Mathematics for Artificial Intelligence

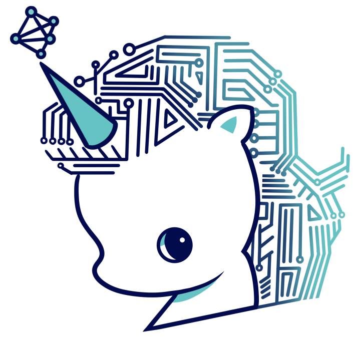
8강: 베이즈 통계학 맛보기

임성빈 UNIST 인공지능대학원 & 산업공학과
Learning Intelligent Machine Lab

조건부 확률이란?

- 베이즈 통계학을 이해하기 위해선 조건부확률의 개념을 이해해야 합니다

$$P(A \cap B) = P(B)P(A|B)$$



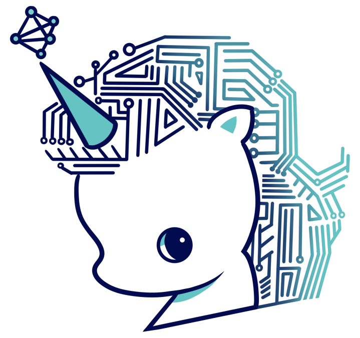
조건부확률 $P(A \mid B)$ 는 사건 B 가 일어난 상황에서 사건 A 가 발생할 확률을 의미한다

조건부 확률이란?

- 베이즈 통계학을 이해하기 위해선 조건부확률의 개념을 이해해야 합니다
- 베이즈 정리는 조건부확률을 이용하여 정보를 갱신하는 방법을 알려줍니다

$$P(A \cap B) = P(B)P(A|B)$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = P(B)\frac{P(A|B)}{P(A)}$$



A 라는 새로운 정보가 주어졌을 때 $P(B)$ 로부터 $P(B|A)$ 를 계산하는 방법을 제공한다

베이즈 정리: 예제

$$P(\theta|\mathcal{D}) = P(\theta)\frac{P(\mathcal{D}|\theta)}{P(\mathcal{D})}$$

가능도 (likelihood)

Evidence

사후확률 (posterior)

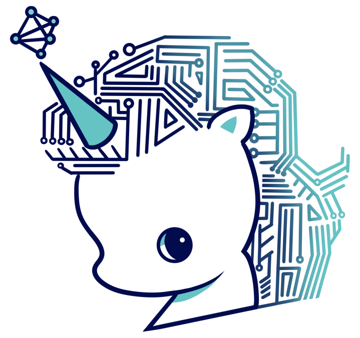
사전확률 (prior)

- COVID-99 의 발병률이 10% 로 알려져있다. COVID-99 에 실제로 걸렸을 때 검진될 확률은 99%, 실제로 걸리지 않았을 때 오검진될 확률이 1% 라고 할 때,

베이즈 정리: 예제

$$\underbrace{P(\theta|\mathcal{D})}_{\text{사후확률 (posterior)}} = \underbrace{P(\theta)}_{\text{사전확률 (prior)}} \frac{\overbrace{P(\mathcal{D}|\theta)}^{\text{가능도 (likelihood)}}}{\underbrace{P(\mathcal{D})}_{\text{Evidence}}}$$

- COVID-99 의 발병률이 10% 로 알려져있다. COVID-99 에 실제로 걸렸을 때 검진될 확률은 99%, 실제로 걸리지 않았을 때 오검진될 확률이 1% 라고 할 때, 어떤 사람이 질병에 걸렸다고 검진결과가 나왔을 때 정말로 COVID-99 에 감염되었을 확률은?



사전확률, 민감도(Recall), 오탐율(False alarm)을 가지고 정밀도(Precision)를 계산하는 문제이다

베이즈 정리: 예제

$$P(\theta|\mathcal{D}) = P(\theta)\frac{P(\mathcal{D}|\theta)}{P(\mathcal{D})}$$

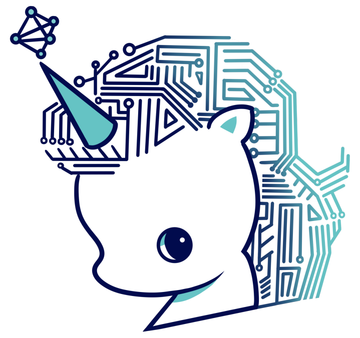
$$P(\theta) = 0.1$$

$$P(\mathcal{D}|\theta) = 0.99$$
$$P(\mathcal{D}|\neg\theta) = 0.01$$

사전확률

가능도

- COVID-99 의 발병률이 10% 로 알려져있다. COVID-99 에 실제로 걸렸을 때 검진될 확률은 99%, 실제로 걸리지 않았을 때 오검진될 확률이 1% 라고 할 때, 어떤 사람이 질병에 걸렸다고 검진결과가 나왔을 때 정말로 COVID-99 에 감염되었을 확률은?



