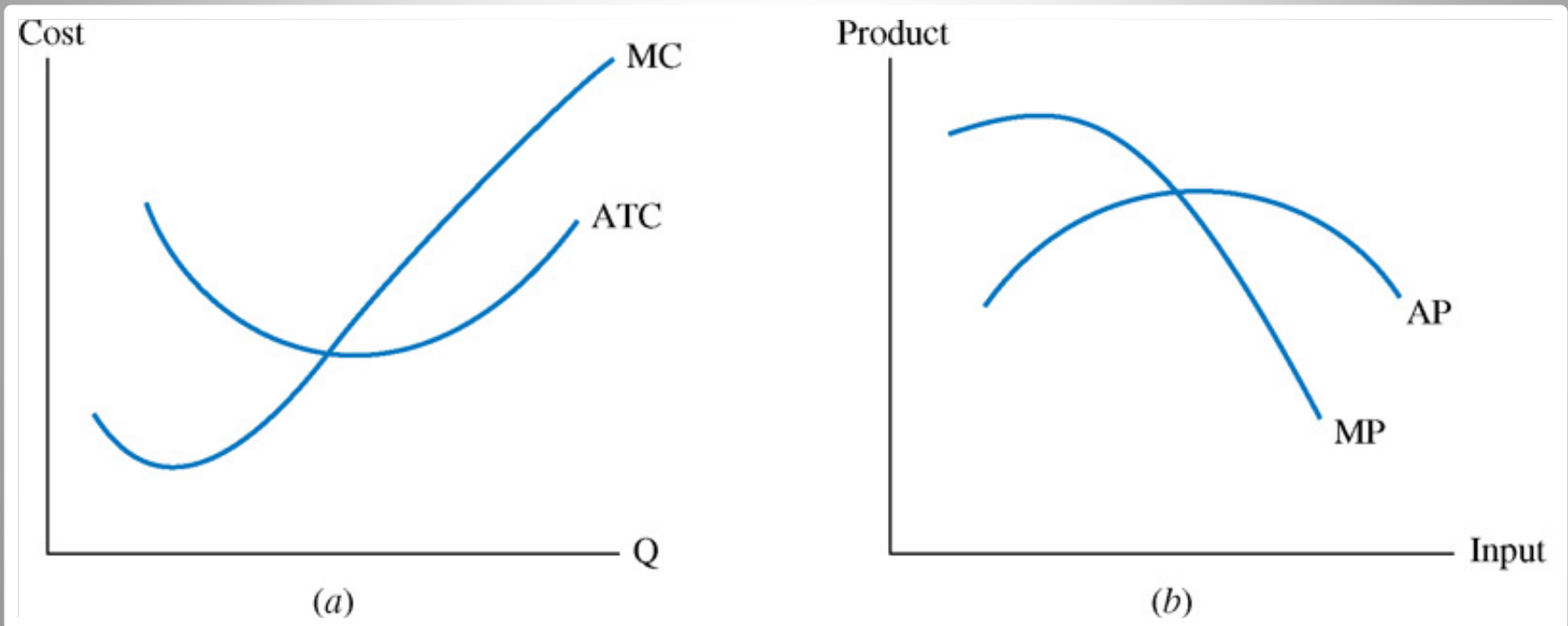
7.1.1 비용곡선 및 생산물곡선

- 총비용곡선(TC) $TC = \alpha_1 + \alpha_2 Q + \alpha_3 Q^2 + \alpha_4 Q^3 + e$
- 총생산물곡선(TP)



■ 평균비용곡선(AC), 한계비용곡선(MC)

- 한계비용곡선(MC) $MC = \frac{dE(TC)}{dQ} = \alpha_2 + 2\alpha_3Q + 3\alpha_4Q^2$
- 평균비용곡선(AC) $AC = \beta_1 + \beta_2Q + \beta_3Q^2 + e$



- TC: $TC = \alpha_1 + \alpha_2 Q + \alpha_3 Q^2 + \alpha_4 Q^3 + e$
- AC: $AC = \beta_1 + \beta_2 Q + \beta_3 Q^2 + e$
- Q^2 , Q^3 를 각각의 새로운 변수로 생각하여 OLS 적용하여 추정하면 됨
- 모수를 해석할 때 주의가 필요함

$$\frac{dE(TC)}{dQ} = \alpha_2 + 2\alpha_3 Q + 3\alpha_4 Q^2$$

$$\frac{dE(AC)}{dQ} = \beta_2 + 2\beta_3 Q$$

7.1.2 임금방정식

- 임금 피크제
 - 젊고 경험이 없는 근로자는 경험이 축적될수록 임금 상승
 - 중년 이후 은퇴에 가까워짐에 따라 임금이 하락하게 됨
- 임금 피크제를 반영한 임금방정식 (2차항이 추가됨)

$$WAGE = \beta_1 + \beta_2 EDUC + \beta_3 EXPER + \beta_4 EXPER^2 + e$$

- 역-U자 형태가 되려면, $\beta_3 > 0$, $\beta_4 < 0$
- 경험이 임금에 미치는 한계효과 $\frac{\partial E(WAGE)}{\partial EXPER} = \beta_3 + 2\beta_4 EXPER$

