

9강. 자료진단

◆ 담당교수 : 김성수 교수

■ 주요용어

용어	해설
잔차벡터	<p>잔차벡터 $\mathbf{e} = \mathbf{Y} - \hat{\mathbf{Y}}$ 이다.</p> $\hat{\mathbf{Y}} = \mathbf{X}\mathbf{b}$ <p>여기서 $\mathbf{b} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$</p> $= \mathbf{H}\mathbf{Y}$ <p>이므로, \mathbf{e}는 다음과 같다.</p> $\mathbf{e} = \mathbf{Y} - \hat{\mathbf{Y}}$ $= \mathbf{Y} - \mathbf{H}\mathbf{Y}$ $= (\mathbf{I} - \mathbf{H})\mathbf{Y}$
스튜던트화 잔차 (studentized residual)	<p>$\text{Var}(\mathbf{e}) = \sigma^2(\mathbf{I} - \mathbf{H})$ 에서 i 번째 잔차인 e_i의 분산 $\text{Var}(e_i) = \sigma^2(1 - h_{ii})$ 이 된다.</p> <p>e_i를 잔차의 표준오차인 $\sigma\sqrt{1 - h_{ii}}$로 나누어 표준화시킨 잔차</p> $r_i = \frac{e_i}{\sqrt{MSE(1 - h_{ii})}}$ <p>를 표준화잔차(standardized residual)라 한다. 여기서</p> $t_i = \frac{e_i}{\sqrt{MSE_{(i)}(1 - h_{ii})}}$ <p>를 스튜던트화 잔차(studentized residuals)라 한다.</p>
영향력 있는 관측값 (influential observations)	<p>회귀모형에서 관측값이 제거될 때 중요한 변화를 가져오는 관측값을 영향력있는 관측값이라고 부른다. 대표적인 통계량이 Cook의 통계량 D_i이다.</p> $D_i = \frac{(\mathbf{b}_{(i)} - \mathbf{b})'(\mathbf{X}\mathbf{X}')(\mathbf{b}_{(i)} - \mathbf{b})}{(k+1)MSE}$
특이점(outlier)	<p>다른 관측값들과는 달리 주어진 모형을 따르지 않는 관측값을 특이점(outlier) 또는 이상점이라 부른다.</p> <p>특이점 검정에 이용되는 검정통계량은 스튜던트화 잔차 t_i로서, 일반적으로 t_i의 값이 2 또는 3보다 큰 경우를 특이점으로 판단한다.</p>

■ 연습문제

1. 표준화잔차(standardized residual)를 구하는 식은 ?

$$r_i = \frac{e_i}{\sqrt{MSE(1-h_{ii})}}$$

정답 및 해설 :

2. 스튜던트화 잔차(studentized residuals)를 구하는 식은 ?

$$t_i = \frac{e_i}{\sqrt{MSE_{(i)}(1-h_{ii})}}$$

정답 및 해설 :

3. 잔차분석을 하고자 한다. 함수(a)는 ?

```
> forbes.res = ( a )(forbes.lm)
> names(forbes.res)
[1] "std.dev" "hat"      "std.res" "stud.res" "cooks"
[6] "dfits"   "correlation" "std.err" "cov.scaled" "cov.unscaled"
```

정답 및 해설 : ls.diag

■ 참고사이트

- 강명욱,김영일,안철환,이용구, 『회귀분석』, 을곡출판사, 1996.
- 박성현, 『회귀분석』 (제3판), 민영사, 2007.
- Faraway, J.J. (2002), Practical Regression and Anova Using R, (www.google.com에서 검색 후, pdf 파일로 다운받을 수 있음)
- Peter Dalgaard (2005), Introductory Statistics with R, Springer, (www.google.com에서 검색 후, pdf 파일로 다운받을 수 있음)

- R 사이트 <http://www.r-project.org/>
- R Studio 사이트 <https://www.rstudio.com/>