θ 를 COVID-99 발병 사건으로 정의(관찰 불가)하고,
$\mathcal{D}$ 를 테스트 결과라고 정의(관찰 가능)한다

베이즈 정리: 예제

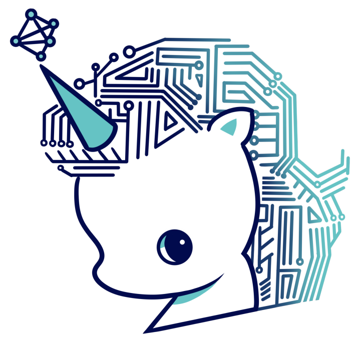
$$P(\theta|\mathcal{D}) = P(\theta)\frac{P(\mathcal{D}|\theta)}{P(\mathcal{D})}$$

$$P(\theta) = 0.1$$

$$P(\mathcal{D}|\theta) = 0.99$$

$$P(\mathcal{D}|\neg\theta) = 0.01$$

$$P(\mathcal{D}) = \sum_{\theta} P(\mathcal{D}|\theta)P(\theta) = 0.99 \times 0.1 + 0.01 \times \overset{P(\neg\theta)}{0.9} = 0.108$$



만일 $P(\mathcal{D}\,|\,\neg\theta)$ 를 모른다면 이 문제는 풀기 어렵다

- COVID-99 의 발병률이 10% 로 알려져있다. COVID-99 에 실제로 걸렸을 때 검진될 확률은 99%, 실제로 걸리지 않았을 때 오검진될 확률이 1% 라고 할 때, 어떤 사람이 질병에 걸렸다고 검진결과가 나왔을 때 정말로 COVID-99 에 감염되었을 확률은?

베이즈 정리: 예제

$$P(\theta|\mathcal{D}) = P(\theta)\frac{P(\mathcal{D}|\theta)}{P(\mathcal{D})}$$

$$P(\theta) = 0.1$$

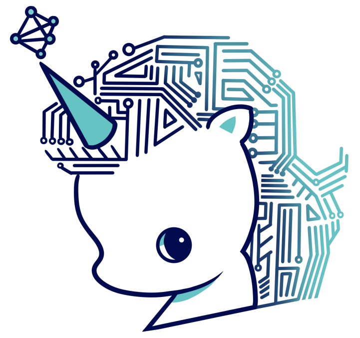
$$P(\mathcal{D}|\theta) = 0.99$$

$$P(\mathcal{D}|\neg\theta) = 0.01$$

$$P(\mathcal{D}) = \sum_{\theta} P(\mathcal{D}|\theta)P(\theta) = 0.99 \times 0.1 + 0.01 \times 0.9 = 0.108$$

$$P(\theta|\mathcal{D}) = 0.1 \times \frac{0.99}{0.108} \approx 0.916$$

- COVID-99 의 발병률이 10% 로 알려져있다. COVID-99 에 실제로 걸렸을 때 검진될 확률은 99%, 실제로 걸리지 않았을 때 오검진될 확률이 1% 라고 할 때, 어떤 사람이 질병에 걸렸다고 검진결과가 나왔을 때 정말로 COVID-99 에 감염되었을 확률은?



만일 오검진될 확률(1종 오류)이 1% 가 아닌 10% 면 어떻게 될까?

베이즈 정리: 예제

$$P(\theta|\mathcal{D}) = P(\theta)\frac{P(\mathcal{D}|\theta)}{P(\mathcal{D})}$$

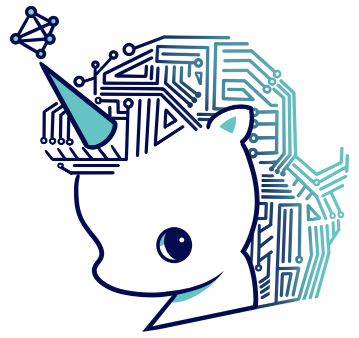
$$P(\theta) = 0.1$$

$$P(\mathcal{D}|\theta) = 0.99$$
$$P(\mathcal{D}|\neg\theta) = 0.1$$

$$P(\mathcal{D}) = \sum_{\theta} P(\mathcal{D}|\theta)P(\theta) = 0.99 \times 0.1 + 0.1 \times 0.9 = 0.189$$

$$P(\theta|\mathcal{D}) = 0.1 \times \frac{0.99}{0.189} \approx 0.524$$

- COVID-99 의 발병률이 10% 로 알려져있다. COVID-99 에 실제로 걸렸을 때 검진될 확률은 99%, 실제로 걸리지 않았을 때 오검진될 확률이 1% 라고 할 때, 어떤 사람이 질병에 걸렸다고 검진결과가 나왔을 때 정말로 COVID-99 에 감염되었을 확률은?