■ 임금방정식 추정 결과

Table 7.1 Wage Equation with Quadratic Experience

Variable	Coefficient	Std. Error	t-Statistic	Prob.
<i>C</i>	-9.8177	1.0550	-9.3062	0.0000
<i>EDUC</i>	1.2101	0.0702	17.2282	0.0000
<i>EXPER</i>	0.3409	0.0514	6.6292	0.0000
<i>EXPER</i> ²	-0.0051	0.0012	-4.2515	0.0000
$R^2 = 0.2709$		$SSE = 28420.08$		

- 교육 받은 기간이 1년씩 증가하면, 시간당 임금은 \$1.21 증가
- 경험 연수 18년에서 경험의 한계효과 (시간당 15.76센트)

$$\left. \frac{\partial E(WAGE)}{\partial EXPER} \right|_{EXPER=18} = 0.3409 + 2 \times (-0.0051) \times (18) = 0.1576$$

- 경험 효과의 전환점 $EXPER = -\beta_3 / 2\beta_4 = -0.3409 / 2 \times (-0.0051) = 33.47(\text{년})$

7.2 더미변수 (모의변수, dummy variable)

- 표본의 일부 관찰값에 대해 회귀모형의 모수값이 달라지는 경우가 있음

(예) 흑인의 한계소비성향은 백인보다 높음

⇒ 그러한 경우를 모형에 반영하기 위해 더미변수 사용

- 더미변수(dummy variables)
 - 질적 차이를 나타내는 변수, 보통 0과 1의 두 가지 값을 가짐

(예) 성: 남성=0, 여성=1 인종: 백인=0, 유색인=1

지역: 도시=0, 지방=1

✓ 여러 가지 값을 가질 수도 있음, 연도더미, 지역더미,....

7.2.1 절편 더미변수

<주택가격 사례>

$$P_t = \beta_1 + \beta_2 S_t + e_t \quad P : \text{주택가격}, \quad S : \text{주택의 크기}$$

- 부동산은 위치가 중요, 위치는 주택의 질적 특성
- 이를 반영하기 위해 더미변수를 다음과 같이 정의할 수 있음

$D_t = 1$ if 특성 존재 (예, 좋은 학군, 좋은 전망,...)

$D_t = 0$ if 특성 없음

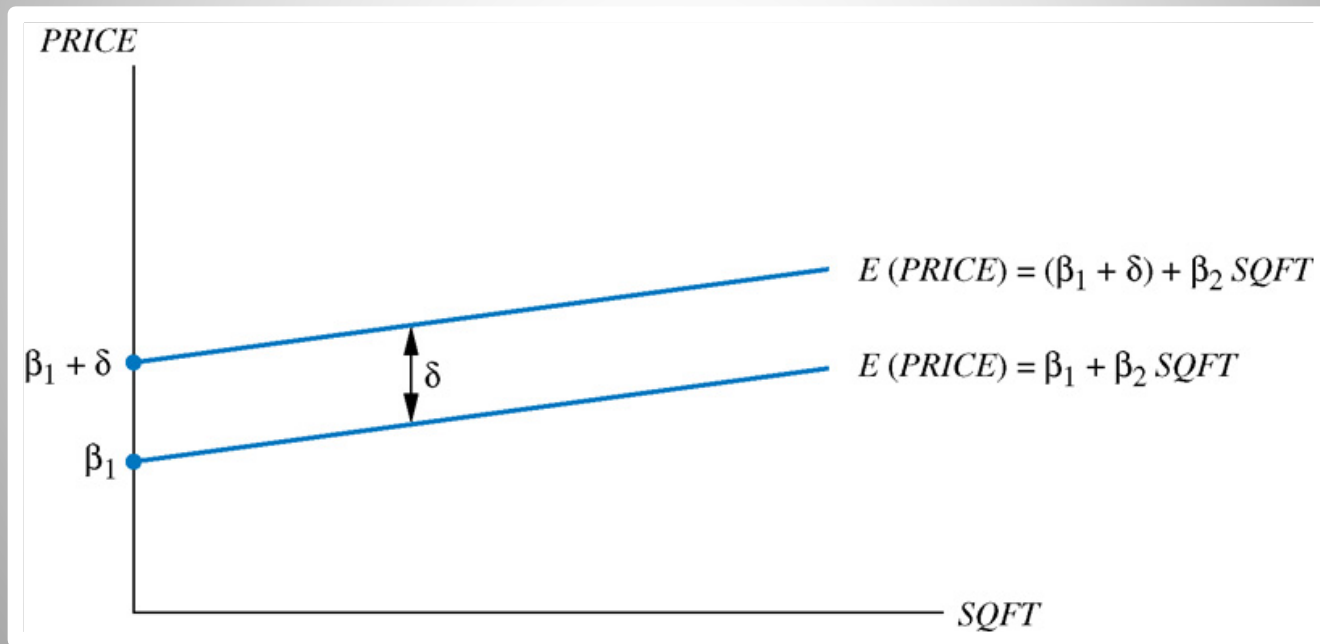
- 더미변수를 추가시킨 모형

$$P_t = \beta_1 + \delta D_t + \beta_2 S_t + e_t$$

■ 주택가격모형

$$P_t = \beta_1 + \delta D_t + \beta_2 S_t + e_t$$

$$E(P_t) = \begin{cases} (\beta_1 + \delta) + \beta_2 S_t & \text{when } D_t = 1 \\ \beta_1 + \beta_2 S_t & \text{when } D_t = 0 \end{cases}$$



