調査観察データの統計科学ゼミ-第 13 回-

Masamichi Ito

Osaka University Graduate School of Human Sciences
Adachi Lab M1

January 16, 2020

担当章

感度分析

" 隠れた共変量" が因果効果の推定値に与える影響を測る方法が必要！

感度分析

-" 隠れた共変量" を明示的にモデリングし，その影響力を変化させた場合に得られる因果効果の推定値がどの範囲で変動するかを調べること．

(偏微分に似た考え方！一つだけ動かしてどれだけ変化するか？どんなふうに変化するか？を考える)

Rosenbaum の方法

結果変数が 2 値 (0,1 とする) の場合に利用できる方法を紹介する．傾向スコアを算出するロジスティック回帰モデルで，

$$logit(p(z_i = 1|\boldsymbol{x}_i)) = \log\left(\frac{p(z_i = 1|\boldsymbol{x}_i)}{1 - p(z_i = 1|\boldsymbol{x}_i)}\right) = g(\boldsymbol{x}_i) + \gamma u_i \qquad (4.4)$$

のように，処置群 ($z_i = 1$) に割り当てられる確率のロジットが，共変量の関数だけでなく，隠れた共変量 $u_i (0 \leq u_i \leq 1)$ にも依存すると考え，その係数 γ を変化させることで，影響の程度を調べる．

Rosenbaum の方法

ここで $\Gamma := \exp\gamma$ とすると，(4.4) 式から，
全く同じ共変量の値を持つ2 つの対象者 i, i' についてのオッズ比が，

$$\frac{1}{\Gamma} \leq \frac{p(z_i = 1|\boldsymbol{x}_i)(1 - p(z_{i'} = 1|\boldsymbol{x}_{i'}))}{p(z_{i'} = 1|\boldsymbol{x}_{i'})(1 - p(z_i = 1|\boldsymbol{x}_i))} \leq \Gamma$$

となる．この性質を利用して，様々なノンパラメトリックな検定に対応する感度分析法を提案している (Rosenbaum, 2002 を参照).

様々なノンパラメトリックな検定

ex. McNemar 検定, Wilcoxon の符号順位検定, Mantel-Haenszel 検定 etc.

次項では，傾向スコアによってマッチングしたペアに対する McNemar 検定での感度分析について紹介

Supplementary explanation for the equation in the previous slide

To ease typing, we denote $p(z_i = 1|\boldsymbol{x}_i)$ and $p(z_{i'} = 1|\boldsymbol{x}_{i'})$ as p and p' respectively. Using this notation, the equation can be expressed as

$$\frac{1}{\Gamma} \leq \frac{p(1-p')}{p'(1-p)} \leq \Gamma$$

Please note that $p/(1-p) = \exp(g(\boldsymbol{x}_i) + \gamma u_i)$. Here, we assume covariates i-th and i'-th individualare are idenical, that is,$\boldsymbol{x}_i = \boldsymbol{x}_{i'}$. Then $\exp(g(\boldsymbol{x}_i)) = \exp(g(\boldsymbol{x}_{i'}))$ and, the term can be rewritten as

$$\frac{\exp(g(\boldsymbol{x}_i))\exp(\gamma u_i)}{\exp(g(\boldsymbol{x}_{i'}))\exp(\gamma u_{i'})} = \frac{\exp(\gamma u_i)}{\exp(\gamma u_{i'})}$$

$$\Leftrightarrow \frac{1}{\Gamma} \leq \frac{\exp(\gamma u_i)}{\exp(\gamma u_{i'})} \leq \Gamma \quad \square$$

In the last transformation, we used $0 \leq u_i \leq 1$ and $e^{\gamma u_i} \leq e^{\gamma} = \Gamma$.

PS でマッチングしたペアに対する McNemar 検定

(PS はもちろん傾向スコア Propensity Score の略)
McNemar 検定では，マッチングされたペアにおいて結果変数の値が一致しない ($(0,1)$ と $(1,0)$) 確率がペアで等質であるという仮定のもとでは $1/2$ となることを仮定する．マッチングされたペアのうち，結果変数の値が一致しないペアの度数を N，$(0,1),(1,0)$ の多い方の度数を a，結果変数の値が一致しないペアの度数を表す確率変数を X とすると，p 値は

$$P(X \geq a) = \sum_{x=a}^{N} {}_N C_x (1/2)^x (1/2)^{N-x} = \sum_{x=a}^{N} {}_N C_x (1/2)^N$$

と計算できる．

Rosenbaum の方法続き

隠れた共変量によってペアにおいて結果変数の値が一致しない確率の最大値は $\Gamma/(1+\Gamma)$，最小値は $1/(1+\Gamma)$ となる．Rosenbaum はの感度分析はこれを用いて，p 値の上限，下限を

$$\sum_{x=a}^{N} {}_N C_x (\Gamma/(1+\Gamma))^x (1/(1+\Gamma))^{N-x} \tag{4.5}$$

および

$$\sum_{x=a}^{N} {}_N C_x (1/(1+\Gamma))^x (\Gamma/(1+\Gamma))^{N-x} \tag{4.6}$$

とするものである．
Γ の値を大きくしていくことで，隠れた共変量の影響が大きい場合でも因果効果の検定結果がどれくらいロバストであるか調べよう!
(もちろん，Γ の大きさによってどちらが上限下限になるかが異なる)