오탐율(False alarm)이 오르면 테스트의 정밀도(Precision)가 떨어진다

조건부 확률의 시각화

$P(\theta) = 0.1$
사전확률

$P(\neg\theta) = 0.9$
사전확률

$P(\mathscr{D}\,|\,\neg\theta)P(\neg\theta) = 0.09$
False Positive
(1종 오류)

$P(\mathscr{D}\,|\,\neg\theta) = 0.1$
오탐(False alarm)

$P(\mathscr{D}\,|\,\theta) = 0.99$
민감도 (Recall)

$P(\mathscr{D}\,|\,\theta)P(\theta) = 0.099$
True Positive

True Negative

$P(\neg\mathscr{D}\,|\,\neg\theta) = 0.9$
특이도(specificity)

False Negative
(2종 오류)

조건부 확률의 시각화

$P(\theta|\mathscr{D})$
정밀도 (Precision)

$$= \frac{\text{TP}}{\text{TP} + \text{FP}} = 0.524$$

$P(\theta) = 0.1$
사전확률

$P(\neg\theta) = 0.9$
사전확률

$P(\mathscr{D}|\neg\theta)P(\neg\theta) = 0.09$
False Positive
(1종 오류)

$P(\mathscr{D}|\neg\theta) = 0.1$
오탐(False alarm)

$P(\mathscr{D}|\theta) = 0.99$
민감도 (Recall)

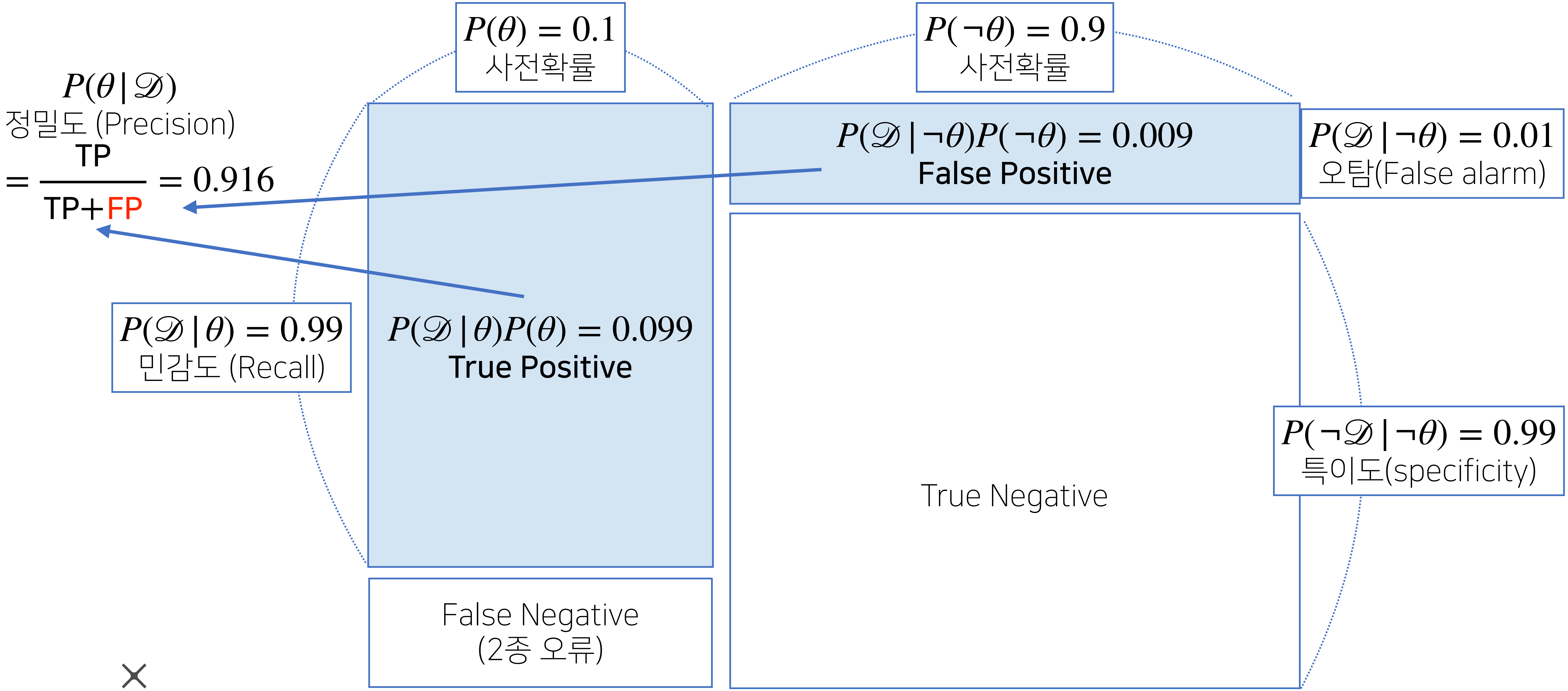
$P(\mathscr{D}|\theta)P(\theta) = 0.099$
True Positive

True Negative

$P(\neg\mathscr{D}|\neg\theta) = 0.9$
특이도(specificity)

False Negative
(2종 오류)

조건부 확률의 시각화


$P(\theta) = 0.1$
사전확률
$P(\neg\theta) = 0.9$
사전확률
$P(\theta \mid \mathscr{D})$
정밀도 (Precision)
$= \frac{\text{TP}}{\text{TP}+\text{FP}} = 0.916$
$P(\mathscr{D} \mid \neg\theta)P(\neg\theta) = 0.009$
False Positive
$P(\mathscr{D} \mid \neg\theta) = 0.01$
오탐(False alarm)
$P(\mathscr{D} \mid \theta) = 0.99$
민감도 (Recall)
$P(\mathscr{D} \mid \theta)P(\theta) = 0.099$
True Positive
$P(\neg\mathscr{D} \mid \neg\theta) = 0.99$
특이도(specificity)
True Negative
False Negative
(2종 오류)