- $\delta = 0$ 이면, 좋은 위치로 인한 프리미엄 존재하지 않는다는 의미

7.2.2 기울기 더미변수

- 질적 특성이 기울기에 영향을 미치는 경우

$$P_t = \beta_1 + \beta_2 S_t + \gamma(S_t D_t) + e_t$$

- 새로운 변수 $S_t D_t$ 는 주택크기와 위치가 주택가격에 미치는 상호작용 효과를 나타내므로,
상호작용변수(interaction variable)라고 함
- $\gamma > 0$ 이라면, 좋은 특성(주거환경)을 가진 지역에서는 주택크기가 주택가격에 추가적인 상승요인을 제공한다는 의미

$$P_t = \beta_1 + \beta_2 S_t + \gamma(S_t D_t) + e_t$$

$$E(P_t) = \beta_1 + \beta_2 S_t + \gamma(S_t D_t) = \begin{cases} \beta_1 + (\beta_2 + \gamma)S_t & \text{when } D_t = 1 \\ \beta_1 + \beta_2 S_t & \text{when } D_t = 0 \end{cases}$$

$$\frac{\partial E(P_t)}{\partial S_t} = \begin{cases} \beta_2 + \gamma & \text{when } D_t = 1 \\ \beta_2 & \text{when } D_t = 0 \end{cases}$$

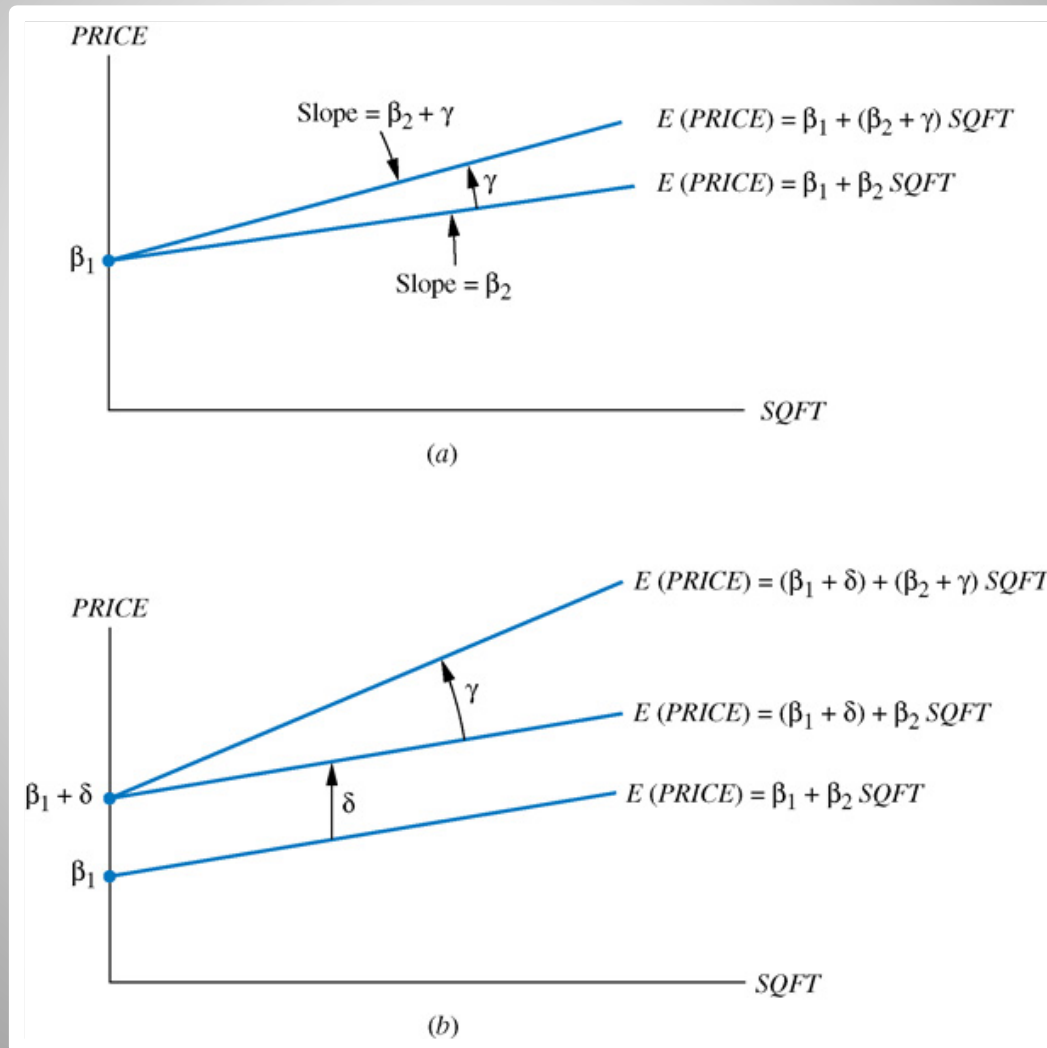
- 좋은 주거환경에 위치한 주택의 단위면적당 가격은 $(\beta_2 + \gamma)$
- 다른 지역에 위치한 주택의 단위면적당 가격은 β_2

