

Part.02  
회귀분석

# I 다중 선형 회귀분석 (회귀 계수/모델에 대한 검정)

FASTCAMPUS  
ONLINE

머신러닝과 데이터분석 A-Z

강사. 이경택

# I 다중 선형 회귀분석

## ■ 다중 선형 회귀 계수 해석

- 광고미디어 예에 다중선형회귀를 적용하면 아래와 같음
- 신문 광고의 경우 단순선형회귀에서는 출력 변수인 매출과 연관이 있었지만, 다중선형회귀에서는 p-value가 0.86으로 높아 매출에 유의미한 영향을 미치지 못함
- 신문 광고의 단순선형회귀에서 TV 광고와 라디오 광고의 영향력을 무시했기 때문에 생긴 결과

$$\hat{Y} = \hat{\beta}_0 + \overset{\text{TV}}{\hat{\beta}_1 X_1} + \overset{\text{radio}}{\hat{\beta}_2 X_2} + \overset{\text{newspaper}}{\hat{\beta}_3 X_3}$$

	Coefficient	Std. error	t-statistic	p-value
Intercept	2.939	0.3119	9.42	< 0.0001
TV	<u>0.046</u>	0.0014	32.81	< 0.0001
radio	<u>0.189</u>	0.0086	21.89	< 0.0001
newspaper	<u>-0.001</u>	0.0059	-0.18	<span style="border: 1px solid red;">0.8599</span>

# I 다중 선형 회귀분석

## ■ 다중 선형 회귀 계수 검정

### ■ $\hat{\beta}_p$ 의 검정

귀무가설 :  $B_p = 0$  (회귀계수는 0이다, 즉 변수의 설명력이 없다)

대립가설 :  $B_p \neq 0$  (회귀계수는 0이 아니다, 즉 변수의 설명력이 존재 한다.)

$$t = \frac{\widehat{B}_p}{\frac{s}{\sqrt{S_{xx}}}} = \frac{\widehat{B}_p}{s.e(\widehat{B}_p)}$$

	Coefficient	Std. error	t-statistic	p-value
Intercept	2.939	0.3119	9.42	< 0.0001
TV	0.046	0.0014	32.81	< 0.0001
radio	0.189	0.0086	21.89	< 0.0001
newspaper	-0.001	0.0059	-0.18	0.8599

# I 다중 선형 회귀분석

## ■ 다중 선형 회귀 모델 검정

귀무가설 :  $B_1 = B_2 \dots B_p = 0$  ( 모든 회귀계수는 0이다, 즉 변수의 설명력이 하나도 존재 하지 않는다)

대립가설 : 하나의 회귀계수라도 0이 아니다. (즉 설명력이 있는 변수가 존재 한다.)

- F검정을 통해서 검정

$$F = \frac{V_1/k_1}{V_2/k_2} \sim F(k_1, k_2)$$

두 확률 변수  $V_1, V_2$ 가 서로 독립인 카이제곱 분포를 따른다고 할 때 확률 변수  $F$ 는  $F$ 분포를 따른다고 한다.  $F$ 검정과 분산분석등에서 주로 사용됨.

Source of Variation	Sum of Squares	DF	Mean Square	F-Statistics	P-value
Regression	SSR	P	MSR(SSR/p)	MSR/MSE	P-value
Error	SSE	N-p-1	MSE(SSE/(N-p-1))		
Total	SST	N-1			

분자, 분모 모두 제곱합의 형태 이기 때문에 F검정사용

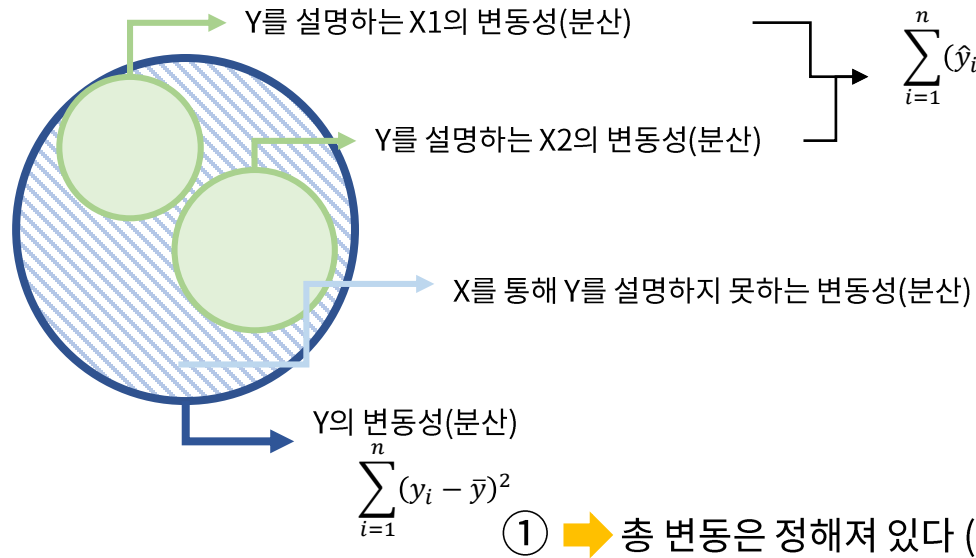
# I 다중 선형 회귀분석

## ■ 다중 선형 회귀 모델 검정

귀무가설 :  $B_1 = B_2 \dots B_p = 0$  ( 모든 회귀계수는 0이다, 즉 변수의 설명력이 하나도 존재 하지 않는다)

대립가설 : 하나의 회귀계수라도 0이 아니다. (즉 설명력이 있는 변수가 존재 한다.)

➡ 기각 하기 너무 쉬운 가설. 변수가 추가 되면 추가 될수록 기각하기 쉬워진다.



➡ ② 제곱합의 형태 이기 때문에 변수가 추가되면 추가될수록 자연스레 증가한다.

③ 제곱합의 형태로 검정을 하는 F검정의 특성상 변수가 추가되면 자연스레 기각하기 쉬워진다. ( $R^2$ 도 커짐)

Part.02  
회귀분석

# I 다중공선성

FASTCAMPUS  
ONLINE

머신러닝과 데이터분석 A-Z

강사. 이경택