베이즈 정리를 통한 정보의 갱신

- 베이즈 정리를 통해 새로운 데이터가 들어왔을 때 앞서 계산한 사후확률을 사전확률로 사용하여 갱신된 사후확률을 계산할 수 있습니다

$$P(\theta|\mathcal{D}) = P(\theta)\frac{P(\mathcal{D}|\theta)}{P(\mathcal{D})}$$

사후확률
(posterior)

$$P(\theta|\mathcal{D}) = P(\theta)\frac{P(\mathcal{D}|\theta)}{P(\mathcal{D})}$$

갱신된 사후확률
(posterior)

베이즈 정리를 통한 정보의 갱신

- 베이즈 정리를 통해 새로운 데이터가 들어왔을 때 앞서 계산한 사후확률을 사전확률로 사용하여 갱신된 사후확률을 계산할 수 있습니다

$$P(\theta|\mathcal{D}) = P(\theta)\frac{P(\mathcal{D}|\theta)}{P(\mathcal{D})}$$

사후확률
(posterior)

$$P(\theta|\mathcal{D}) = P(\theta)\frac{P(\mathcal{D}|\theta)}{P(\mathcal{D})}$$

갱신된 사후확률
(posterior)

$$P(\theta|\mathcal{D}) = 0.1 \times \frac{0.99}{0.189} \approx 0.524$$

- 앞서 COVID-99 판정을 받은 사람이 두 번째 검진을 받았을 때도 양성이 나왔을 때 진짜 COVID-99 에 걸렸을 확률은?

베이즈 정리를 통한 정보의 갱신

- 베이즈 정리를 통해 새로운 데이터가 들어왔을 때 앞서 계산한 사후확률을 사전확률로 사용하여 갱신된 사후확률을 계산할 수 있습니다

$$P(\theta|\mathcal{D}) = 0.1 \times \frac{0.99}{0.189} \approx 0.524$$

$$P(\mathcal{D}|\theta) = 0.99$$

$$P(\mathcal{D}|\neg\theta) = 0.1$$

- 앞서 COVID-99 판정을 받은 사람이 두 번째 검진을 받았을 때도 양성이 나왔을 때 진짜 COVID-99 에 걸렸을 확률은?

베이즈 정리를 통한 정보의 갱신

- 베이즈 정리를 통해 새로운 데이터가 들어왔을 때 앞서 계산한 사후확률을 사전확률로 사용하여 갱신된 사후확률을 계산할 수 있습니다

$$P(\theta|\mathcal{D}) = 0.1 \times \frac{0.99}{0.189} \approx 0.524$$

$$P(\mathcal{D}|\theta) = 0.99$$

$$P(\mathcal{D}|\neg\theta) = 0.1$$

$$P(\mathcal{D}^*) = 0.99 \times 0.524 + 0.1 \times 0.476 \approx 0.566$$

갱신된 evidence

- 앞서 COVID-99 판정을 받은 사람이 두 번째 검진을 받았을 때도 양성이 나왔을 때 진짜 COVID-99 에 걸렸을 확률은?

베이즈 정리를 통한 정보의 갱신

- 베이즈 정리를 통해 새로운 데이터가 들어왔을 때 앞서 계산한 사후확률을 사전확률로 사용하여 갱신된 사후확률을 계산할 수 있습니다

$$P(\theta|\mathcal{D}) = 0.1 \times \frac{0.99}{0.189} \approx 0.524$$

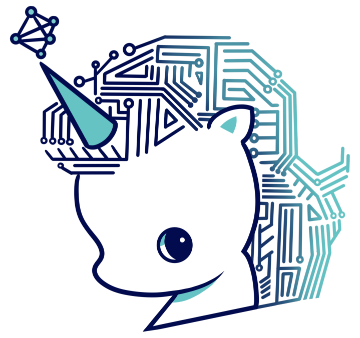
$$P(\mathcal{D}|\theta) = 0.99$$

$$P(\mathcal{D}|\neg\theta) = 0.1$$

$$P(\mathcal{D}^*) = 0.99 \times 0.524 + 0.1 \times 0.476 \approx 0.566$$

갱신된 사후확률 (posterior) $$P(\theta|\mathcal{D}^*) = 0.524 \times \frac{0.99}{0.566} \approx 0.917$$

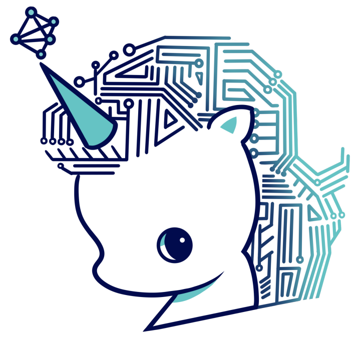
- 앞서 COVID-99 판정을 받은 사람이 두 번째 검진을 받았을 때도 양성이 나왔을 때 진짜 COVID-99 에 걸렸을 확률은?



