



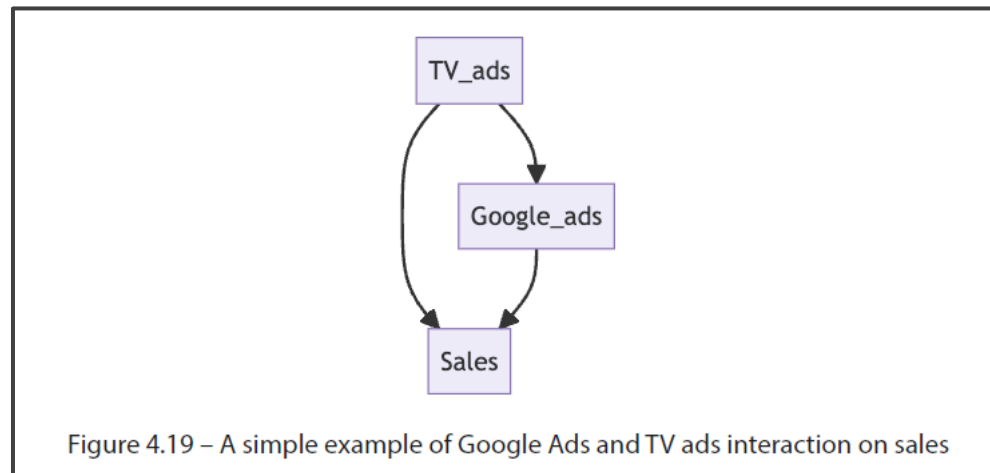
『Marketing Science』

Week 04 : PyMC & TVC

1. DAG
2. PyMC - Plate Notation
3. TVC - 오늘의집
4. TVC - 게임사 사례

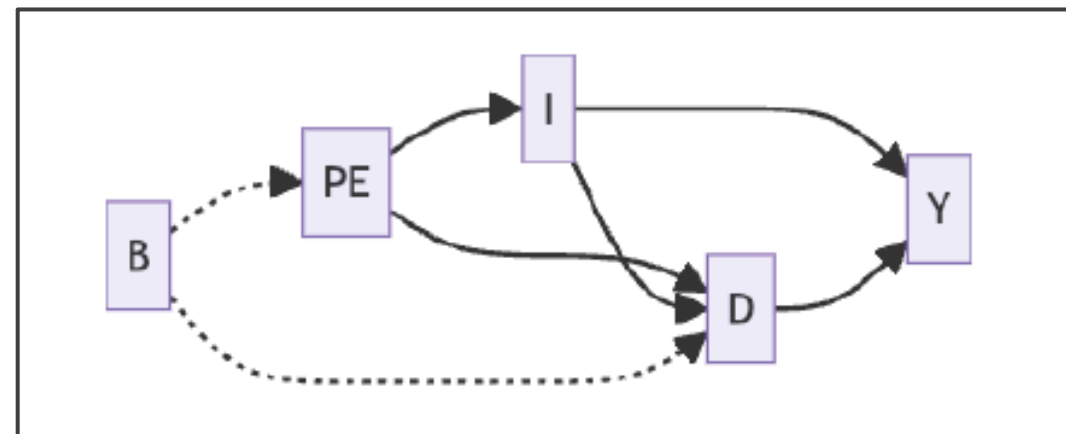
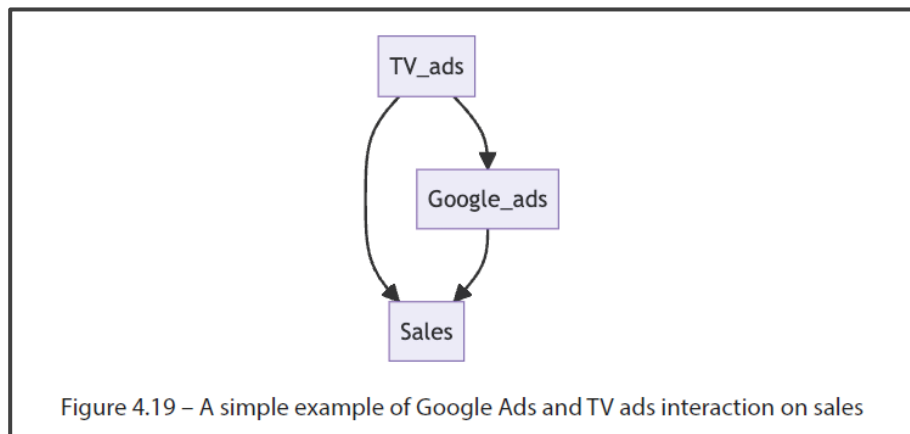
인과관계를 시각적으로 표현하는 수단

- 인과적 구조들을 시각적으로 나타내기 위해 사용
- DAG (유향 비순환 그래프)를 통해 변수간의 영향 관계를 시각적으로 이해하고, 어떤 변수가 인과효과 식별을 위한 보정의 대상이 되어야 하는지를 말해줌