- 주택의 위치가 절편과 기울기 모두에 영향을 미친다면, 두 효과를 동시에 모형에 포함시킬 수 있음

$$P_t = \beta_1 + \delta D_t + \beta_2 S_t + \gamma(S_t D_t) + e_t$$

$$E(P_t) = \begin{cases} (\beta_1 + \delta) + (\beta_2 + \gamma)S_t & \text{when } D_t = 1 \\ \beta_1 + \beta_2 S_t & \text{when } D_t = 0 \end{cases}$$

- 기울기 더미 $P_t = \beta_1 + \beta_2 S_t + \gamma(S_t D_t) + e_t$
- 절편 및 기울기 더미 $P_t = \beta_1 + \delta D_t + \beta_2 S_t + \gamma(S_t D_t) + e_t$



7.2.3 사례: 주택가격에 위치(대학 소재)가 미치는 영향

- 두 개의 유사한 주택지구에서 1,000개의 표본 수집
 - 대규모 주립대 인접 주택지구(519), 대학에서 먼 주택지구(481)

Table 7.2 Representative real estate data values

Price	Sqft	Age	Utown	Pool	Fplace
205452	2346	6	0	0	1
185328	2003	5	0	0	1
301037	2987	6	1	0	1
264122	2484	4	1	0	1
253392	2053	1	1	0	0
257195	2284	4	1	0	0
263526	2399	6	1	0	0
300728	2874	9	1	0	0
220987	2093	2	1	0	1

PRICE: 주택가격 (달러)

SQFT: 주택크기 (평방 피트)

AGE: 주택 나이 (년)

UTOWN=1 if 대학 인접

POOL=1 if 수영장 있음

FPLACE=1 if 백난로 있음

* patio, pet, garage, fitness, ...

- 주택가격 회귀모형

$$PRICE_t = \beta_1 + \delta_1 UTOWN_t + \beta_2 SQFT_t + \gamma(SQFT_t \times UTOWN_t) + \beta_3 AGE_t + \delta_2 POOL_t + \delta_3 FPLACE_t + e_t$$

Table 7.3 House Price Equation Estimates

Variable	Coefficient	Std. Error	t-Statistic	Prob.
<i>C</i>	24.5000	6.1917	3.9569	0.0001
<i>UTOWN</i>	27.4530	8.4226	3.2594	0.0012
<i>SQFT</i>	7.6122	0.2452	31.0478	0.0000
<i>SQFT</i> × <i>UTOWN</i>	1.2994	0.3320	3.9133	0.0001
<i>AGE</i>	−0.1901	0.0512	−3.7123	0.0002
<i>POOL</i>	4.3772	1.1967	3.6577	0.0003
<i>FPLACE</i>	1.6492	0.9720	1.6968	0.0901

$R^2 = 0.8706$

$SSE = 230184.4$

- 대학 인접 주택의 추정된 회귀함수

$$\begin{aligned} \hat{PRICE} &= (24500 + 27453) + (76.12 + 12.99)SQFT - 190.09AGE + 4377.16POOL + 1649.17FPLACE \\ &= 51953 + 89.11SQFT - 190.09AGE + 4377.16POOL + 1649.17FPLACE \end{aligned}$$

- 대학에서 먼 지역 주택의 추정된 회귀함수

$$\hat{PRICE} = 24500 + 76.12SQFT - 190.09AGE + 4377.16POOL + 1649.17FPLACE$$

<추정결과 해석>

- 대학 인접 프리미엄은 \$27,453
- 단위면적당 가격: 대학 인접 주택=\$89.11, 타 지역=\$76.12
- 주택은 매년 \$190.09씩 감가상각이 발생한다고 추정됨
- 수영장 프리미엄은 \$4,377.16
- 벽난로 프리미엄은 \$1,649.17

7.3 더미변수 사용법

7.3.1 질적인 요소간의 상호작용

- 단순한 임금모형: 임금($WAGE$)은 생산성($EDUC$)에 의존
- 인종별, 성별 격차가 존재함
 - 백인이면서 남성인 경우, 특별한 임금대우를 해주는 상황

$$WAGE = \beta_1 + \beta_2 EDUC + \delta_1 BLACK + \delta_2 FEMALE + \gamma (BLACK \times FEMALE) + e$$

$$BLACK = \begin{cases} 0 & \text{white} \\ 1 & \text{black} \end{cases} \quad FEMALE = \begin{cases} 0 & \text{male} \\ 1 & \text{female} \end{cases}$$

$$E(WAGE) = \begin{cases} \beta_1 + \beta_2 EDUC & \text{WHITE - MALE} \\ (\beta_1 + \delta_1) + \beta_2 EDUC & \text{BLACK - MALE} \\ (\beta_1 + \delta_2) + \beta_2 EDUC & \text{WHITE - FEMALE} \\ (\beta_1 + \delta_1 + \delta_2 + \gamma) + \beta_2 EDUC & \text{BLACK - FEMALE} \end{cases}$$