세번째 검사해도 양성이 나오면 정밀도가 99.1% 까지 갱신된다

조건부 확률 → 인과관계?

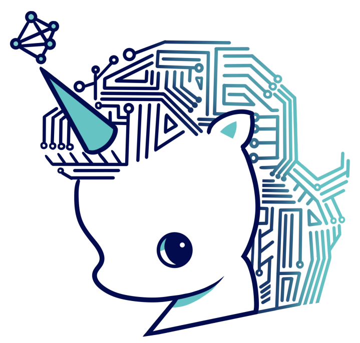
- 조건부 확률은 유용한 통계적 해석을 제공하지만 인과관계(causality)를 추론할 때 함부로 사용해서는 안 됩니다



데이터가 많아져도 조건부 확률만 가지고
인과관계를 추론하는 것은 불가능합니다

조건부 확률 → 인과관계?

- 조건부 확률은 유용한 통계적 해석을 제공하지만 인과관계(causality)를 추론할 때 함부로 사용해서는 안 됩니다
- 인과관계는 데이터 분포의 변화에 강건한 예측모형을 만들 때 필요합니다



단, 인과관계만으로는 높은 예측 정확도를 담보하기는 어렵습니다

99% 예측정확도 테스트시
→ **95%** 예측정확도 시나리오 A
→ 72% 예측정확도 시나리오 B

조건부확률 기반 예측모형