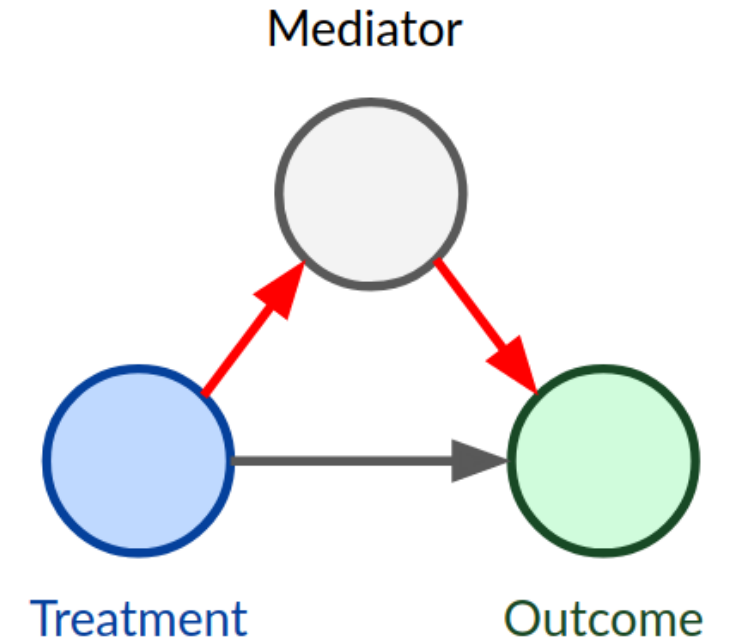
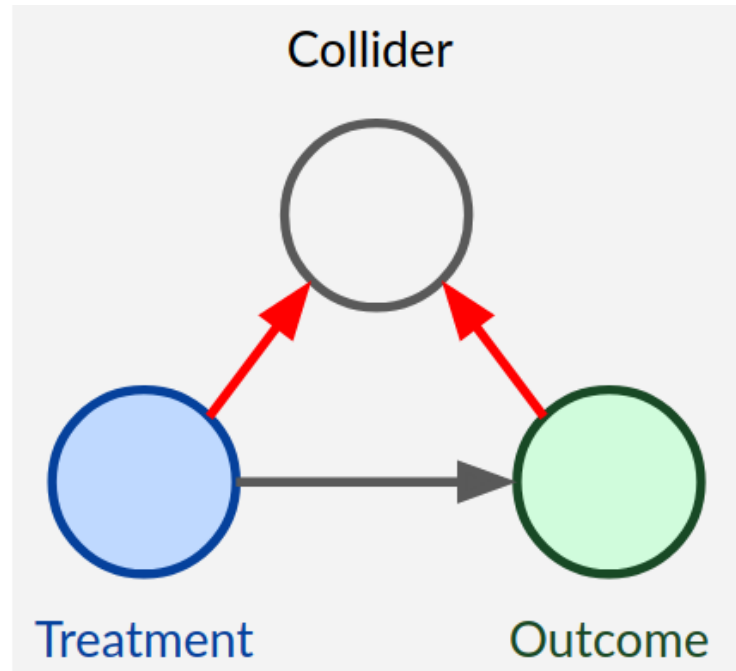
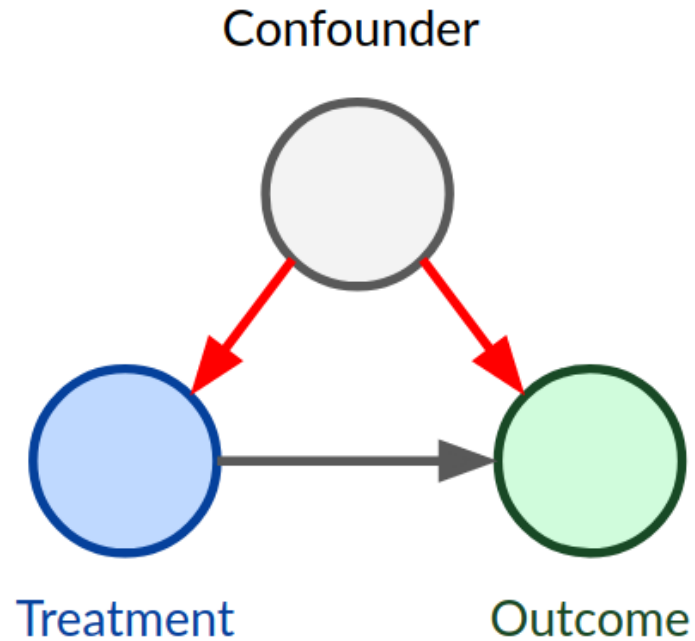


인과적 관계를 직관적으로 표현하는 수단

- 책에서도 인과적 구조들을 직관적으로 나타내기 위해 사용
- DAG 자체는 확률모형인 “베이저안 네트워크” 와 같이 이해하면 좋지만,
보정 (조건부화)에 따른 영향도만 알면 문제없음



대표적인 변수관계 3종

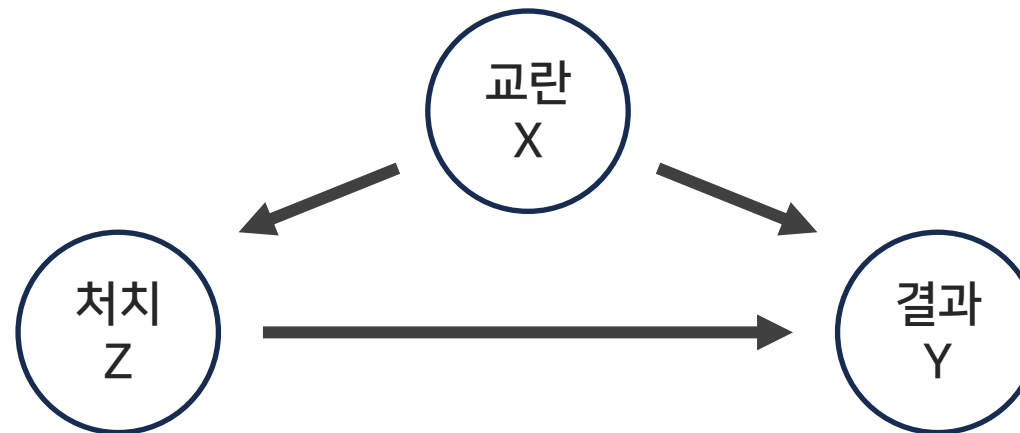


대표적인 변수관계 3종

- Confounder (교란변수) : 처치와 결과 모두에 영향을 주는 변수
 - 인과추론의 적, 보정하지 않으면 문제를 일으킴
- Collider (충돌변수) : 처치와 결과에 영향을 받은 변수
 - 보정하면 안됨. 변수로 보정되는 경우는 없지만, Selection Bias 같은 상황과 연관
- Mediator (매개변수) : 처치에 영향을 받고, 결과에 영향을 주는 변수
 - 보정하면 처치의 직접적 인과효과만을 식별가능

