베이즈 정리

Outline

I. 베이즈 정리 개념

II. 조건부 독립

I. 베이즈 정리

$$P(a \land b) = P\,(a \mid b)P(b) = P\,(b \mid a)P(a)$$ (곱셈정리)

$$\Downarrow$$

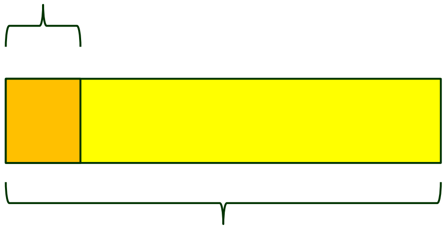
$$P\,(a \mid b)P(b) = P\,(b \mid a)P(a)$$

$$\Downarrow$$

$$\boxed{P\,(b \mid a) = \frac{P\,(a \mid b)P(b)}{P(a)}}$$ (베이즈 정리)

a가 발생한 상황에서
b가 발생할 확률

사건 b에 의해 발생된
부분: $P(a \mid b)P(b)$



사건 a의 발생 확률: $P(a)$

아래의 정보가 알려져 있다면, 사건 a가 발생한 상황에서 b가 얼마나 자주 발생하는지를 알 수있음

- b가 발생한 상황에서 a가 얼마나 자주 발생하는지
- a 자체가 독자적으로 얼마나 자주 일어나는지
- b 자체가 독자적으로 얼마나 자주 발생하는지

다중 값 변수를 위한 베이즈 정리

$$\mathbf{P}(Y \mid X) = \frac{\mathbf{P}(X \mid Y)\mathbf{P}(Y)}{\mathbf{P}(X)}$$

X와 Y의 가능한 값 쌍에 대해 각각 적용되는 방정식의 집합

특정 사건 $\boldsymbol{e}$에 대해 조건화된, 좀더 일반화된 버전:

$$\mathbf{P}(Y \mid X, \boldsymbol{e}) = \frac{\mathbf{P}(X \mid Y, \boldsymbol{e})\mathbf{P}(Y \mid \boldsymbol{e})}{\mathbf{P}(X \mid \boldsymbol{e})} = \frac{\mathbf{P}(X, Y \mid \boldsymbol{e})}{\mathbf{P}(X \mid \boldsymbol{e})}$$

베이즈 정리 적용하기

- 어떤 알려지지 않은 *원인*에 의해 발생한 *결과*를 증거(사건)로 인식
- 원인 규명

진단 측면의 관계성　　　인과관계 측면의 관계성

$$P(cause \mid effect) = \frac{P(effect \mid cause)P(cause)}{P(effect)}$$

확률적 정보는 종종 P(결과 | 원인)의 형태로 제공되지만, P(원인 | 결과)의 형태로는 제공되지 않는 경우가 많음

의사는 P(증상 | 질병)에 대한 정보를 알고 있으며, 이를 기반으로 P(질병 | 증상)을 추론하여 진단을 내리고자 함

예제

어떤 의사가 다음의 사실을 미리 알고 있다고 가정

- 뇌수막염은 환자에게 70%의 확률로 목 경직 증상을 유발
- 임의의 환자가 뇌수막염에 걸릴 사전 확률은 1/50,000
- 임의의 환자가 목이 뻣뻣해질 사전 확률은 1/100

s: 환자가 목 경직 증상을 보임
m: 환자가 뇌수막염에 걸림

$$P(s \mid m) = 0.7$$

$$P(m) = 1/50000$$

$$P(s) = 0.01$$

$$P(m \mid s) = \frac{P(s \mid m)P(m)}{P(s)} = \frac{0.7 \cdot 1/50000}{0.01} = 0.0014$$

뇌수막염이 목 경직 증상을 유발하는 비율을 이용하여, 목이 뻣뻣한 경우 뇌수막염에 걸릴 확률을 알아냄

사건(증거)의 사전 확률

$$P(m \mid s) = \frac{P(s \mid m)P(m)}{P(s)}$$?

$P(s \mid m)$과 $P(s \mid \neg m)$를 알면 $P(s)$를 계산할 수 있음:

$$P(m, s) = P(m \mid s)\, P(s) = P(s \mid m)\, P(m) \quad P(\neg m \mid s)\, P(s) = P(s \mid \neg m)\, P(\neg m)$$

$$P(m \mid s)\, P(s) + P(\neg m \mid s)\, P(s) = P(s \mid m)\, P(m) + P(s \mid \neg m)\, P(\neg m)$$

$$P(m \mid s) + P(\neg m \mid s) = 1$$

$$P(s) = P(s \mid m)\, P(m) + P(s \mid \neg m)\, P(\neg m)$$

베이즈 정리 (일반형)

$P(s)$를 모를 때 베이즈 규칙에 정규화를 적용하기:

$$\mathbf{P}(M \mid s) = \alpha \langle P(s \mid m)P(m), P(s \mid \neg m)P(\neg m) \rangle$$

두 가지 값 m과 $\neg m$을 갖는 확률 변수

$$\alpha = \frac{1}{P(s)} = \frac{1}{P(s \mid m)\, P(m) + P(s \mid \neg m)\, P(\neg m)}$$