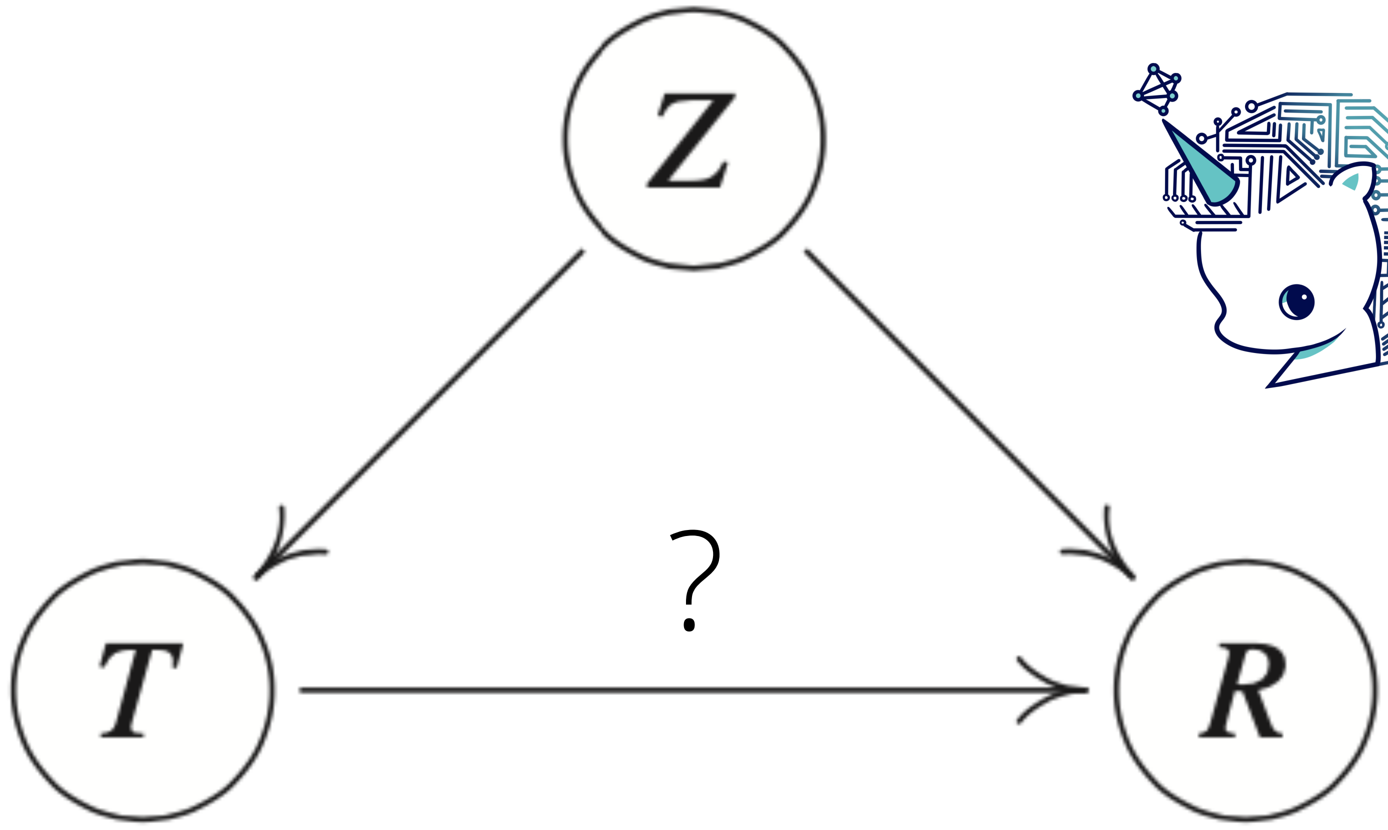
vs

85% 예측정확도 테스트시
→ **82%** 예측정확도 시나리오 A
→ 81% 예측정확도 시나리오 B

인과관계 기반 예측모형

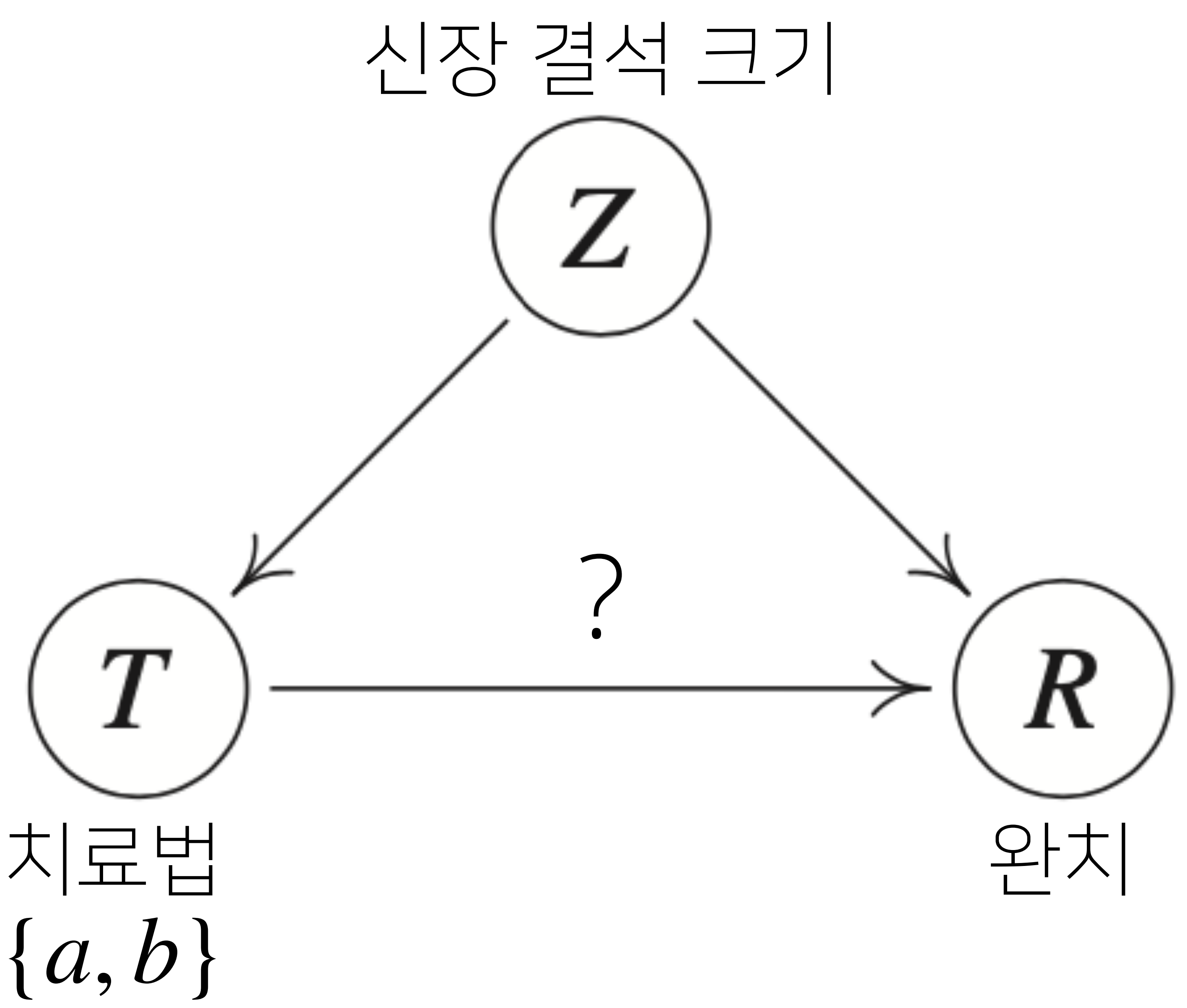
조건부 확률 → 인과관계?

- 조건부 확률은 유용한 통계적 해석을 제공하지만 인과관계(causality)를 추론할 때 함부로 사용해서는 안 됩니다
- 인과관계는 데이터 분포의 변화에 강건한 예측모형을 만들 때 필요합니다
- 인과관계를 알아내기 위해서는 중첩요인(confounding factor)의 효과를 제거하고 원인에 해당하는 변수만의 인과관계를 계산해야 합니다



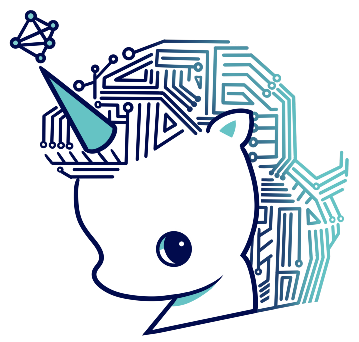
만일 Z 의 효과를 제거하지 않으면 가짜 연관성(spurious correlation)이 나온다