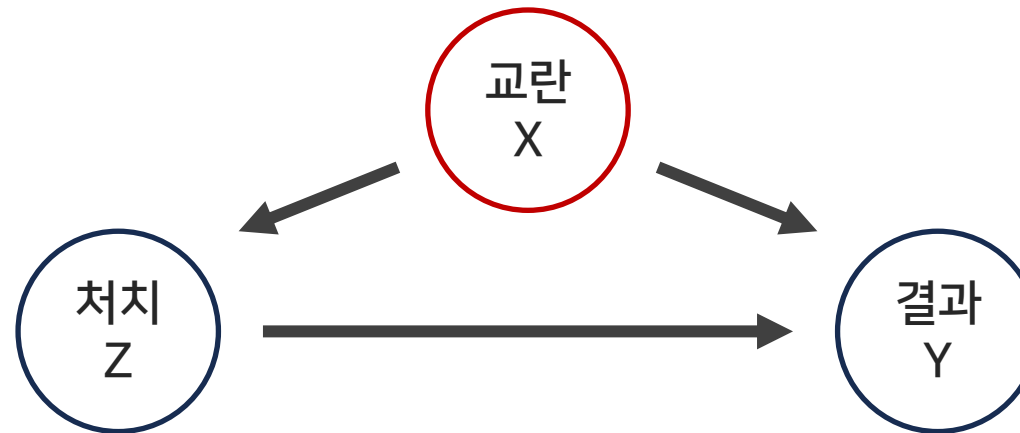
대표적인 변수관계 3종

- 교란은 보정을 통해 무조건 끊어줘야함



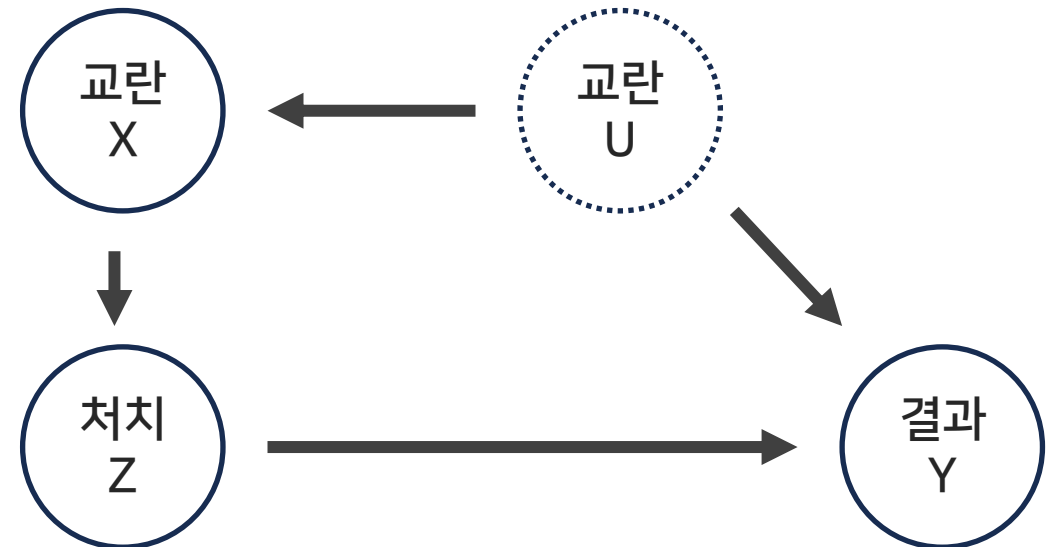
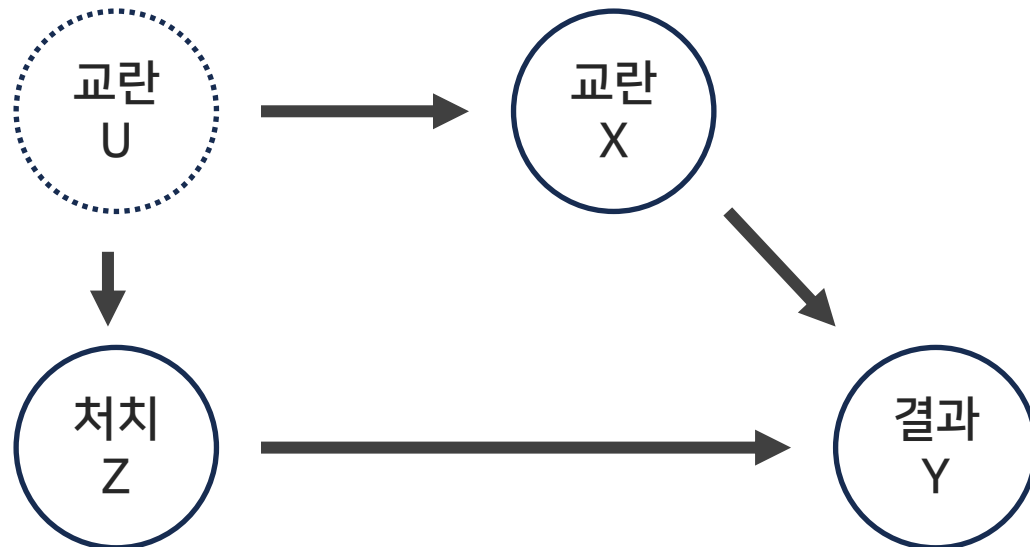
대표적인 변수관계 3종

- 교란은 보정을 통해 무조건 끊어줘야함



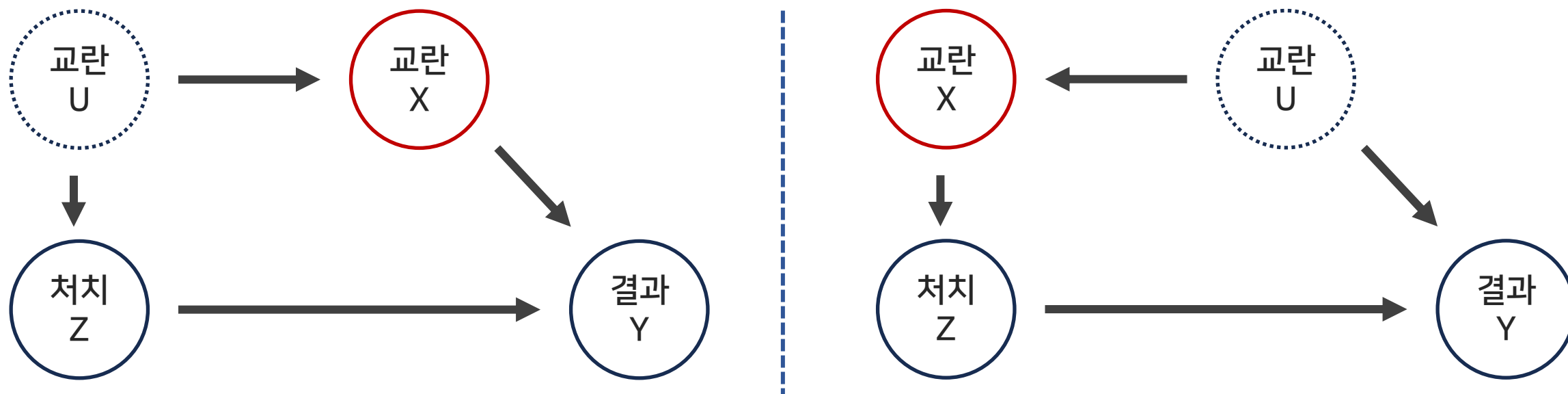
관찰되지 않는 교란변수가 존재한다면?

- 교란은 무조건 끊어줘야함
- 관찰되지 않는 교란이 존재하는 경우...



