

머신러닝과 빅데이터분석(R)

13주차 로지스틱 회귀



박길식 교수



고려사이버대학교
THE CYBER UNIVERSITY OF KOREA



학습 목표

-  로지스틱 회귀분석의 원리를 이해하고 구현할 수 있다.
-  활성화 함수를 설명할 수 있다.



학습 목차

1 로지스틱 회귀 분석

2 로지스틱 회귀 실습

CHAPTER

로지스틱 회귀 분석

로지스틱 회귀

회귀모델에서 종속변수가 연속형 값이 아닌 범주형 값의 경우를 다루기 위해 개발된 통계적 방법
(Iris 데이터 셋에서 품종을 예측)

⊕ 일반적으로 종속변수가 두 가지 값만 가지는 경우에 적용

이
예

참/거짓, 성공/실패, 환자/정상, 사망/생존, 승리/패배 등

[01] 로지스틱 회귀

② 독립변수를 x , 종속변수를 l 로 표기

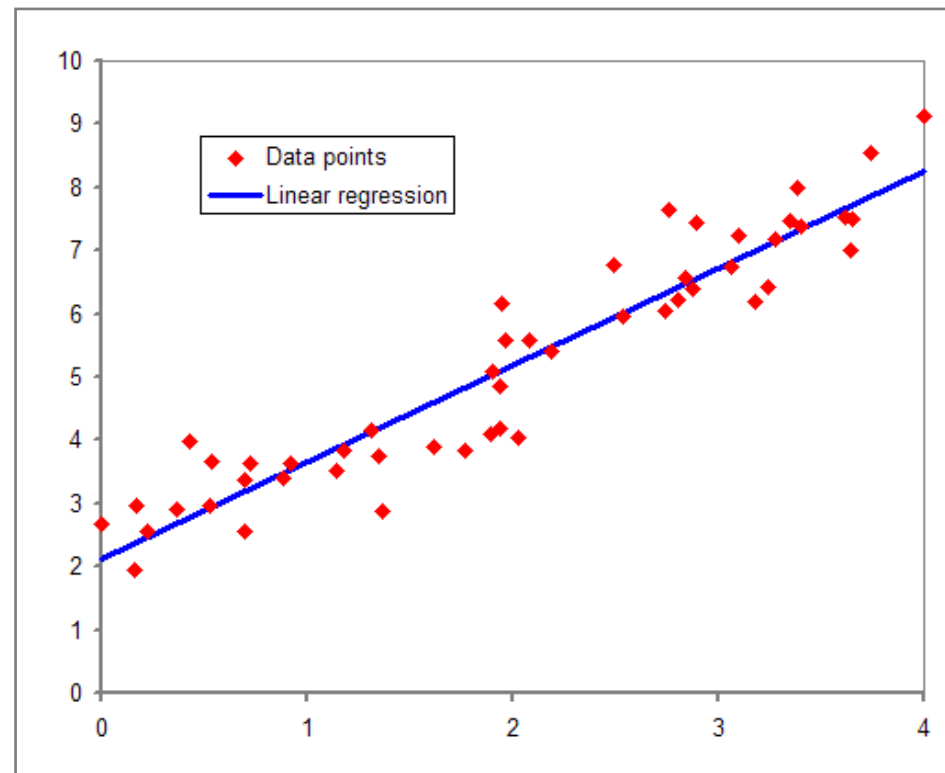
$$l = a_1x + a_0$$