인과관계 추론: 예제



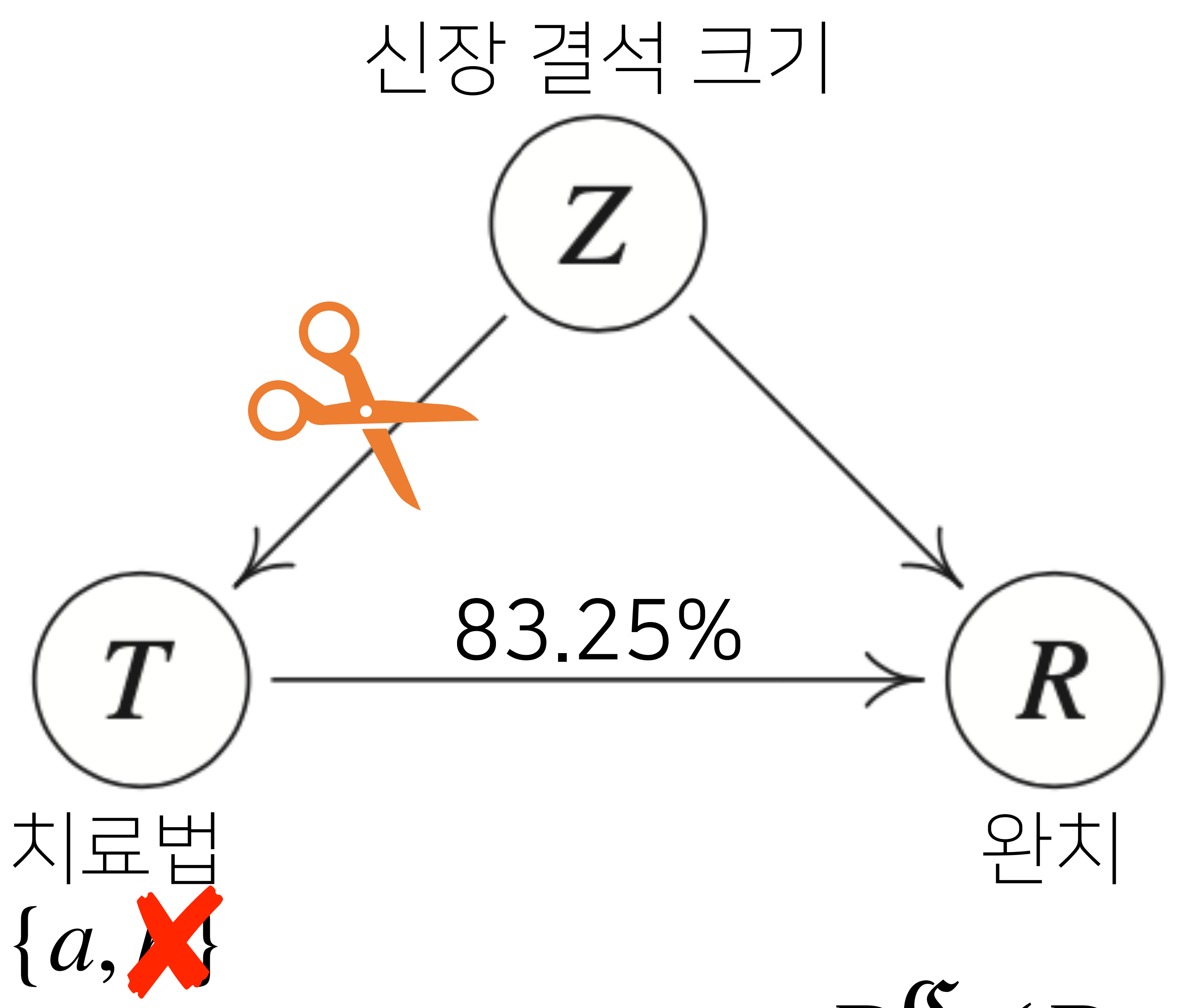
	Overall	Patients with small stones ($Z = 0$)	Patients with large stones ($Z = 1$)
Treatment a: Open surgery	78% (273/350)	**93%** (81/87)	**73%** (192/263)
Treatment b: Percutaneous nephrolithotomy	**83%** (289/350)	87% (234/270)	69% (55/80)

출처: Elements of Causal Inference, Peters et al.



치료법 a 와 b 중 어느 것이 더 나은가?

