

12강. 일반선형모형 (1)

◆ 담당교수 : 김성수 교수

■ 주요용어

용어	해설
일반화선형모형	일반화선형모형은 세 가지 구성성분 즉, 1)반응변수의 분포 2)선형예측자 3)연결함수로 이루어지며 이들 세 성분을 확정함으로써 구체적인 분석모형이 설정된다. 이러한 모형구성방법을 통하여 일반화선형모형은 고전적인 회귀모형인 선형회귀모형을 확장한 모형을 말한다.
선형예측자	선형예측자(linear predictor)는 모형에 있는 설명변수들의 선형결합을 말한다. η 로 표기하며 모형에 포함될 설명변수가 결정되면 해당 설명변수의 효과를 크기를 나타내는 미지의 모수 β 들과 결합하여 선형 식으로 표시된다.
연결함수	일반화선형모형에서 반응변수의 평균과 선형예측자와의 수학적 함수관계를 설정하는 연결함수 (link function)를 $g(\cdot)$ 로 표기하면 다음의 방정식이 성립한다고 가정한다. $\eta = g(\mu) = x'\beta = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_px_p$
로지스틱회귀모형	로지스틱회귀모형은 로짓함수를 연결함수로 설정하는 모형으로 다음과 같다. $\eta = \text{logit}(\pi) = \log\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1x_1 + \beta_2x_2$ 여기서 $\pi = \mu = E(Y)$ 이다.

■ 연습문제

1. 일반화선형모형은 세 가지 구성성분은 ?

정답 및 해설 : 반응변수의 분포, 선형예측자, 연결함수

2. 다음 자료에서 occur를 반응변수(1=yes, 0=no)로 하여 로지스틱 회귀모형을 적합하고자 한다. (a)에 들어갈 옵션은 ?

```
> head(glider,3)
  p_no  occur  con_metric  p_size_km
1  1  1  0.650  130.9
2  2  0  0.610  104.1
3  3  0  0.744  132.3
> #모형적합
> logit_m1 <- glm(occur~p_size_km+con_metric, family= ( a ) (link=logit),
data=glider)
```

정답 및 해설 : binomial

3. 다음은 로지스틱 회귀모형 적합 결과이다. 적합된 로지스틱 회귀식을 쓰시오.

```
> summary(logit_m1)
Call:
glm(formula = occur ~ p_size_km+con_metric, family = binomial(link =
logit), data = glider)
Deviance Residuals:
Min 1Q Median 3Q Max
-1.4969 -0.8829 -0.3884 0.8766 2.0515
Coefficients:
Estimate Std. Error z value Pr(>|z|)
(Intercept) -3.606207 1.436391 -2.511 0.01205 *
p_size_km 0.023566 0.007462 3.158 0.00159 **
con_metric 1.631800 1.642758 0.993 0.32055

(Dispersion parameter for binomial family taken to be 1)
Null deviance : 68.994 on 49 degrees of freedom
Residual deviance : 54.661 on 47 degrees of freedom
AIC : 60.661
Number of Fisher Scoring iterations : 4
```

정답 및 해설 :

$$\log\left(\frac{\hat{\pi}}{1-\hat{\pi}}\right) = -3.606 + 0.024x_1 + 1.632x_2$$

■ 참고사이트

- 강명옥,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)
- McCullagh, P. and Nelder, J.A., Generalized Linear Models, 2nd ed,
Chapman & Hall / CRC., 1999.
- Neter et al. Applied Linear Statistical Models, 4th ed. IRWIN, 1996.
- R 사이트 <http://www.r-project.org/>
- R Studio 사이트 <https://www.rstudio.com/>