■ 임금식의 추정 결과

$$WAGE = \beta_1 + \beta_2 EDUC + \delta_1 BLACK + \delta_2 FEMALE + \gamma (BLACK \times FEMALE) + e$$

Table 7.4 Wage Equation with Race and Gender

Variable	Coefficient	Std. Error	t-Statistic	Prob.
<i>C</i>	-3.2303	0.9675	-3.3388	0.0009
<i>EDUC</i>	1.1168	0.0697	16.0200	0.0000
<i>BLACK</i>	-1.8312	0.8957	-2.0444	0.0412
<i>FEMALE</i>	-2.5521	0.3597	-7.0953	0.0000
<i>BLACK</i> × <i>FEMALE</i>	0.5879	1.2170	0.4831	0.6291
$R^2 = 0.2482$		$SSE = 29307.71$		

▪ 임금 결정에 인종별, 성별 격차가 존재하는지에 대한 검정

$$WAGE = \beta_1 + \beta_2 EDUC + \delta_1 BLACK + \delta_2 FEMALE + \gamma(BLACK \times FEMALE) + e$$

$$H_0 : \delta_1 = 0, \delta_2 = 0, \gamma = 0$$

$$H_1 : \delta_1 \neq 0, \text{ or } \delta_2 \neq 0, \text{ or } \gamma \neq 0$$

- (귀무가설이 옳다는 가정하의) 제약된 식의 추정결과

$$\widehat{WAGE} = -4.9122 + 1.1385 EDUC, \quad SSE_R = 31093$$

- 유의수준 1%, 자유도 = $N - K = 1,000 - 5 = 995$

$$F = \frac{(SSE_R - SSE_U) / J}{SSE_U / (N - K)} = \frac{(31093 - 29308) / 3}{29308 / 995} = 20.2$$

$$F = 20.2 > F_{(0.99; 3, 995)} = 3.80$$

- 귀무가설 기각, 따라서 인종별 및 성별 임금격차가 존재함

7.3.2 여러 개의 범주를 가지는 질적인 요소

- 지역: 동, 서, 남, 북
- 학력: 고졸 미만, 고졸, 대졸, 대학원졸업

$$E_0 = \begin{cases} 1 & \text{less than high school} \\ 0 & \text{otherwise} \end{cases} \quad E_1 = \begin{cases} 1 & \text{high school diploma} \\ 0 & \text{otherwise} \end{cases}$$

$$E_2 = \begin{cases} 1 & \text{college degree} \\ 0 & \text{otherwise} \end{cases} \quad E_3 = \begin{cases} 1 & \text{postgraduate degree} \\ 0 & \text{otherwise} \end{cases}$$

$$WAGE = \beta_1 + \beta_2 EXP + \delta_1 E_1 + \delta_2 E_2 + \delta_3 E_3 + e$$

- 네 가지 더미변수를 모두 포함시키면,

더미변수의 함정(dummy variable trap)에 빠지게 됨

- $E_0 + E_1 + E_2 + E_3 = 1$ 이므로, 더미변수들 사이의 **다중공선성**을 피하기 위해 임의의 한 더미변수를 제외시켜야 함

- 학력: 고졸 미만, 고졸, 대졸, 대학원졸업

$$E_0 = \begin{cases} 1 & \text{less than high school} \\ 0 & \text{otherwise} \end{cases} \quad E_1 = \begin{cases} 1 & \text{high school diploma} \\ 0 & \text{otherwise} \end{cases}$$

$$E_2 = \begin{cases} 1 & \text{college degree} \\ 0 & \text{otherwise} \end{cases} \quad E_3 = \begin{cases} 1 & \text{postgraduate degree} \\ 0 & \text{otherwise} \end{cases}$$

$$WAGE = \beta_1 + \beta_2 EXP + \delta_1 E_1 + \delta_2 E_2 + \delta_3 E_3 + e$$

$$E(WAGE) = \begin{cases} (\beta_1 + \delta_3) + \beta_2 EXP & \text{postgraduate degree} \\ (\beta_1 + \delta_2) + \beta_2 EXP & \text{college degree} \\ (\beta_1 + \delta_1) + \beta_2 EXP & \text{high school diploma} \\ \beta_1 + \beta_2 EXP & \text{less than high school} \end{cases}$$

■ 지역별 임금격차가 존재하는지에 대한 검정

$$WAGE = \beta_1 + \beta_2 EDUC + \delta_1 SOUTH + \delta_2 MIDWEST + \delta_3 WEST + e$$

