

선형회귀

하나의 독립변수만을 가지고 종속변수를 예측하기 위한 회귀 모형을 만들기 위한 것

형태 : $Y = a + b * x$

다중선형회귀

여러 개의 독립변수(x)들을 가지고 종속변수(y)를 예측하기 위한 회귀 모형

* 종속 변수는 반드시 연속형 변수

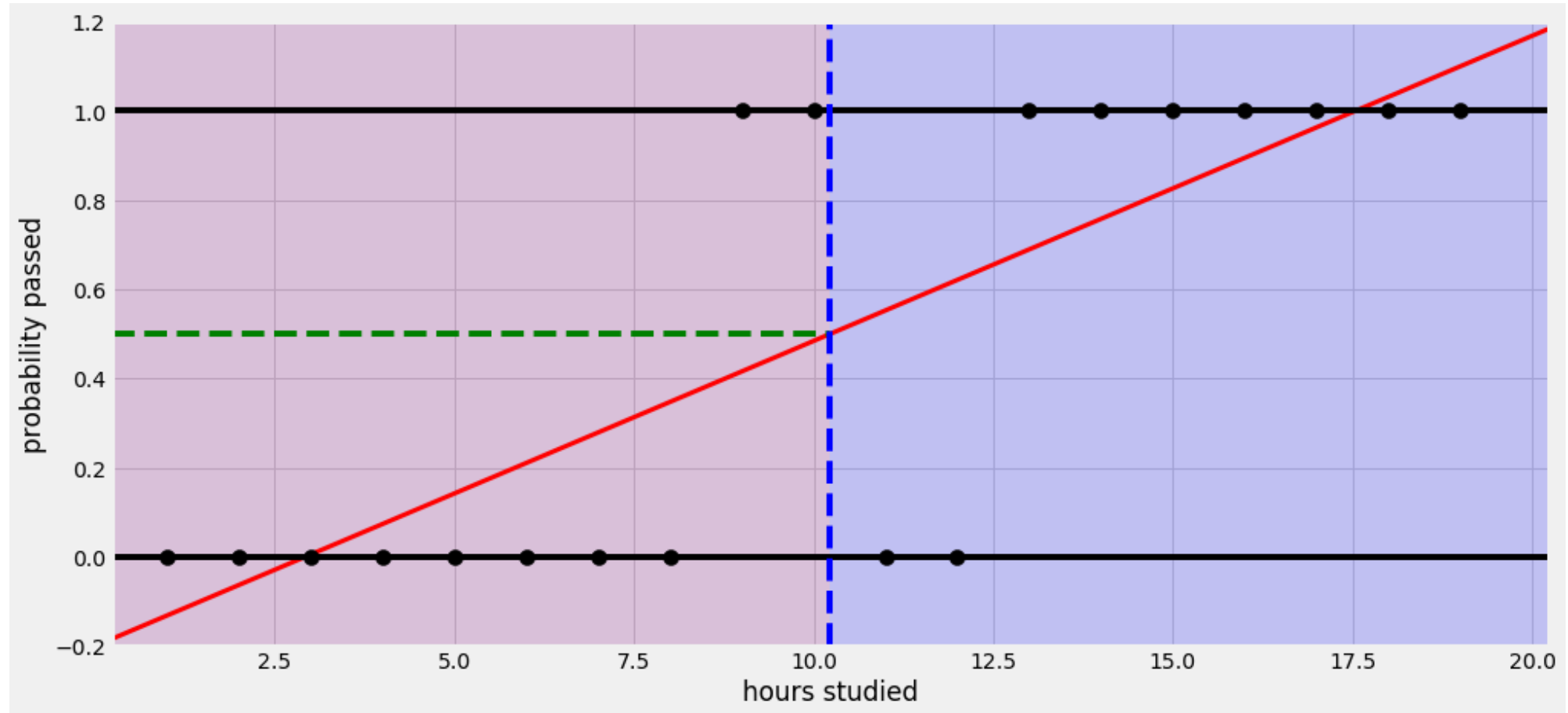
형태 : $Y = a + b_1 * x_1 + b_2 * x_2 + \dots$

:: 어떤 설명변수가 종속변수에 영향을 미치는지, 설명변수가 종속 변수에 얼마나 영향을 미치는지, 종속 변수를 예측할 하게 할수 있는 분석 방법

다중선형회귀 분석 과정

1. 회귀 모형에 포함되어질 설명변수 설정
2. 편회귀 계수 추정치를 구함
3. 유의성을 검사
4. 적합도 검사
5. wald 통계량을 이용하여 편회귀 계수 추정치들이 모집단에서 유의한지 안한지 판단

