

베이즈 통계학이란...

- ① 1763년 영국인 Bayes가 제시한 통계적 방법
- ② 표본에서 얻은 정보뿐 아니라 사전지식이나 정보를 활용해 확률을 결정하는 통계학
- ③ 개인적 경험이 확률에 영향을 줌
- ④ 실행 횟수(경험)이 많아질수록 기존의 통계와 비슷해짐
- ⑤ 생활에서 무의식적으로 사용하는 예
 - 복권의 당첨이 많이 되는 복권방(경험)에서 복권을 구매
 - 의사의 경험에 의해 질병을 판단

베이지안 정리

- ① 조건부 확률 : 특정 조건에서 사건이 일어날 확률
 - $P(A|B)$: B(조건)가 나타날 때, A가 나타날 확률
 - $P(A|B) = \frac{\text{특정 사건의 경우의 수}}{\text{전체 사건의 경우의 수}} = \frac{n(A \cap B)}{n(B)} = \frac{P(A \cap B)}{P(B)}$
 - ex) 질병 검사에서 양성인 경우 실제 병에 걸릴 확률
A : 병에 걸리는 것, B : 질병 검사에서 양성인 경우
$$P(A|B) = \frac{\text{양성이 나오면서 병에 걸리는 경우}}{\text{질병 검사에서 양성인 경우}} = \frac{P(A \cap B)}{P(B)}$$

② 베이지안 정리

- 조건부 확률에서 $P(A \cap B) = P(A|B)P(B)$ 이며,
 $P(A \cap B) = P(B \cap A)$ 이므로, $P(A|B)P(B) = P(B|A)P(A)$
 $\therefore P(A|B) = \frac{P(B|A)P(A)}{P(B)}$: 베이지안 정리

이때, $P(B|A)$ 가 사전 지식(경험)에 해당

예를 들어, 재채기 증상이 있는 사람이 감기에 걸릴 확률을 구한다면,

A는 재채기하는 것, B는 감기에 걸리는 것이라 놓고 구하면,

구하는 것은 $P(A|B)$ 이며, 알아야하는 것에서

$P(A)$ [재채기하는 확률], $P(B)$ [감기에 걸릴 확률]는 구할 수 있지만,

$P(B|A)$ [감기에 걸렸을 때 재채기를 할 확률]은 기존의 경험에 의존해야 함.

이때, $P(B|A)$ 의 경험이 많을수록 정확도는 올라가게 됨

- ③ 나온 결과($P(A|B)$)를 기존 경험($P(B|A)$)으로 처리함으로써 정확도를 올릴 수 있음
 \Rightarrow 인공지능, 빅데이터 등의 수학적 모델의 기초가 됨