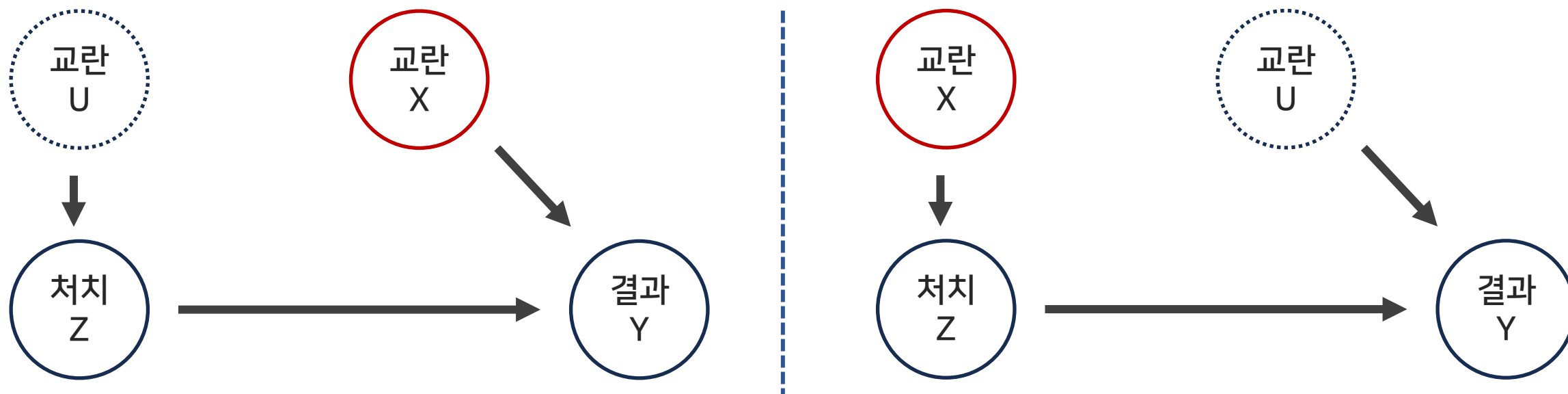
관찰되지 않은 교란변수가 존재한다면? => Proxy

- 관찰되지 않는 교란 U가 존재하는 경우, U의 Proxy 인 X를 보정하는 것으로 충분함



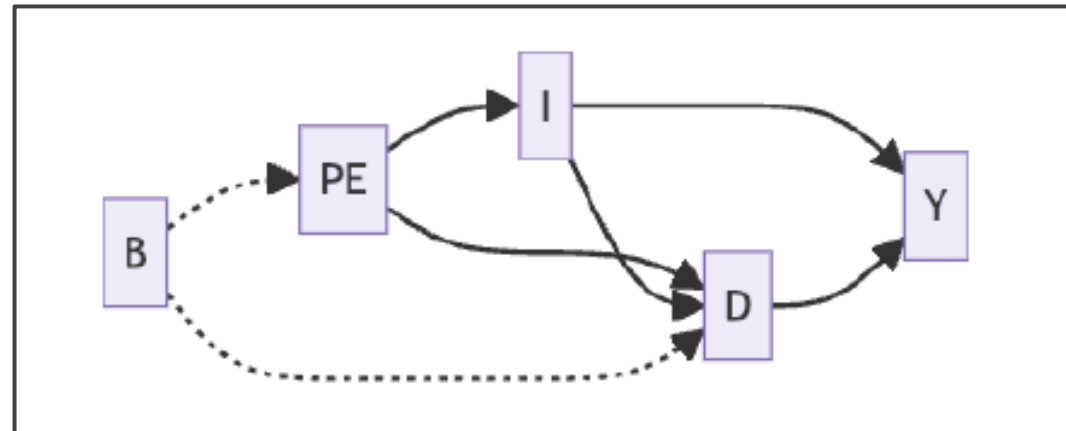
관찰되지 않은 교란변수가 존재한다면? => Proxy

- 관찰되지 않는 교란 U가 존재하는 경우, U의 Proxy 인 X를 보정하는 것으로 충분함



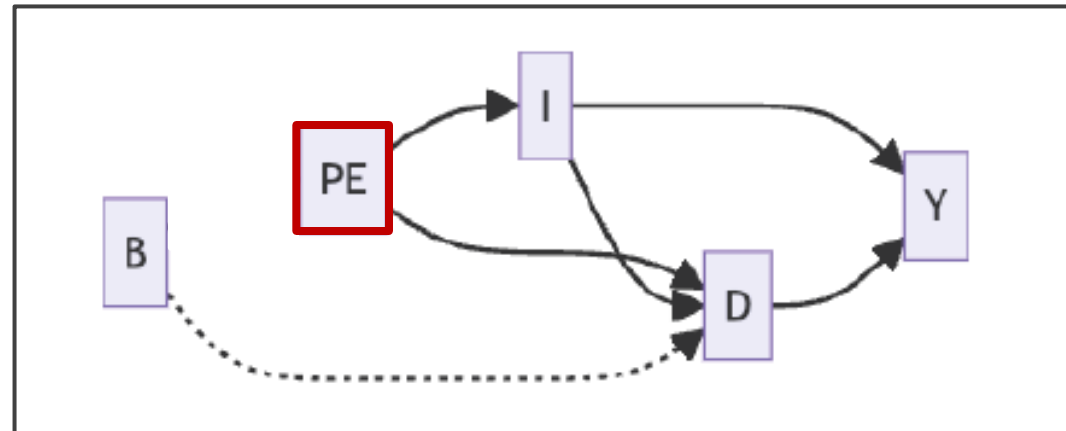
책에서 잘못 설명하고 있어서...

- PE, I 를 보정하면, $D \Rightarrow Y$ 의 인과효과를 추정가능
- B는 PE를 보정하는 순간, 교란변수로서 기능하지 않음



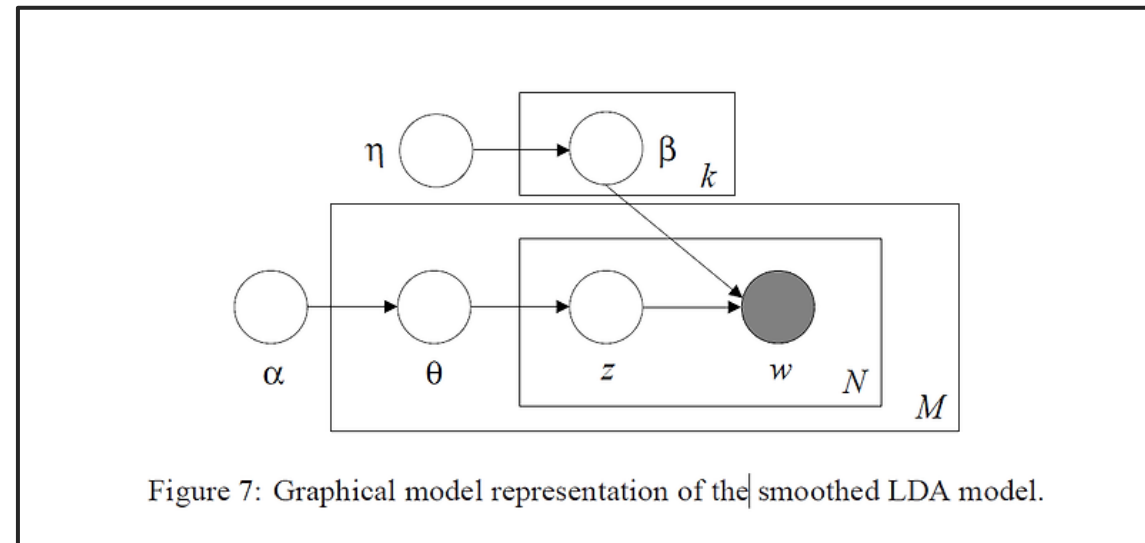
책에서 잘못 설명하고 있어서...

