

## ⚙ 베이즈 이론식

- $P(A) * P(B|A) = P(B) * P(A|B)$

## ⚙ 나이브 베이즈 알고리즘

- $P(\text{특정단어})$ 가 나타날 확률은 일일이 구할 수 없으므로 모두 같다고 생각함
- $X\text{문서} = \{\text{'단어1'}, \text{'단어2'}, \text{'단어3'}, \dots\}$  일 때에  $P(X|\text{스팸})$ 인 확률은 서로 독립이라고 가정하고,  
단순하게  $P(\text{'단어1'}|\text{스팸}) * P(\text{'단어2'}|\text{스팸}) * P(\text{'단어3'}|\text{스팸}) * \dots$  으로 계산함

## ⚙ 나이브 베이즈의 적용

- 콜센터의 상담전화가 걸려 올 때에 쓰는 단어들 목록을 보고 유추하여 해당하는 전문상담원을 연결시켜 주는 모델임
- 결혼을 준비 중인 30대 커플들의 선호하는 신혼가구들의 분포를 예측하는 모델임

## ⚙ 감성분석

- 영화 댓글을 판별하여 긍정/부정을 예측하는 모델을 말함
- 텍스트 마이닝의 영역에서 '감성분석'은 나아가 NLP(자연어 처리)까지 발전될 수도 있음