一般化傾向スコアの利用法と広告効果測定における実践

経済学部4年　原田悠介

内容

1. 一般化傾向スコアとは？
2. 広告効果測定での利用(別スライド)

1.一般化傾向スコアとは？

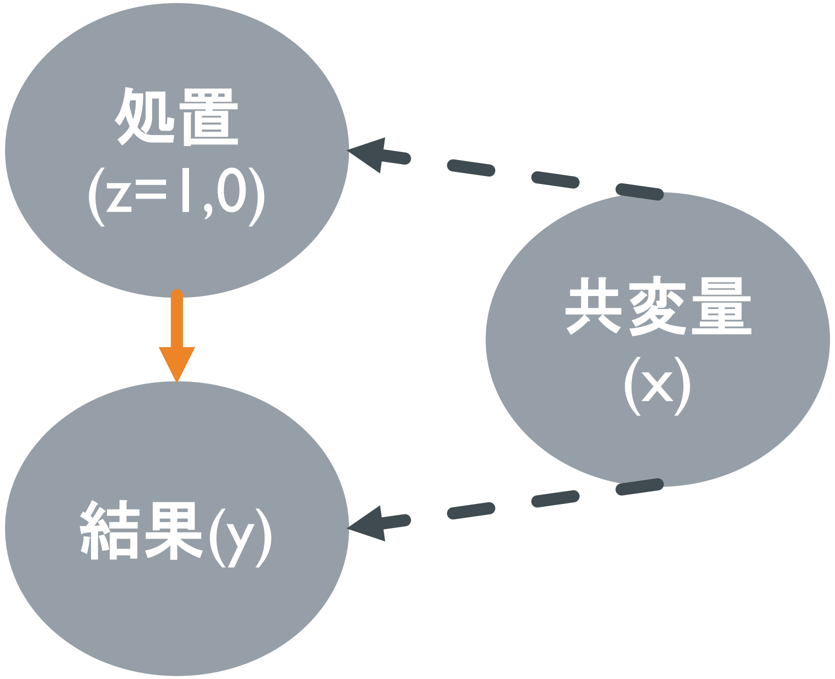
1. 因果効果とは
2. 因果効果の推定
3. 無作為割り当て
4. 観察研究における因果効果の推定
5. 傾向スコアの利用法
6. 一般化傾向スコアの利用法

因果効果とは

ある処置が結果に与える因果関係の強さを測りたい！

相関分析、(重)回帰分析

→だいたいの場合交絡による内生性のため因果関係の強さとは不一致



zの例	yの例	xの例
投薬	病気の快復	体力・年齢
保育園通園	子供の知能	親の学歴・世帯収入
広告	購買	ライフスタイル

因果効果とは

ルービンの因果モデル

y_{1i} ：個体 i が処置を受けたとき$(z = 1)$の(決定論*的)結果

y_{0i} : 個体 i が処置を受けないとき$(z = 0)$の(決定論的)結果

個体 i における因果関係の強さ = 因果効果は

$$y_{1i} - y_{0i}$$

で計算される

(*決定論:あらゆる事象の世紀と帰結が魔rもって決定されているという立場)

因果効果の推定

“fundamental problem of causal inference”(Holland,1986)

個人 i について y_{1i} , y_{0i} の両方を同時に観測することはできない

つまり、個人の因果効果を直接計算することはできない

→集団で因果効果の期待値を考える

平均処置効果 (Average Treatment Effect)

$$\frac{1}{n}\sum_{i=1}^{n}(y_{1i}-y_{0i})=E[y_1]-E[y_0]$$

１はこのATEをどうやって推定するのかについての話です

因果効果の推定

ATE : $E[y_1] - E[y_0]$ も欠測を含む(下図)が、

推定できる場合がある

	処置群(z = 1)	非処置群(z = 0)
y_1	観測 $E[y_1 \mid z=1]$	欠測 $E[y_1 \mid z=0]$
y_0	欠測 $E[y_0 \mid z=1]$	観測 $E[y_0 \mid z=0]$
	共変量	

$E[y_1]$

$E[y_0]$

無作為割り当て

ATEを観測データから推定するための条件

$$E[y_1|z=1]-E[y_0|z=1]$$

$$= \underbrace{E[y_1-y_0|z=1]}_{\text{A}} + \underbrace{E[y_0|z=1]-E[y_0|z=0]}_{\text{B選択性バイアス}}$$