선형 회귀 예시

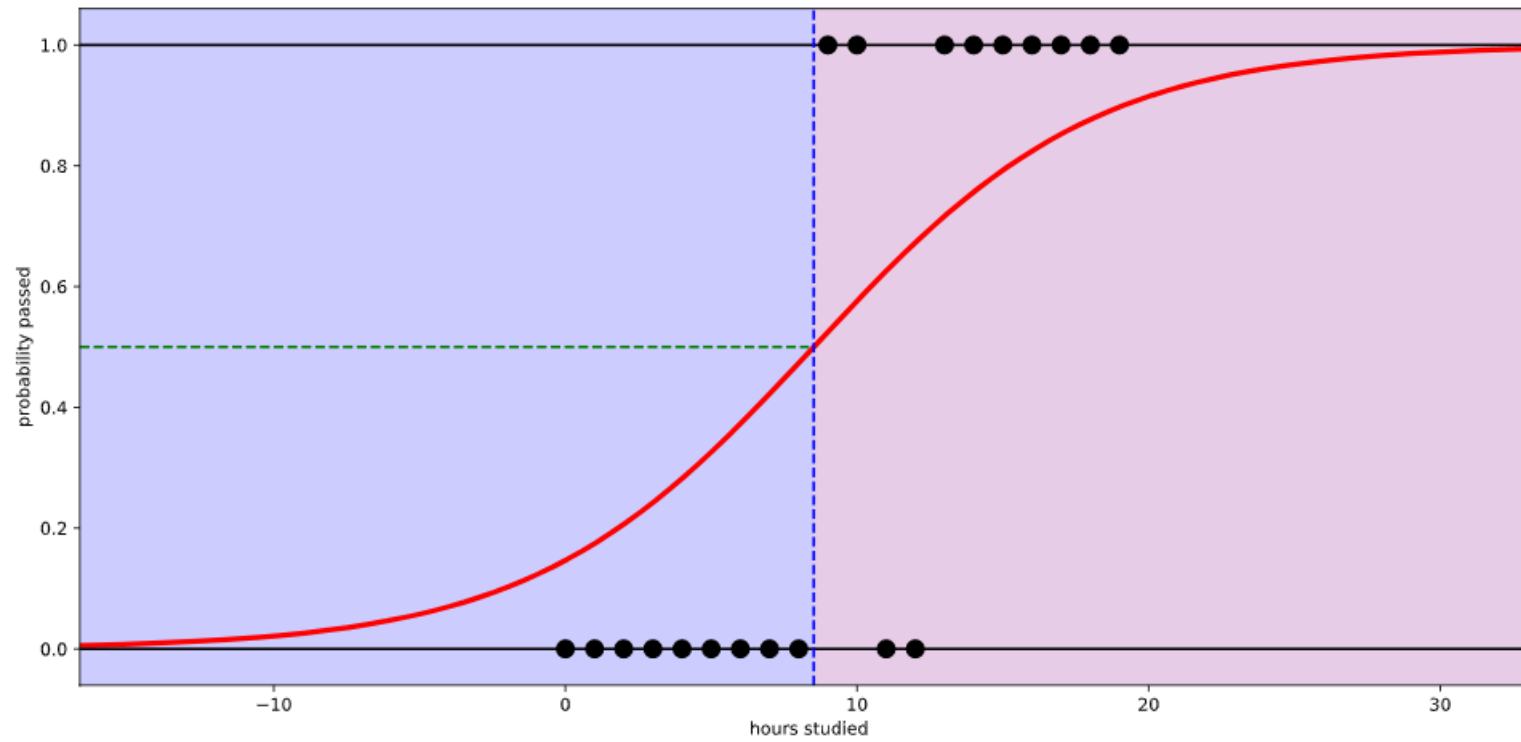


로지스틱 회귀

회귀를 사용하여 데이터가 어떤 범주에 속할 확률을
0에서 1 사이의 값으로 예측하고 그 확률에 따라 가능성이
더 높은 범주에 속하는 것으로 분류해주는 지도 학습 알고리즘

$$\text{logistic function} = \frac{e^{\beta \cdot X_i}}{1 + e^{\beta \cdot X_i}}$$

로지스틱 회귀 예시



의사결정나무

나무구조로 도표화하여 분류와 예측을 수행하는 분석 방법

*예측이 나무구조에 의한 추론규칙에 의해서 표현

- 나무구조로 표현되어 모델을 사용자가 쉽게 이해
- 두 개 이상의 변수가 결합하여 목표변수에 어떻게 영향을 주는지 해석
- 선형성, 정규성, 등분산성 등의 가정이 불필요

의사결정나무 예시