- PE, I 를 보정하면, $D \Rightarrow Y$ 의 인과효과를 추정가능
- B는 PE를 보정하는 순간, 교란변수로서 기능하지 않음



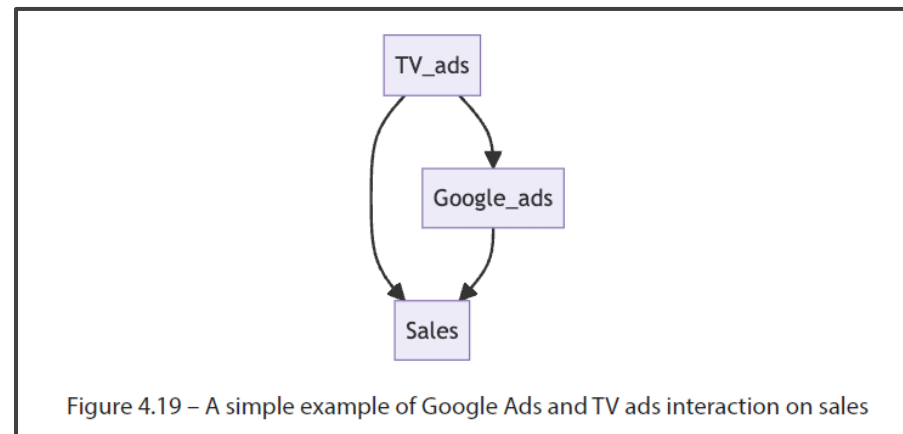
확률분포 모델링을 시각화하는 도구

- 가장 대표적인 예시가 LDA (잠재 디리클레 할당)
- N개의 단어, M개의 문서, k개의 토픽 => 반복되는 변수들에 대해서 사각형으로 묶어서 표현



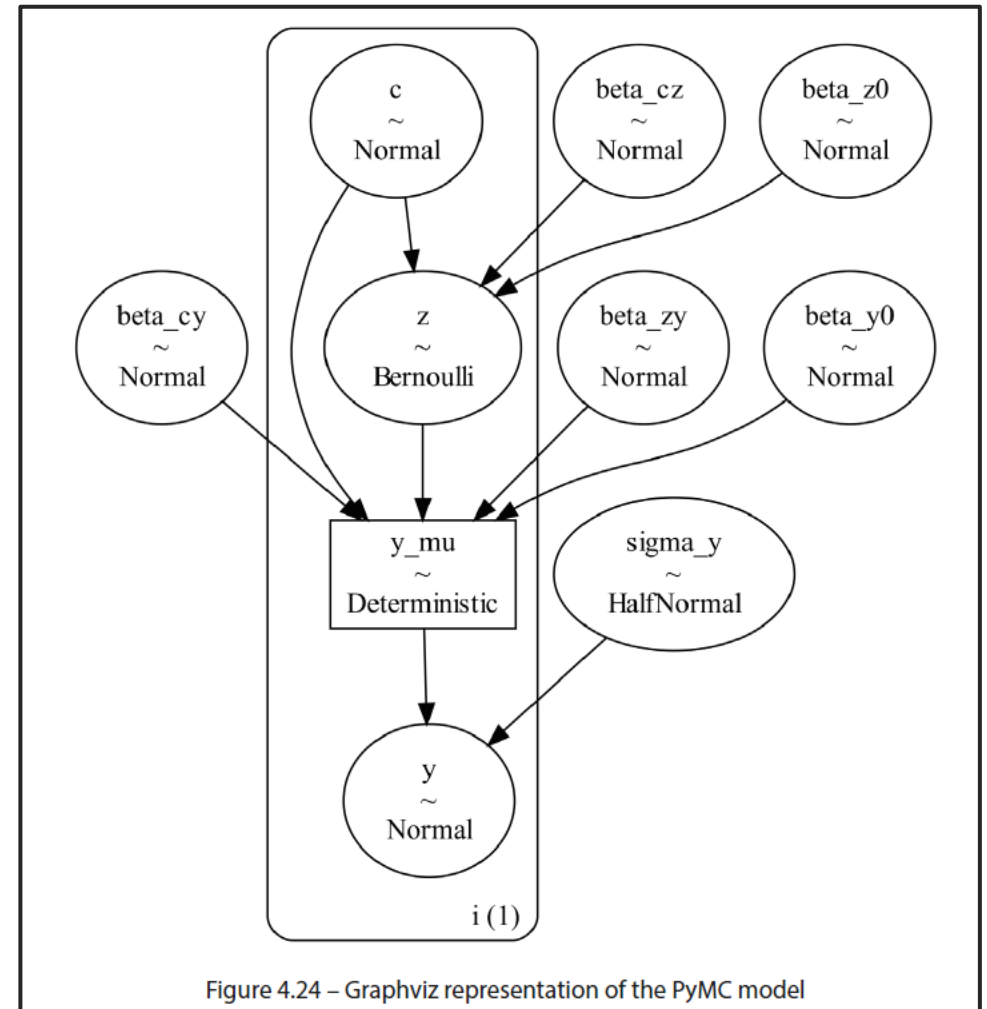
책 예제 – TV광고 + Google 광고효과

- C (confounder) – TV Ads : $C \sim \text{Normal}(0,1)$
- Z (Treatment) – Google On/Off : $Z \sim \text{Ber}(\text{InvLogit}(\beta_z + \beta_c z C))$
- Y (Outcome) – Sales : $Y \sim \text{Normal}(\beta_y + \beta_{cy} C + \beta_{zy} Z, \sigma_y)$