- $E[y_1-y_0|z=1]=E[y_1-y_0|z=0]$ であればAがATEと一致
- B選択性バイアスについて、$E[y_0|z=1]-E[y_0|z=0]=0$

→ yとzが独立 $(y_1,y_0)\perp z$ であれば満たせる

無作為割り当て

yとzが独立 $(y_1, y_0) \perp z$ のとき欠測を無視してATEを推定できる

	処置群(z = 1)		非処置群(z = 0)		
y_1	観測 $E[y_1 \mid z=1]$	=	欠測 $E[y_1 \mid z=0]$	=	$E[y_1]$
y_0	欠測 $E[y_0 \mid z=1]$	=	観測 $E[y_0 \mid z=0]$	=	$E[y_0]$
	共変量				

$$\boxed{E[y_1 \mid z=1] - E[y_0 \mid z=0] \ = \ E[y_1] - E[y_0]}$$

無作為割り当て

yとzを独立 $(y_1, y_0) \perp z$にするためには…

処置zを各個体 i にランダムに割り当てる無作為割り当て(ランダム化実験)をすればよい

無作為割り当てが行えれば共変量xに関わらず群間平均の差が因果効果になっている

→ “fundamental problem of causal inference”を解決！

観察研究における因果効果の推定

無作為割り当ては優れた手法であるが、社会科学や臨床医学などの観察研究では様々な問題から行えないことが多い

- 倫理的な問題(医療や教育など)
- 不遵守
- 研究の生態学的妥当性を欠く(ホーソン効果、ローゼンタール効果など)

観察研究における因果効果の推定

無作為割り当てが行えない場合は共変量xの影響を除去する必要がある

条件付独立 $(y_1, y_0) \perp z|x$ を仮定

→「共変量で調整すれば結果yは処置の有無zに依存しない」という仮定

共変量が同じ個体は$E[y_1], E[y_0]$も同じと言い換えられる

観察研究における因果効果の推定

yとzがxによる条件付独立 $(y_1, y_0) \perp z|x$ のとき欠測を無視してATEを推定できる

	処置群(z = 1)		非処置群(z = 0)		
y_1	観測 $E[y_1 \| z = 1, x]$	=	欠測 $E[y_1 \| z = 0, x]$	=	$E[y_1 \| x]$
y_0	欠測 $E[y_0 \| z = 1, x]$	=	観測 $E[y_0 \| z = 0, x]$	=	$E[y_0 \| x]$
	共変量(x)				

$$E[E[y_1|z = 1, x] - E[y_0|z = 0, x]] = E[y_1] - E[y_{]}]$$

観察研究における因果効果の推定

共変量調整による因果効果の推定法

(1)マッチング・層別化 : 共変量が「近い」個体を比較する

- 共変量が増えるほど「近い」個体が見つかりにくい
  (次元の呪い)
- 「近い」の決め方が恣意的

(2)回帰モデルに共変量を組み込む

- 結果変数yと共変量xの正しいモデリングが必要

(3)傾向スコア

傾向スコアの利用法

傾向スコアとは

個体 i の共変量ベクトルをx_iとしたとき、処置群へと割り当てられる条件付確率

$$e_i = p(z_i = 1|x_i)$$

→ロジスティック回帰やプロビット回帰で推定できる

複数の共変量を一変数に集約しているため、次元の呪いが起こらない

傾向スコアの利用法

傾向スコアによる共変量調整はすべての共変量を用いた調整と同じ効果がある

	処置群(z = 1)		非処置群(z = 0)		
y_1	観測 $E[y_1 \mid z=1, e]$	=	欠測 $E[y_1 \mid z=0, e]$	=	$E[y_1 \mid e]$
y_0	欠測 $E[y_0 \mid z=1, e]$	=	観測 $E[y_0 \mid z=0, e]$	=	$E[y_0 \mid e]$
	共変量(x)→傾向スコア(e)				

$$E[E[y_1 \mid z=1, e] - E[y_0 \mid z=0, e]] = E[y_1] - E[y]$$

傾向スコアの利用法

傾向スコアを用いた調整方法

①傾向スコアの推定

処置の有無 z を共変量 x によって説明するモデルを設定

$$z_i \sim \frac{1}{\{1+\exp[-(\alpha+\sum \beta x_i)]\}}$$ (ロジスティック回帰)