- $P(\neg m) = 1 - P(m)$
- $P(s)$ 대신 $P(s \mid \neg m)$을 추정

  더 쉬울 수도 있고 더 어려운 경우도 있음

$$\mathbf{P}(Y \mid X) = \alpha\, \mathbf{P}(X \mid Y)\mathbf{P}(Y)$$

$\mathbf{P}(Y \mid X)$의 값들이 1이 되도록 합을 맞추기 위한 정규화 상수

II. 결합 사건(증거 조합하기)

두 가지 이상의 증거(사건)를 사용하는 베이즈 정리

- 전체 결합 확률 분포를 알고 있다고 가정

	$toothache$		$\neg toothache$	
	$catch$	$\neg catch$	$catch$	$\neg catch$
$cavity$	0.108	0.012	0.072	0.008
$\neg cavity$	0.016	0.064	0.144	0.576

$$\mathbf{P}(Cavity \mid toothache \wedge catch) = \alpha\langle 0.108, 0.016\rangle \approx \langle 0.871, 0.129\rangle$$

$$\frac{1}{\alpha} = P(toothache \wedge catch) = 0.108 + 0.016 = 0.124$$

♠ 다수의 증거 변수를 다루기에는 확장성이 부족함

지수적 증가(Exponential Growth)

- 베이즈 정리(규칙)의 적용:

$$\mathbf{P}(Cavity \mid toothache \wedge catch) = \alpha \mathbf{P}(toothache \wedge catch \mid Cavity)\, \mathbf{P}(Cavity)$$

즉,

$$\langle P(cavity \mid toothache \wedge catch), P(\neg cavity \mid toothache \wedge catch) \rangle$$

$$= \alpha \langle P(toothache \wedge catch \mid cavity), P(toothache \wedge catch \mid \neg cavity) \rangle$$

$$\cdot \langle P(cavity), P(\neg cavity) \rangle$$ // 요소별 곱셈

- n 개의 부울리언 증거 변수: *Toothache*, *Catch*, *X-rays*, *Diet*, …

$\Longrightarrow$ 관측값에 대한 2^n 개의 가능한 조합이 필요함!

$$\mathbf{P}(toothache \wedge catch \mid Cavity \wedge Toothache \wedge \cdots)$$

조건부 독립

- *충치의 존재 여부*와는 무관하게 **<u>치통</u>**과 **<u>Catch</u>**는 상호 독립
  - 각각은 충치에 의해 직접적으로 발생
  - 상호간에 직접적인 영향은 없음

    치통은 치아의 신경 상태에 따라 달라짐

    탐침(Probe)의 정확도는 치통과 무관하게 치과 의사의 기술에 따라 달라짐

- 충치(Cavity)를 고려한 치통과 Catch의 조건부 독립:

$$\mathbf{P}(toothache \wedge catch \mid Cavity) = \mathbf{P}(toothache \mid Cavity)\, \mathbf{P}(catch \mid Cavity)$$

베이즈정리: $\mathbf{P}(Cavity \mid toothache \wedge catch) = \alpha \boxed{\mathbf{P}(toothache \wedge catch \mid Cavity)}\, \mathbf{P}(Cavity)$

$$\mathbf{P}(Cavity \mid toothache \wedge catch) \\ = \alpha \mathbf{P}(toothache \mid Cavity)\, \mathbf{P}(catch \mid Cavity) \mathbf{P}(Cavity)$$

조건부 독립의 정의

세 번째 변수 Z가 주어졌을 때 두 변수 X와 Y의 조건부 독립:

$$\underbrace{P(X,Y}_{X \wedge Y} \mid Z) = P(X \mid Z)\, P(Y \mid Z)$$

$$\mathbf{P}(\mathit{Toothache}, \mathit{Catch} \mid \mathit{Cavity}) = \mathbf{P}(\mathit{Toothache} \mid \mathit{Cavity})\, \mathbf{P}(\mathit{Catch} \mid \mathit{Cavity})$$

// 치통과 Catch의 모든 가능한 값 조합에 대해 독립임을 나타내는 명제

동치 관계:

$$P(X \mid Y,Z) = P(X \mid Z)$$ and $$P(Y \mid X,Z) = P(Y \mid Z)$$

더 작은 조건부 명제로 분해도 가능

결합 확률 분포의 분해

$$\boldsymbol{P}(Toothache, Catch, Cavity)$$

$$= \boldsymbol{P}(Toothache, Catch \mid Cavity)\ \boldsymbol{P}(Cavity)$$

$$= \underbrace{\boldsymbol{P}(Toothache \mid Cavity)\ \boldsymbol{P}(Catch \mid Cavity)\ \boldsymbol{P}(Cavity)}$$

2×2, 2×2, 그리고 2×1 크기의 3개의 테이블에는 각각 첫 번째 행에 있는 2+2+1=5개의 독립적인 숫자가 존재

n개의 징조가 있을 때, 전체 크기는 $O(2^n)$가 아닌 $O(n)$으로 증가

- 조건부 독립 명제로 인해 확률 시스템의 규모 확장이 가능
- 조건부 독립이 절대적 독립보다 더 자주 사용되고, 더 용이함
- 조건부 독립을 활용한 대규모 확률적 도메인의 분해 기술은 인공지능(AI) 분야에서 중요한 기술로 인정됨