- 더미변수의 함정 피하기 위해 *NORTHEAST* 제외

$$NORTHEAST + SOUTH + MIDWEST + WEST = 1$$

$$E(WAGE) = \begin{cases} (\beta_1 + \delta_3) + \beta_2 EDUC & WEST \\ (\beta_1 + \delta_2) + \beta_2 EDUC & MIDWEST \\ (\beta_1 + \delta_1) + \beta_2 EDUC & SOUTH \\ \beta_1 + \beta_2 EDUC & NORTHEAST \end{cases}$$

■ 지역별 임금격차가 존재하는지에 대한 검정

$$WAGE = \beta_1 + \beta_2 EDUC + \delta_1 SOUTH + \delta_2 MIDWEST + \delta_3 WEST + e$$

- 가설 $H_0 : \delta_1 = 0, \delta_2 = 0, \delta_3 = 0$

$$H_1 : \delta_1 \neq 0, \text{ or } \delta_2 \neq 0, \text{ or } \delta_3 \neq 0$$

- 검정통계량

$$F = \frac{(SSE_R - SSE_U) / J}{SSE_U / (N - K)} \sim F(1 - \alpha; J, N - K)$$

■ 임금식의 추정 결과

$$WAGE = \beta_1 + \beta_2 EDUC + \delta_1 BLACK + \delta_2 FEMALE + \gamma(BLACK \times FEMALE) + e$$

Table 7.4 Wage Equation with Race and Gender

Variable	Coefficient	Std. Error	t-Statistic	Prob.
<i>C</i>	-3.2303	0.9675	-3.3388	0.0009
<i>EDUC</i>	1.1168	0.0697	16.0200	0.0000
<i>BLACK</i>	-1.8312	0.8957	-2.0444	0.0412
<i>FEMALE</i>	-2.5521	0.3597	-7.0953	0.0000
<i>BLACK</i> × <i>FEMALE</i>	0.5879	1.2170	0.4831	0.6291
$R^2 = 0.2482$		$SSE = 29307.71$		

- 인종별 및 성별 임금격차가 존재한다는 전제하에서
지역별 임금격차가 추가로 존재하는지에 대한 검정

Table 7.5 Wage Equation with Regional Dummy Variables

Variable	Coefficient	Std. Error	t-Statistic	Prob.
<i>C</i>	-2.4557	1.0510	-2.3365	0.0197
<i>EDUC</i>	1.1025	0.0700	15.7526	0.0000
<i>BLACK</i>	-1.6077	0.9034	-1.7795	0.0755
<i>FEMALE</i>	-2.5009	0.3600	-6.9475	0.0000
<i>BLACK</i> × <i>FEMALE</i>	0.6465	1.2152	0.5320	0.5949
<i>SOUTH</i>	-1.2443	0.4794	-2.5953	0.0096
<i>MIDWEST</i>	-0.4996	0.5056	-0.9880	0.3234
<i>WEST</i>	-0.5462	0.5154	-1.0597	0.2895

$R^2 = 0.2535$

$SSE = 29101.3$

- <Table 7.4>와 비교

$F_{(0.95; 3, 992)} = 2.61$, 지역별 격차 없음

$$F = \frac{(SSE_R - SSE_U) / J}{SSE_U / (N - K)} = \frac{(29308 - 29101) / 3}{29101 / (1000 - 8)} = 2.35$$

7.3.3 더미변수를 이용한 두 회귀식의 등가 검정

(예) 주택의 위치가 절편 및 기울기 모두에 영향을 미친다면,

$$P_t = \beta_1 + \delta D_t + \beta_2 S_t + \gamma(S_t D_t) + e_t$$

- 두 개의 다른 주택지구에 위치한 주택의 회귀함수

519개

$$E(P_t) = \begin{cases} (\beta_1 + \delta) + (\beta_2 + \gamma)S_t = \alpha_1 + \alpha_2 S_t & \text{desirable neighborhood data} \\ \beta_1 + \beta_2 S_t & \text{other neighborhood data} \end{cases}$$

481개

$$P_t = \alpha_1 + \alpha_2 S_t + e_t$$

$$P_t = \beta_1 + \beta_2 S_t + e_t$$

- If $\delta=0$ then $\alpha_1 = \beta_1$, and if $\gamma=0$, then $\alpha_2 = \beta_2$.
 $\Rightarrow$ 1,000개의 표본으로 $P_t = \beta_1 + \beta_2 S_t + e_t$ 을 추정하면 됨 (위치와 무관)
- 위의 두 회귀의 등가성에 대한 검정 $\Rightarrow$ **Chow test** (F-검정)

- 남부지역의 임금회귀식이 타지역의 임금식과 차이나는가?
- 차이가 없다면, 모든 지역의 자료를 통합하여
더미변수 없는 다음 식으로 한 번 추정하면 됨

$$WAGE = \beta_1 + \beta_2 EDUC + \delta_1 BLACK + \delta_2 FEMALE + \gamma(BLACK \times FEMALE) + e$$

■ 남부지역의 임금회귀식이 타지역의 임금식과 차이나는가?