인과관계 추론: 예제



		$Z=0$	$Z=1$
	Overall	Patients with small stones	Patients with large stones
Treatment a: Open surgery	78% (273/350)	**93%** (81/87)	**73%** (192/263)
Treatment b: Percutaneous nephrolithotomy	**83%** (289/350)	87% (234/270)	69% (55/80)

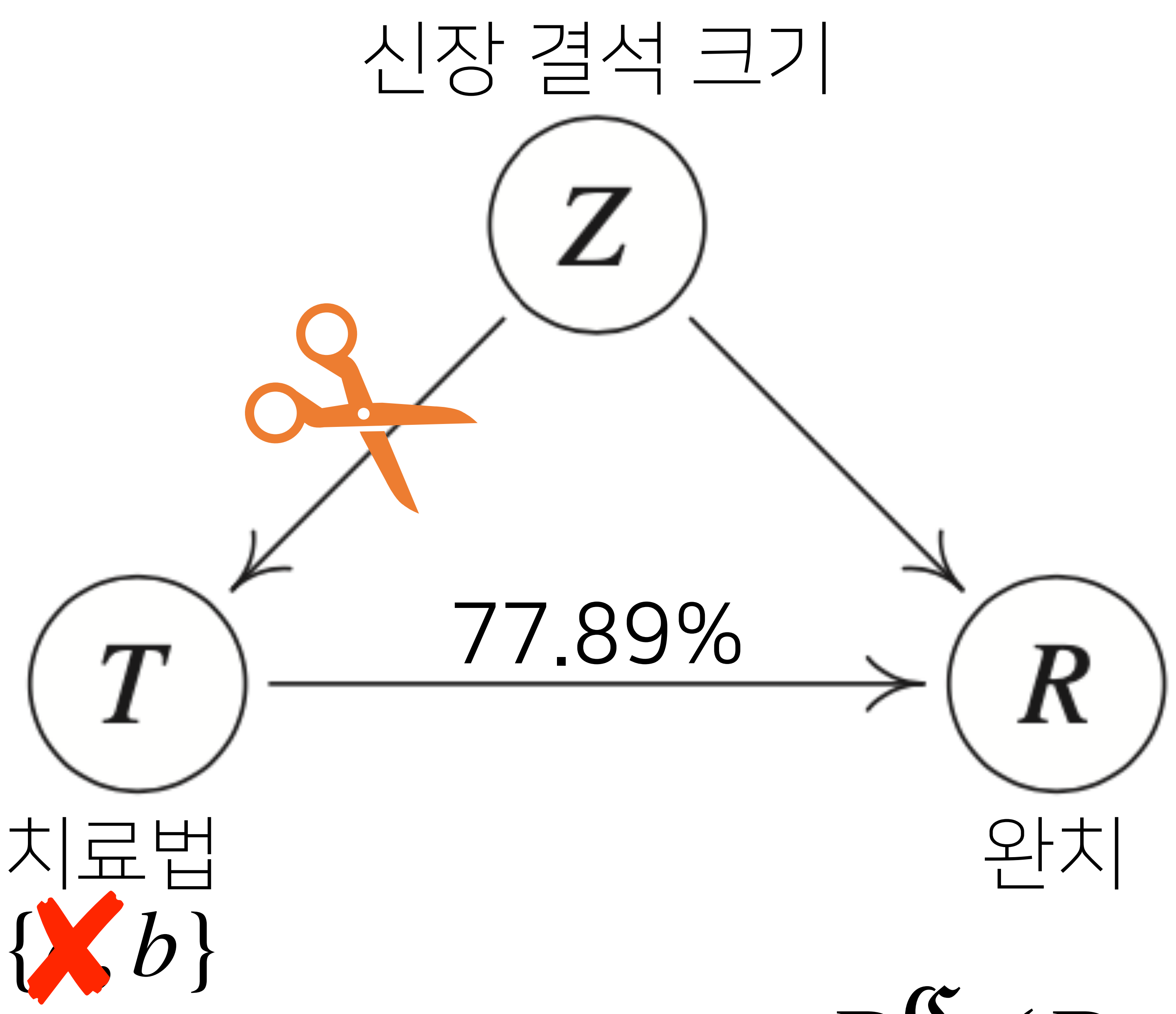
출처: Elements of Causal Inference, Peters et al.

$$P^{\mathfrak{C};do(T:=a)}(R=1) = \sum_{z \in \{0,1\}} P^{\mathfrak{C}}(R=1 \mid T=a, Z=z) P^{\mathfrak{C}}(Z=z)$$

$$= \frac{81}{87} \times \frac{(87+270)}{700} + \frac{192}{263} \times \frac{(263+80)}{700} \approx 0.8325$$

do$(T=a)$ 라는 조정(intervention)
효과를 통해 Z 의 개입을 제거한다

인과관계 추론: 예제



		$Z = 0$	$Z = 1$
	Overall	Patients with small stones	Patients with large stones
Treatment a: Open surgery	78% (273/350)	**93%** (81/87)	**73%** (192/263)
Treatment b: Percutaneous nephrolithotomy	**83%** (289/350)	87% (234/270)	69% (55/80)

출처: Elements of Causal Inference, Peters et al.

$$P^{\mathfrak{C}_b}(R=1) = \sum_{z \in \{0,1\}} P^{\mathfrak{C}}(R=1 \mid T=b, Z=z) P^{\mathfrak{C}}(Z=z)$$

$$= \frac{234}{270} \times \frac{(87+270)}{700} + \frac{55}{80} \times \frac{(263+80)}{700} \approx 0.7789$$

조건부확률로 계산한 치료효과와
정반대의 결과가 나오게 된다

THE END

다음 시간에 보아요!