② 종속변수 l 의 범위는 $[-\infty, \infty]$ 이므로
로지스틱 회귀를 모델링할 수 없음

② 해결책 : 로짓 함수(Logit Function)를
추가로 사용

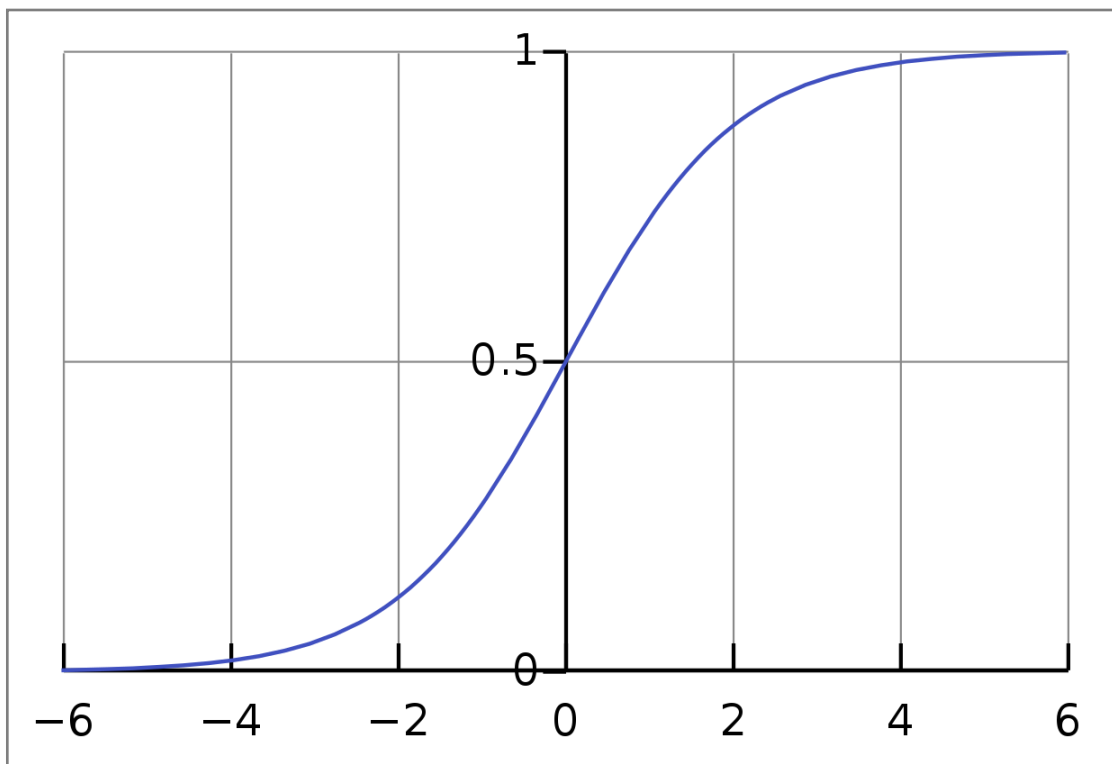
➡ 범위를 $[0,1]$ 로 축소

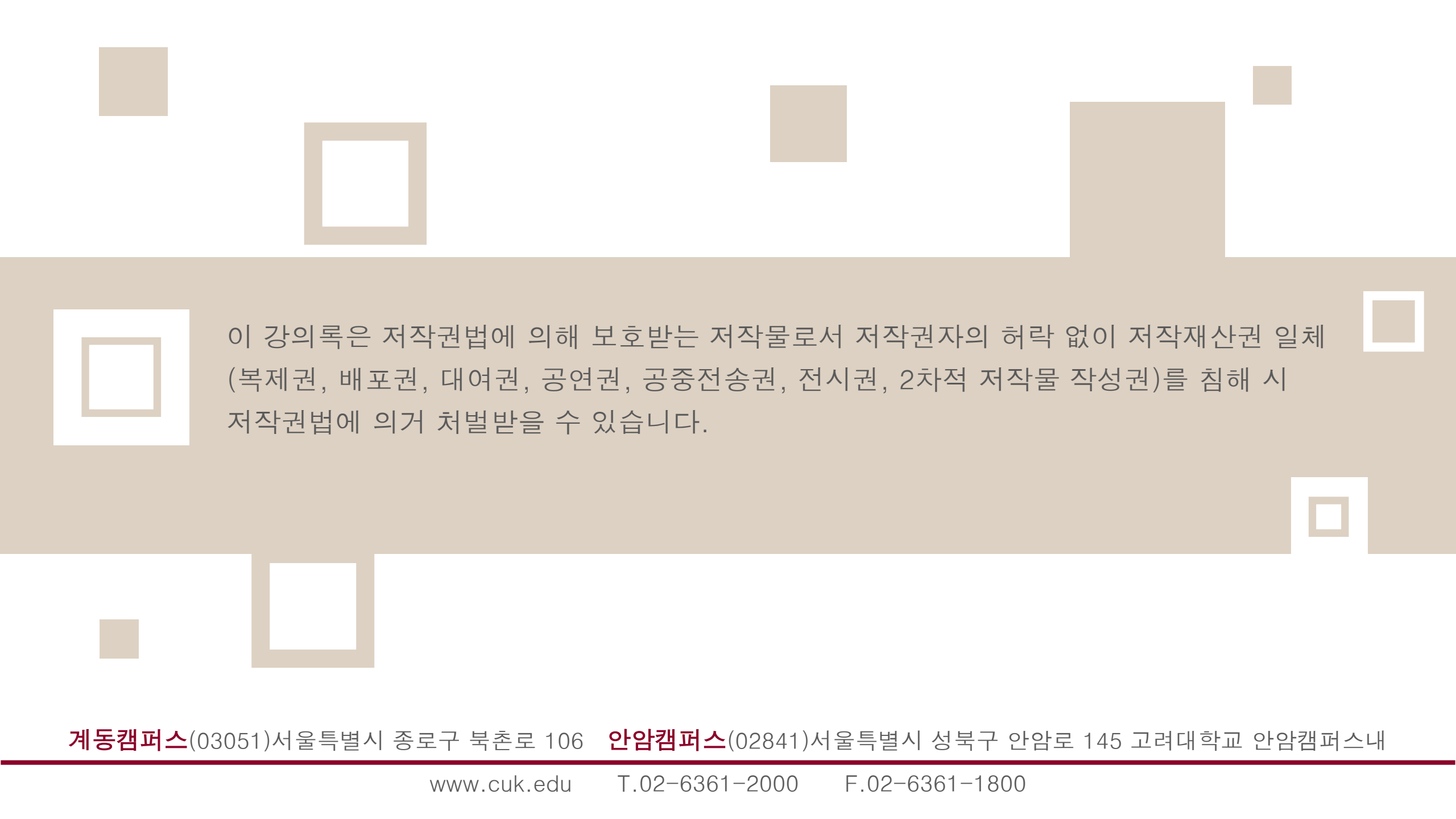
$$y = \frac{1}{1 + e^{-l}}$$



— [01] 로지스틱 회귀

- ⊕ 가로축은 l , 세로축은 y 를 나타내며 y 는 $[0,1]$ 사이로 축소되었음에 주목
- ⊕ l 을 잠복(Latent) 변수 또는 은닉(Hidden) 변수라 부름





이 강의록은 저작권법에 의해 보호받는 저작물로서 저작권자의 허락 없이 저작재산권 일체 (복제권, 배포권, 대여권, 공연권, 공중전송권, 전시권, 2차적 저작물 작성권)를 침해 시 저작권법에 의거 처벌받을 수 있습니다.