책 예제 - TV광고 + Google 광고효과

```
import pymc as pm
with pm.Model(coords_mutable={"i": [0]}) as model_generative:
    beta_y0 = pm.Normal("beta_y0")
    beta_cy = pm.Normal("beta_cy")
    beta_cz = pm.Normal("beta_cz")
    # priors on Z -> Y causal path
    beta_z0 = pm.Normal("beta_z0")
    beta_zy = pm.Normal("beta_zy")
    # Obs noise on Y
    sigma_y = pm.HalfNormal("sigma_y")
    # core nodes and causal relationships
    c = pm.Normal("c", mu=0, sigma=1, dims="i")
    z = pm.Bernoulli("z", p=pm.invlogit(beta_z0 + beta_cz * c),
                    dims="i")
    y_mu = pm.Deterministic("y_mu", beta_y0 + (beta_zy * z) +
                           (beta_cy * c), dims="i")
    y = pm.Normal("y", mu=y_mu, sigma=sigma_y, dims="i")
```



책 예제 - TV광고 + Google 광고효과

- 책보다 ChatGPT가 더 깔끔하게 잘 작성해줌
 - 수식형태로 지정하지 않아도, 적절하게 작 수행해줌

야 내가 다음과 같은 구조를 지니는 베이지안 모형을 만들건데,
PyMC로 짜줘.

C (confounder) - TV Ads : $C \sim \text{Normal}(0,1)$

Z (Treatment) - Google On/Off : $Z \sim \text{Ber}(\text{InvLogit}(\beta_z + \beta_c zC))$

Y (Outcome) - Sales : $Y \sim \text{Normal}(\beta_y + \beta_c yC + \beta_z y Z, \sigma_y)$

책 예제 - TV광고 + Google 광고효과

```
import pymc as pm
import numpy as np
import arviz as az

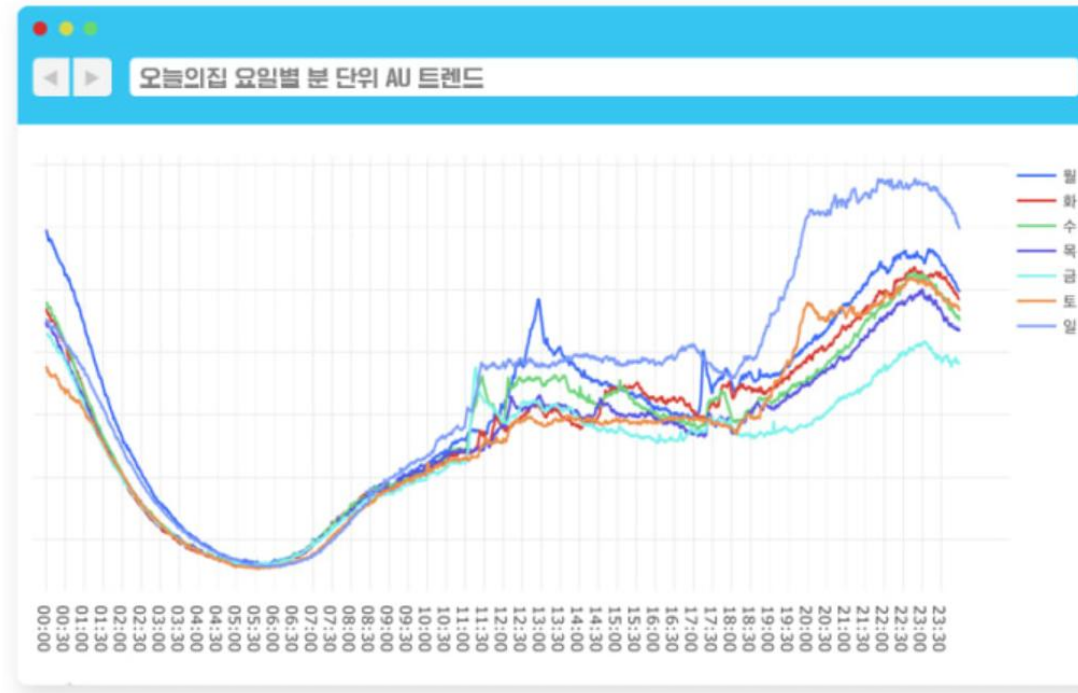
with pm.Model() as model:
    # Priors
    beta_z = pm.Normal("beta_z", 0, 1)
    beta_cz = pm.Normal("beta_cz", 0, 1)
    beta_y = pm.Normal("beta_y", 0, 1)
    beta_cy = pm.Normal("beta_cy", 0, 1)
    beta_zy = pm.Normal("beta_zy", 0, 1)
    sigma_y = pm.HalfNormal("sigma_y", 1)
    # Confounder
    C = pm.Normal("C", 0, 1, shape=N)
    # Treatment assignment (Google On/Off)
    Z_prob = pm.math.invlogit(beta_z + beta_cz * C)
    Z = pm.Bernoulli("Z", Z_prob, observed=Z_true)
    # Outcome (Sales)
    mu_y = beta_y + beta_cy * C + beta_zy * Z
    Y = pm.Normal("Y", mu_y, sigma_y, observed=Y_true)
    # Inference
    trace = pm.sample(1000, return_inferencedata=True, target_accept=0.9)
# 결과 확인
az.summary(trace)
```

오늘의집 사례

- 모바일 UA 광고는 성과추적이 가능하지만, TVC는 직접적 성과추적이 불가능
=> TV 광고를 통한 증분을 측정
- 1. "TV광고가 없었다면 성과가 어땠을까?"
 - 액티브 유저의 증분 / 유입의 증분
- 2. 광고 슬롯별 성과배분
 - 슬롯에 따른 시청자수 보정 위험

