

Lecture 09 보충 - 조건부 확률

조건부 확률(conditional probability)이란?

- 특정 사건이 일어났다는 조건(condition) 하에서 계산하는 확률
- 사건 B가 일어났다는 조건 하에서 사건 A가 일어날 확률: $P(A|B)$
- 일반적으로, $P(A) \neq P(A|B)$

독감 발병 확률 계산하기

가정

- 독감에 걸리는 사건: flu
- 백신 접종을 하는 사건: vaccine
- $P(\text{flu})$: 독감에 걸릴 확률(= 독감에 걸린 비율)
- $P(\text{vaccine})$: 예방 접종을 할 확률(= 예방 접종을 한 비율)

데이터

조건	vaccine	no vaccine	-
flu	30	2,000	2,030
no flu	4,000	30,000	34,000
-	4,030	32,000	36,030

- 예방 접종하고 독감에 걸린 사람 수:
 $n(\text{flu} \& \text{vaccine}) = 30$
- 예방 접종하지 않고 독감에 걸린 사람 수:
 $n(\text{flu} \& \text{no vaccine}) = 2,000$
- 예방 접종하고 독감에 걸리지 않은 사람 수:
 $n(\text{no flu} \& \text{vaccine}) = 4,000$
- 예방 접종하지 않고 독감에 걸리지 않은 사람 수:
 $n(\text{no flu} \& \text{no vaccine}) = 30,000$

질문

독감에 걸릴 확률은 얼마인가?

- $P(\text{flu}) = n(\text{flu}) / n(\text{total}) = 2,030 / 36,030 = 0.0563$

예방 접종을 했을 때, 독감에 걸릴 확률은 얼마인가?

- $P(\text{flu} | \text{vaccine}) = n(\text{flu} \& \text{vaccine}) / n(\text{vaccine}) = 30 / 4,030 = 0.0074$

예방 접종을 하지 않았을 때, 독감에 걸릴 확률은 얼마인가?

- $P(\text{flu} | \text{no vaccine}) = n(\text{flu} \& \text{no vaccine}) / n(\text{no vaccine}) = 2,000 / 32,000 = 0.0625$

결론

- 예방 접종과 상관 없이, 독감에 걸릴 확률은 5.63%
- 만약 예방 접종을 맞았다면 독감에 걸릴 확률은 0.74%
- 예방 접종을 맞지 않았다면 독감에 걸릴 확률은 6.25%
- 예방 접종 유무에 따라 독감에 걸릴 확률은 8.4배(= 0.0625 / 0.0074) 차이가 난다.

다른 질문

P(vaccine | flu) = ?

- 조건부 확률에서 조건 사건과 관심 사건이 바뀔 수 있다.
- $P(\text{vaccine} | \text{flu}) = n(\text{vaccine} \& \text{flu}) / n(\text{flu}) = 30 / 2,030 = 0.0148$
- 독감에 걸린 사람이 예방 접종을 맞았을 확률: 1.48%
- 예방 접종을 맞았어도 독감에 걸릴 확률은 1.48%

독감에 걸린 사람 중, 백신을 맞지 않은 비율은?

- $P(\text{no vaccine} | \text{flu}) = n(\text{no vaccine} \& \text{flu}) / n(\text{flu}) = 2,000 / 2,030 = 0.9852$

유의할 점

조건부 확률에서 사건의 선후 관계

- 조건부 확률에서 조건부 사건이 항상 시간 상 먼저 일어나야 하는 것은 아니다.
- $P(\text{rain} | \text{wet})$: 오늘 땅이 젖어 있을 때, 어제 비가 왔을 확률
- $P(\text{wet} | \text{rain})$: 어제 비가 왔을 때, 오늘 땅이 젖어 있을 확률

일상에서의 조건부 확률

- 조건부 확률이라고 명시하지 않더라도, 사람들은 자연스럽게 조건부로 데이터를 살펴 본다.
- 특정 그룹으로 데이터를 나누어서 비교하는 행위가 그룹이라는 조건으로 데이터를 관찰하는 과정.
예시) iris data