- 차이가 있다면, 다음 (1)식을 추정하는 것은 (2)식과 (3)식을 각각 추정하는 것과 같음

$$(1) \quad WAGE = \beta_1 + \beta_2 EDUC + \delta_1 BLACK + \delta_2 FEMALE + \gamma(BLACK \times FEMALE) +$$

$$\theta_1 SOUTH + \theta_2(EDUC \times SOUTH) + \theta_3(BLACK \times SOUTH) +$$

$$\theta_4(FEMALE \times SOUTH) + \theta_5(BLACK \times FEMALE \times SOUTH) + e$$

$$(2) \quad E(WAGE) = \begin{cases} \beta_1 + \beta_2 EDUC + \delta_1 BLACK + \delta_2 FEMALE + \gamma(BLACK \times FEMALE) & SOUTH = 0 \end{cases}$$

$$(3) \quad E(WAGE) = \begin{cases} (\beta_1 + \theta_1) + (\beta_2 + \theta_2) EDUC + (\delta_1 + \theta_3) BLACK + (\delta_2 + \theta_4) FEMALE + (\gamma + \theta_5)(BLACK \times FEMALE) & SOUTH = 1 \end{cases}$$

■ 남부지역의 임금회귀식이 타지역의 임금식과 차이나는가?

Table 7.6 Comparison of Fully Interacted to Separate Models

Variable	(1) Full sample		(2) Non-south		(3) South	
	Coefficient	Std. Error	Coefficient	Std. Error	Coefficient	Std. Error
<i>C</i>	-3.5775	1.1513	-3.5775	1.2106	-2.2752	1.5550
<i>EDUC</i>	1.1658	0.0824	1.1658	0.0866	0.9741	0.1143
<i>BLACK</i>	-0.4312	1.3482	-0.4312	1.4176	-2.1756	1.0804
<i>FEMALE</i>	-2.7540	0.4257	-2.7540	0.4476	-1.8421	0.5896
<i>BLACK</i> × <i>FEMALE</i>	0.0673	1.9063	0.0673	2.0044	0.6101	1.4329
<i>SOUTH</i>	1.3023	2.1147				
<i>EDUC</i> × <i>SOUTH</i>	-0.1917	0.1542				
<i>BLACK</i> × <i>SOUTH</i>	-1.7444	1.8267				
<i>FEMALE</i> × <i>SOUTH</i>	0.9119	0.7960				
<i>BLACK</i> × <i>FEMALE</i> × <i>SOUTH</i>	0.5428	2.5112				
<i>SSE</i>	29012.7		22031.3		6981.4	
<i>N</i>	1000		685		315	

$$(\hat{\delta}_1 + \hat{\theta}_3) = -0.4312 - 1.7444 = -2.1756$$

$$(\hat{\delta}_2 + \hat{\theta}_4) = -2.7540 + 0.9119 = -1.8421$$

■ 남부지역의 임금회귀식이 타지역의 임금식과 차이나는가?

$$\begin{aligned} WAGE = & \beta_1 + \beta_2 EDUC + \delta_1 BLACK + \delta_2 FEMALE + \gamma (BLACK \times FEMALE) + \\ & \theta_1 SOUTH + \theta_2 (EDUC \times SOUTH) + \theta_3 (BLACK \times SOUTH) + \\ & \theta_4 (FEMALE \times SOUTH) + \theta_5 (BLACK \times FEMALE \times SOUTH) + e \end{aligned}$$

$$H_0 : \theta_1 = \theta_2 = \theta_3 = \theta_4 = \theta_5 = 0$$

$$F = \frac{(SSE_R - SSE_U) / J}{SSE_U / (N - K)} = \frac{(29307.7 - 29012.7) / 5}{29012.7 / 990} = 2.0132$$

$$F_{(0.90; 5, 990)} = 1.85$$

- 10% 유의수준에서 남부지역과 타지역의 임금방정식이 동일하다는 가설을 기각함

7.3.4 시간에 대한 통제

- 계절 더미변수 예

$AUG=1$ if month is August, $AUG=0$ otherwise

$SUMMER=1$ if month = June, July or August; $SUMMER=0$ otherwise

- 연도 더미변수

$D98=1$ if year = 1998; $D98 = 0$ otherwise

- 제도 더미변수

- (예) 투자세액공제가 실시된 시기: 1962-65, 1970-86

$$ITC = \begin{cases} 1 & 1962-1965, 1970-1986 \\ 0 & otherwise \end{cases}$$

$$INV_t = \beta_1 + \delta ITC_t + \beta_2 GNP_t + \beta_3 GNP_{t-1} + e_t$$

■ 질적효과의 존재에 대한 검정

<단일 질적효과의 존재에 대한 검정>

- 해당 더미변수의 유의성에 대한 t-검정 실시

(예) $INV_t = \beta_1 + \delta ITC_t + \beta_2 GNP_t + \beta_3 GNP_{t-1} + e_t$

투자세액공제제도의 효율성은 $H_0 : \delta = 0$ 을 검정하면 됨

<복수의 질적효과의 존재에 대한 결합검정>

- 모든 더미변수가 동시에 유의한지에 대한 F-검정 실시

(예) $WAGE = \beta_1 + \beta_2 EXP + \delta_1 RACE + \delta_2 SEX + \gamma(RACE \times SEX) + e$