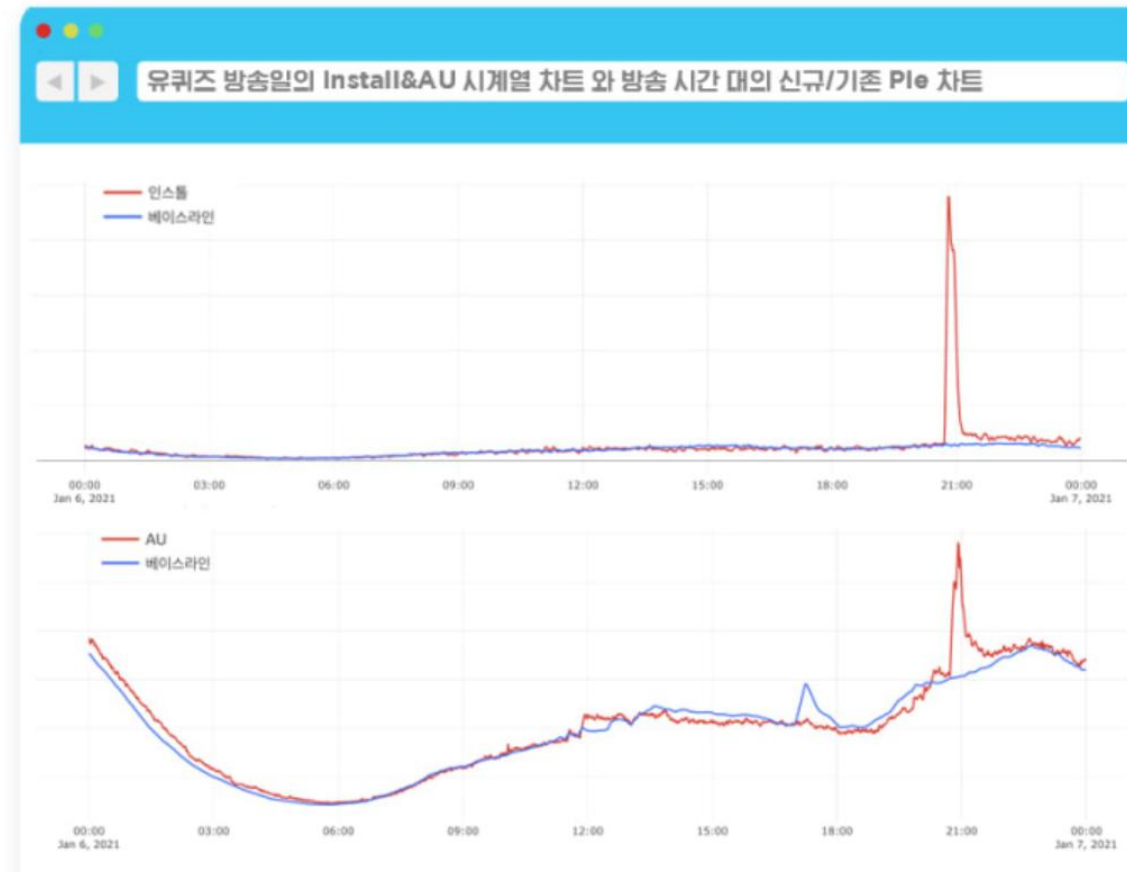
오늘의집 사례 - 시간대별 증분

- 기존데이터들을 통해 베이스라인을 추정
- 베이스라인을 counterfactual로 생각



오늘의집 사례 - 시간대별 증분

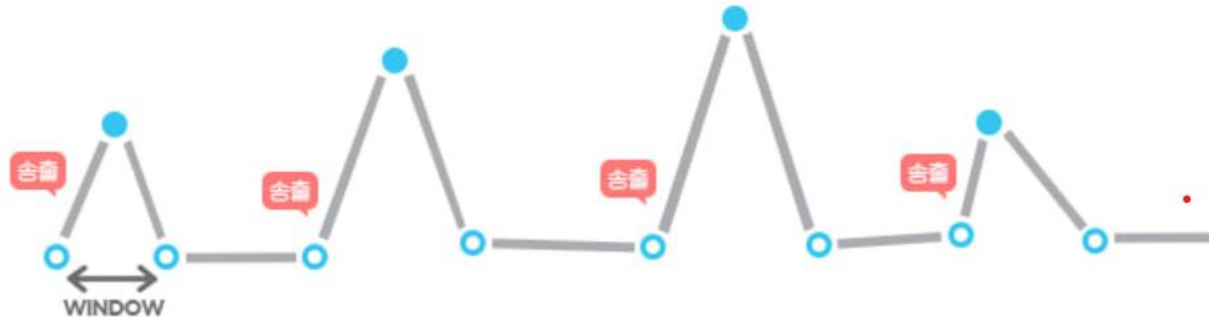
- 이승재 대표의 유퀴즈 출연을 통해 성과 확인방법 검증



오늘의집 사례 - 슬롯별 증분

- 광고간의 간격이 있는 경우, 별도의 보정작업 없이도 효과추정이 가능

중복 없이 송출되는 광고의 spike 예시



오늘의집 사례 - 슬롯별 증분

- 광고간의 간격이 없는 경우, 광고 슬롯별 성과 파악이 어려움
- 슬롯별 성과파악이 중요한 이유

=> 추정 : 광고 슬롯별로 비용이 다를 것이기 때문

- 비용에 따라 시청자수/유입효과가 비례하지 않을 수 있어, 추가적인 성과 파악이 필요

오늘의집 사례 - 슬롯별 증분

광고슬롯별 송출시간 및 시청자수		송출시간		시청자수			
	A	19:00		200			
	B	19:00		400			
	C	19:01		300			
	D	19:03		100			
시간대		19:00	19:01	19:02	19:03	19:04	19:05
총 증분		100	150	140	70	50	30
슬롯별 증분 분배	A	33	40	23			
	B	67	80	47			
	C	30		70	53		
	D				18	50	30
분배된 증분의 합		100	150	140	70	50	30

게임사 TVC 사례

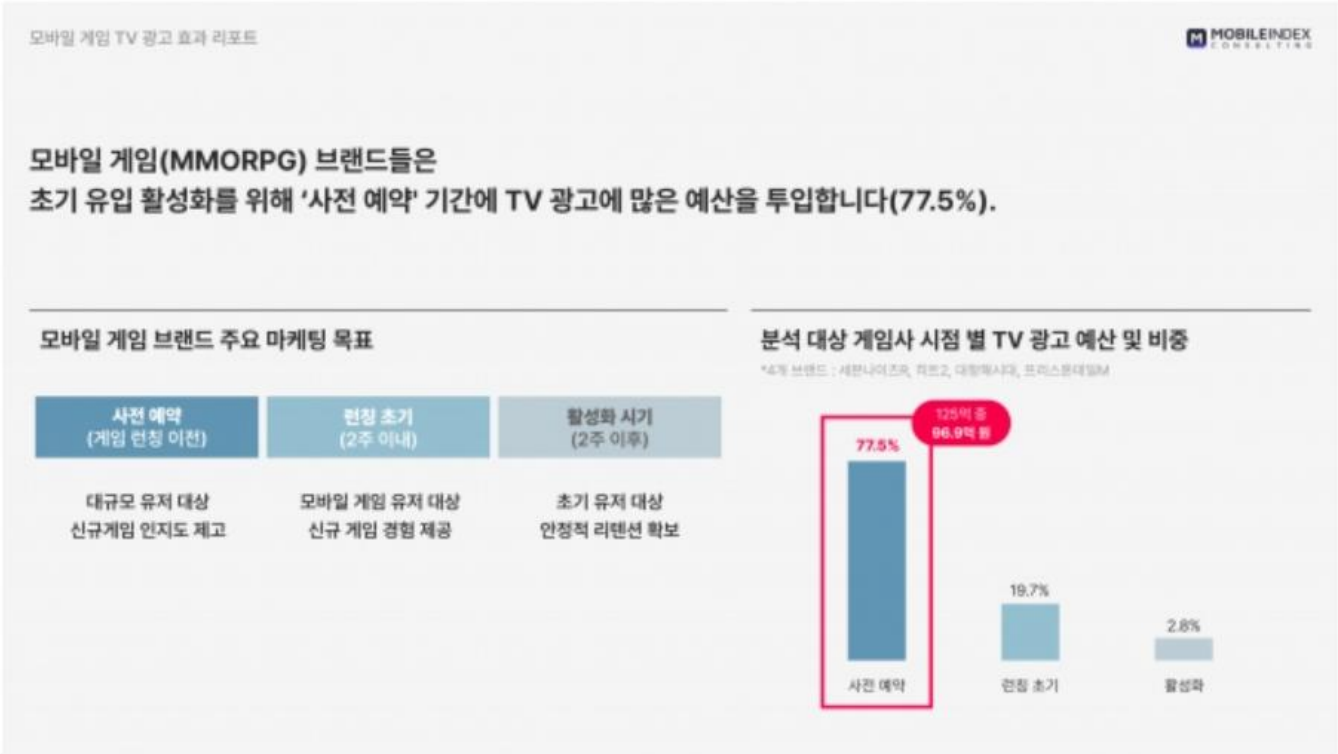
- MMORPG는 대세감 - 주로大作마케팅 성향이 있는 MMORPG에서 TVC 진행
- 정상급 배우를 모델로 쓰고, 웅-장해지는 느낌



