

Part.02
회귀분석

| 회귀계수에 대한 검정

FASTCAMPUS
ONLINE

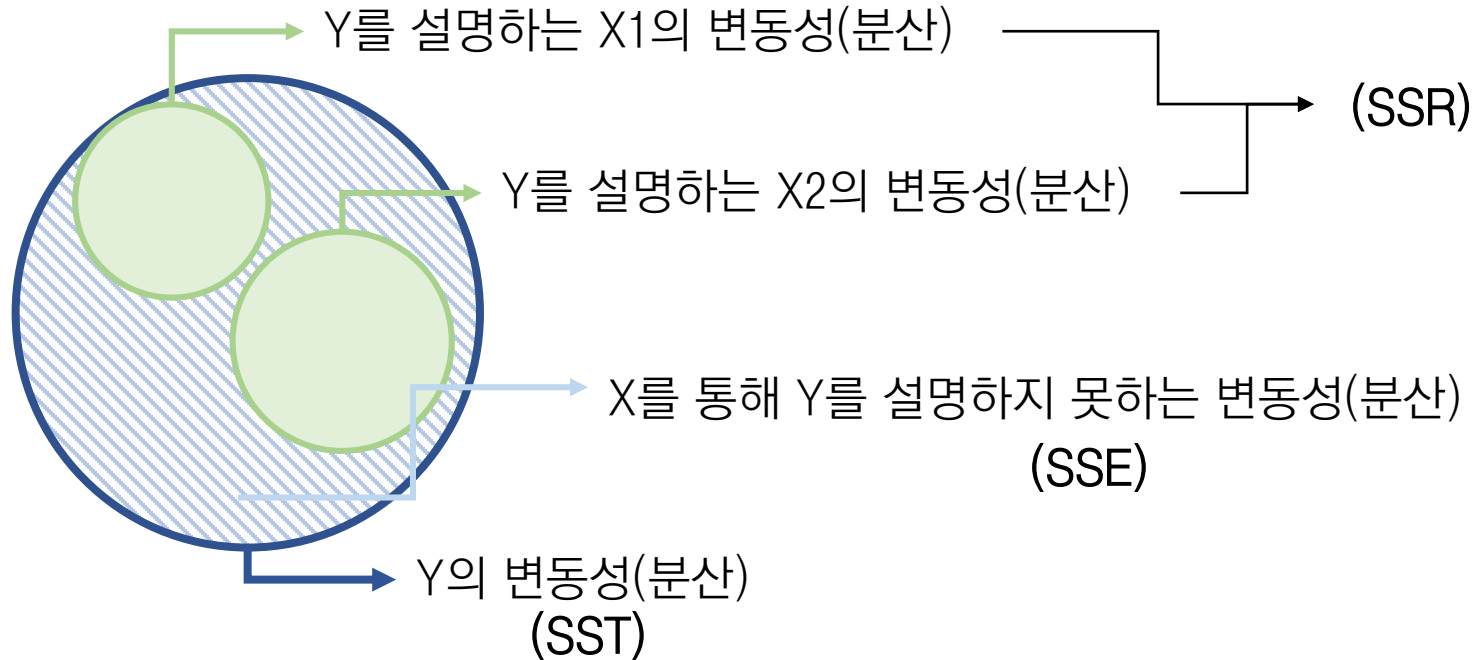
머신러닝과 데이터분석 A-Z

강사. 이경택

I 회귀계수에 대한 검정

■ 선형 회귀의 정확도 평가

- 회귀 분석은 결국 Y의 변동성을 얼마나 독립변수가 잘 설명하느냐가 중요
- 변수가 여러 개일 때 각각 Y를 설명하는 변동성이 크면 좋은 변수 -> p-value자연스레 낮아짐



$$R^2 = \frac{SSR}{SST}$$

I 회귀계수에 대한 검정

- 단순 선형 회귀분석의 검정

- $\hat{\beta}_1$ 의 표준오차

$$S.E(\widehat{B}_1) = \frac{\sigma}{\sqrt{S_{xx}}}$$

오차의 표준편차가 알려져 있지 않은 경우, $s = \sqrt{\frac{SSE}{n-2}}$ 를 대입하여 추정

- $\hat{\beta}_1$ 의 표본분포

$$t = \frac{\frac{(\widehat{B}_1 - B_1)}{s}}{\sqrt{S_{xx}}} \sim t(n-2)$$

I 회귀계수에 대한 검정

■ 단순 선형 회귀분석의 검정

■ $\hat{\beta}_1$ 의 검정

귀무가설: $B_1 = 0$ (회귀계수는 0이다, 즉 변수의 설명력이 없다)

대립가설: $B_1 \neq 0$ (회귀계수는 0이 아니다, 즉 변수의 설명력이 존재 한다.)

$$t = \frac{\widehat{B}_1}{\frac{s}{\sqrt{S_{xx}}}} = \frac{\widehat{B}_1}{s.e(\widehat{B}_1)}$$

회귀 계수

회귀 계수의 표준 오차

Simple regression of sales on radio

회귀 계수의 의성을 판단하는 통계치

| | Coefficient | Std. error | t-statistic | p-value |
|-----------|-------------|------------|-------------|----------|
| Intercept | 9.312 | 0.563 | 16.54 | < 0.0001 |
| radio | 0.203 | 0.020 | 9.92 | < 0.0001 |

Simple regression of sales on newspaper

| | Coefficient | Std. error | t-statistic | p-value |
|-----------|-------------|------------|-------------|----------|
| Intercept | 12.351 | 0.621 | 19.88 | < 0.0001 |
| newspaper | 0.055 | 0.017 | 3.30 | < 0.0001 |

I 회귀계수에 대한 검정

- 단순 선형 회귀분석의 검정

- $\hat{\beta}_1$ 의 신뢰구간

B_1 의 100(1- α)% 신뢰구간:
$$\widehat{B}_1 \pm t_{\frac{\alpha}{2}}(n-2) \times \frac{s}{\sqrt{S_{xx}}}$$

- $\hat{\beta}_0$ 의 신뢰구간

B_0 의 100(1- α)% 신뢰구간:
$$\widehat{B}_0 \pm t_{\frac{\alpha}{2}}(n-2) \times s \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}}$$

Part.02
회귀분석

I 다중 선형 회귀분석

FASTCAMPUS
ONLINE

머신러닝과 데이터분석 A-Z

강사. 이경택