인종별 혹은 성별 임금격차 모두 존재하지 않는다는

귀무가설 $H_0 : \delta_1 = 0, \delta_2 = 0, \gamma = 0$ 을 검정하면 됨

7.4 변수 사이의 상호작용

- 두 가지 상이한 변수의 곱이 상호작용 변수

$$PIZZA = \beta_1 + \beta_2 AGE + \beta_3 INCOME + e$$

$PIZZA$: 피자에 대한 연간 지출액, AGE : 연령, $INCOME$: 개인 소득

- 모수의 의미 $\frac{\partial E(PIZZA)}{\partial AGE} = \beta_2$ $\frac{\partial E(PIZZA_i)}{\partial INCOME} = \beta_3$
- β_3 는 한계소비성향
- 만약 한계소비성향이 나이에 따라 변화한다면,
즉 소득이 미치는 영향이 개인의 연령에 의존한다면
상호작용 변수 ($AGE \times INCOME$) 가 모형에 추가되어야 함

- 상호작용 변수가 추가된 모형

$$PIZZA = \beta_1 + \beta_2 AGE + \beta_3 INCOME + \beta_4 (AGE \times INCOME) + e$$

- $\frac{\partial E(PIZZA)}{\partial AGE} = \beta_2 + \beta_4 INCOME$
 - 나이의 영향은 소득에 따라 달라짐
- $\frac{\partial E(PIZZA)}{\partial INCOME} = \beta_3 + \beta_4 AGE$
 - 한계소비성향은 나이에 따라 달라짐
- 예상 부호: $\beta_2 < 0$, $\beta_3 > 0$, $\beta_4 < 0$

- 피자 지출액 자료(표본 수는 40명)를 이용한 추정 결과

Table 7.7 Pizza Expenditure Data

<i>PIZZA</i>	<i>INCOME</i>	<i>AGE</i>
109	15000	25
0	30000	45
0	12000	20
108	20000	28
220	15000	25

$$\widehat{PIZZA} = 342.88 - 7.58AGE + 0.0024INCOME$$

(t) (-3.27) (3.95)

$$\widehat{PIZZA} = 161.47 - 2.98AGE + 0.009INCOME - 0.00016(AGE \times INCOME)$$

(t) (-0.89) (2.47) (-1.85)

The AUTOREG Procedure

Model r10_1

Dependent Variable pizza

Ordinary Least Squares Estimates

SSE	635636.721	DFE	37
MSE	17179	Root MSE	131.07010
SBC	511.521842	AIC	506.455204
Regress R-Square	0.3293	Total R-Square	0.3293
Durbin-Watson	0.9009		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > t
Intercept	1	342.8848	72.3434	4.74	<.0001
age	1	-7.5756	2.3170	-3.27	0.0023
income	1	0.002382	0.000604	3.95	0.0003

The AUTOREG Procedure

Model r10_2

Dependent Variable pizza

Ordinary Least Squares Estimates

SSE	580608.65	DFE	36
MSE	16128	Root MSE	126.99613
SBC	511.588711	AIC	504.833194
Regress R-Square	0.3873	Total R-Square	0.3873
Durbin-Watson	0.9320		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > t
Intercept	1	161.4654	120.6634	1.34	0.1892
age	1	-2.9774	3.3521	-0.89	0.3803
income	1	0.009074	0.003670	2.47	0.0183
age*income	1	-0.000160	0.0000867	-1.85	0.0730

■ 추정결과

$$\widehat{PIZZA} = 342.88 - 7.58AGE + 0.0024INCOME$$

(t) (-3.27) (3.95)

$$\widehat{PIZZA} = 161.47 - 2.98AGE + 0.009INCOME - 0.00016(AGE \times INCOME)$$

(t) (-0.89) (2.47) (-1.85)

- 두 추정결과를 비교하면, AGE 의 유의성이 크게 하락
 $\Rightarrow AGE$ 가 그 자체로 피자수요에 영향을 미치는 것이 아니라
 소득과의 상호작용을 통해 피자수요에 영향을 미친다는 것을
 의미함,
 즉 나이는 한계지출성향에 영향을 미친다는 의미

- 소득이 \$25,000인 사람과 \$90,000인 사람의 경우
피자지출에 대한 나이의 한계효과를 추정할 수 있음

$$\begin{aligned}\frac{\partial E(\text{PIZZA})}{\partial \text{AGE}} &= b_2 + b_4 \text{INCOME} \\ &= -2.98 - 0.00016 \text{INCOME} \\ &= \begin{cases} -6.98 & \text{for } \text{INCOME} = \$25,000 \\ -17.40 & \text{for } \text{INCOME} = \$90,000 \end{cases}\end{aligned}$$

- 소득이 \$25,000인 경우(저소득층),
나이가 한 살 많아지면 피자지출은 연간 \$6.98 감소
- 소득이 \$90,000인 경우(고소득층),
나이가 한 살 많아지면 피자지출은 연간 \$17.40 감소

<과제>

7.10

7.16

**Eviews output을 출력하고,
출력물의 빈 여백에 간단하게 답을 적으시오.**

※ 참고: 필요한 data는 WILEY 교과서 홈페이지에 있음

<http://principlesofeconometrics.com/>