母数の推定量を用いて予測確率を計算(= 傾向スコアの推定値)

$$e_i = \frac{1}{\{1+\exp[-(\hat{\alpha}+\sum \widehat{\beta^T} x_i)]\}}$$

傾向スコアの利用法

②推定された傾向スコアを用いた調整

1. マッチング

2群で傾向スコアが近い個体をペアにして

その差の平均をATEの推定量とする

＜欠点＞

・$E[y_1], E[y_0]$を直接計算できない

・「近い」の基準が恣意的

・個体数の少ない集団に合わせるため、個体数が減少する

傾向スコアの利用法

2.層別化

　傾向スコアの大小で5つ以上のサブクラスにわけ、各クラスでの処置群と非処置群の平均を計算、全体としての効果の推定量を計算する

$$ATE = \sum_{k=1}^{K} \frac{N_k}{N} (\bar{y}_{1k} - \bar{y}_{0k})$$

(N: 全個体数、N_k: 第kクラスのサンプル数、$\bar{y}_{1k}, \bar{y}_{0k}$: kクラスでの処置群と非処置群の平均)

＜欠点＞

・$E[y_1], E[y_0]$を直接計算できない

・傾向スコアが離れた個体が同クラスに分類される可能性がある、また5つ以上の層別が適切であるという理論的な保証はない

傾向スコアの利用法

3.共分散分析

処置 z と傾向スコア e を説明変数とした線形回帰を行う

＜欠点＞

傾向スコアと目的変数が線形な関係にあるという仮定に無理がある

傾向スコアの利用法

4.IPW推定量

傾向スコアの逆数で重みづけ平均を取ることで$E[y_1], E[y_0]$
を直接計算する

$$\hat{E}[y_1] = \sum_{i=1}^{N} \frac{z_i y_i}{e_i} / \frac{z_i}{e_i} \ , \ \hat{E}[y_0] = \sum_{i=1}^{N} \frac{(1-z_i) y_i}{1-e_i} / \frac{1-z_i}{1-e_i}$$

$$ATE = \hat{E}[y_1] - \hat{E}[y_0]$$

一般化傾向スコアの利用法

通常の傾向スコアは処置zの有無の二群でしか利用できない

→複数の処置を比較できるように傾向スコアを拡張したものが
一般化傾向スコア

傾向スコア

$$z = \begin{cases} 1 & (\text{処置有}) \\ 0 & (\text{処置無}) \end{cases}$$

一般化傾向スコア

$$z_j = \begin{cases} 1 & (\text{処置}\, j\, \text{有}) \\ 0 & (\text{処置}\, j\, \text{無}) \end{cases}$$

$$j = 1,2,\dots,J$$

一般化傾向スコアの利用法

一般化傾向スコアとは

個体 i の共変量ベクトルをx_iとしたとき、第 j 群へと割り当てられる条件付確率

$$R(j, x_i) = p(z_{ji} = 1 | x_i)$$

→多項ロジスティック回帰で推定できる

一般化傾向スコアの利用法

一般化傾向スコアを用いた調整方法

①一般化傾向スコアの推定

処置 j の有無 z_j を共変量 x によって説明するモデルを設定

母数の推定量を用いて予測確率を計算

(= 一般化傾向スコアの推定値)

$$R(t, x_i) = \frac{exp(z_j)}{1 + \sum_{k=1} exp(z_k)}, \quad z_j = \widehat{\alpha_j} + \widehat{\beta_j^T} x_i$$

一般化傾向スコアの利用法

②推定された一般化傾向スコアを用いた調整

→IPW推定量を用いる

$$\hat{E}[y_j] = \sum_{i=1}^{N} \frac{z_{ji} y_i}{R(j, x_i)} \Big/ \frac{z_{ji}}{R(j, x_i)}$$

第 a 群と第 b 群の $ATE = \hat{E}[y_a] - \hat{E}[y_b]$

2.広告効果測定での利用

1. 研究の背景と目的
2. 広告効果測定における因果推論の必要性
3. リサーチクエスチョンと分析方法
4. データ概要と変数の定義
5. 分析
6. 考察