게임사 TVC 사례

- TV광고 예산 중에 사전예약 기간 80%, 론칭 초기 20% 사용 (모바일인덱스 결과)
 - 인지도 및 대세감을 위한 광고 (초기 모객단가 낮추기 위함)
- 모바일게임은 TV광고를 통한 전환율이 타업종 대비 120배 높음
- 모바일게이머는 상대적으로 실내 활동을 즐겨서 더 효율이 좋음

게임사 TVC 사례



2022.06~2022.09 : 세븐나이츠레볼루션 / 히트2 / 대항해시대 / 프리스톤테일M

참고자료 : 모바일인덱스 ([매드타임스 기사](#))

TV광고 성과를 추적하고 있을까?

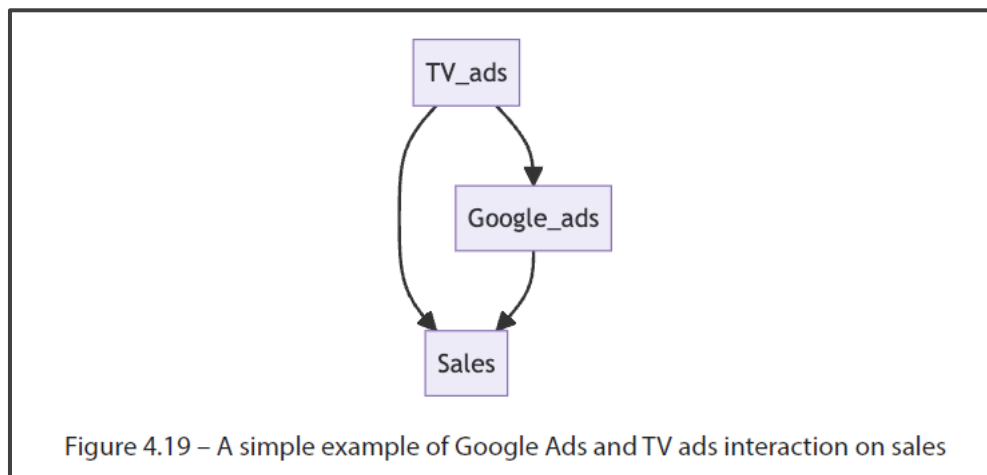
- MMORPG 게임의 TVC / OOH (옥외광고)는 기본이라고 여겨짐
 - TVC + 강남역 옥외광고를 여러 구좌에 수행하는 전략
- '스픽'같은 애드테크 회사에서도 강남역 옥외광고를 여러 구좌에 수행



TVC 성과 확인을 위한 고려 사항

1. 성과를 전체 증분만으로 성과를 판단해서는 안됨

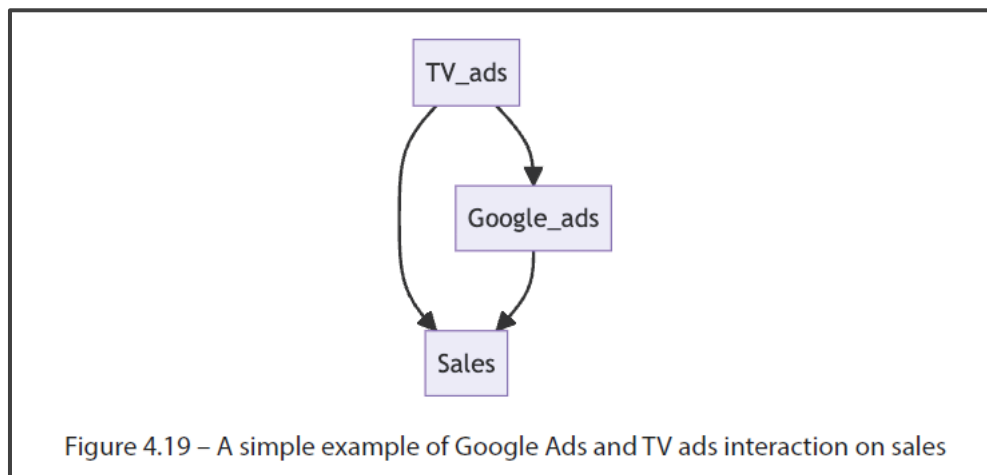
- TVC를 통해 자연유입도 늘어나지만, 광고유입도 늘어나기 때문



TVC 성과 확인을 위한 고려 사항

2. 성과를 비용 기반으로 파악해야함

- TVC 비용은 매우 비싸기 때문에, TVC를 통해 얻은 증분과 미시행시의 성과를 비교해서 판단해야함



베이지안 툴 PyMC, Numpyro 기반 (마케팅) 분석

- [Juan Camilo Orduz 블로그](#) : PyMC Labs 소속
 - Cohort Retention Analysis [1](#), [2](#), [3](#)
 - [Product Life Cycle Modeling](#), MMM , CLV , ...
- [PyMC-Marketing](#) 공식 페이지
 - Marketing Mix Models (MMM), Customer Lifetime Value (CLV, pLTV),
Customer Choice (Multivariate Interrupted Time Series)

감사합니다.