Rosenbaum の方法続き

(一旦落ち着いて)，(4.4) 式だけでは，「隠れた共変量が結果変数に与える影響」を考慮できていない．
$\rightarrow$ 隠れた共変量が因果効果の推定に及ぼす効果を過小評価する可能性有り
$\rightarrow$「割り当て (z_i)を説明するモデル」「結果変数 y_iを説明するモデル」の両方において 隠れた共変量を説明変数とするモデルを利用した方法も提案されている．

Imbens(2003) の方法

両方において隠れった共変量を説明変数とするモデルとして，例えば，Imbens(2003) は，隠れた共変量 u_i は 2 値変数 (=0,1) とし，(4.4) と同様に，割り当てが，

$$Pr(z_i = 1|\boldsymbol{x}_i, u_i) = \frac{1}{1 + \exp(-\{\boldsymbol{\alpha}^t \boldsymbol{x}_i + \gamma u_i\})} \tag{4.7}$$

に従うと考える．さらに結果変数は，割り当てと，観測された/されない共変量によって説明される，つまり，

$$y_i = \tau_z z_i + \boldsymbol{\beta}^t \boldsymbol{x}_i + \delta u_i + \epsilon_i$$

と考える．ただし，$\epsilon \sim N(0, \sigma)$．ここで，$\gamma, \delta$ を固定すれば，他のパラメータは最尤法で推定できる．

Imbens(2003) の方法続き

$\rightarrow \gamma, \delta$ を様々な値に変えながら推定を行うことで，

- 「隠れた共変量が割り当て z や結果変数 y にどれくらい影響力があるか」
- 「因果効果の推定値がどの程度変化するか」

を推論可能.

ただし，γ, δ はそのままでは解釈しにくい…

$\rightarrow$Imbens(2003) では，隠れた共変量の影響力の指標として隠れた共変量による重相関係数を利用している.

他の隠れた共変量の扱い方

以上 2 つの例は，隠れた共変量をモデリングすることによる感度分析法
これ以外には，

- 「割り当てが結果変数に依存する場合」のモデルを利用する (4.4 節)
  - パラメトリックなモデルを利用する場合においては，完全尤度において，ψ などの識別されない母数をいろいろ変化させることで，因果効果の推定値がどう変化するかを見る
- 仮定したモデル (y, z の同時分布) が，真のモデルからある程度逸脱している場合にどのような結果が得られるかを調べる．
  - 強く無視できる割り当て条件に従うモデル $p(y_1, y_0, z|\boldsymbol{x}) = p(y_1, y_0)p(z|\boldsymbol{x})$ が正しいのではなく，それにある程度近いモデル $g(y_1, y_0, z|\boldsymbol{x})$ が正しいとする．このように複数モデルを用意して，推定値どう変化するかを見る．

当然の疑問達

これまで，調査観察データから" 因果効果" の推定を行う方法や具体例を扱ってきた．でも，

- ほんとに調査観察データから因果推論を行なってもいいの？
- 研究者が独立変数を操作して行った実験研究でしか因果推論を行えないと断定してもいいの？

→調査観察データから因果関係を推論するためには，どのような条件が必要なの??

ヒュームの因果の定義と実験研究

因果関係の定義として，最もよく知られているのは，David Hume の 3 つの条件

1. 原因と結果が空間的・時間的に近接していること (spatial/temporal contingency)
2. 原因が結果よりも時間的に先行しており，継続して結果が起こること (temporal succession)
3. 同じ原因から必ず同じ結果が生じること (恒常的連接: constant conjunction)

である.

ヒュームの因果の定義と実験研究 (無意味スライド)

Figure: David Hume

(初めて図を入れられて嬉しい)

ヒュームの因果の定義と実験研究続き

以上の三つの条件の中には，原因と結果それぞれの生起に関わる”第3の要因 (= 共変量)”の議論がない．しかし，条件3を「共変量が同じである場合に」と条件付きの命題として捉えることで，共変量の議論も包摂されうる．

一方，実験可能な分野では...

1. 「独立変数の操作性」つまり，独立変数を操作することで従属変数が変化することを確認すること (= 実験研究)
2. 独立変数と従属変数の「時間的順序性」

の2点が因果関係を同定するために要件
→ これらを満たす研究で得られた関係こそ因果関係である!! という考え

ヒュームの因果の定義と実験研究続き

「時間的順序性」はヒュームの条件 2 に対応し,「独立変数の操作性」によってのみ恒常的連接を理解することができる,という暗黙の了解を考えれば,「独立変数の操作性」は条件 3 に対応している.

じゃあ「操作性のある研究のみが因果推論を可能にする」?
$\rightarrow$ んなわけない

批判 s

1. 社会科学など多くの学問分野では現実には「独立変数の操作性」という条件を満たすことは不可能．しかし (?) 実験研究よりも調査観察研究の方が結果の (生態学的な) 妥当性が高くなる可能性がある
2. 「独立変数の操作性」にこだわるなら，例えば物理学で確立されている多くの因果関係すらも単なる相関関係となってしまう (月の位置と潮の満ち引きは操作できるか？)
3. 自然の斉一性原理 (principle of the uniformity of nature)(統計学的には，交換可能性 (exchangeability)) の仮定が成り立つとして，一般的な因果関係を示すのは，帰納法の考え方によるもの．ヒューム自身は，帰納法では恒常的連接 (条件 3) は証明できないとしている．(因果関係は観察されたものだけに" たまたま" 成り立つものでは？)

批判 s をふまえて

3 番目の批判 (いわゆる Hume の懐疑論) に対して反論することは難しい．極論，Hume のいうとおり，どのような研究も因果関係を立証できないということになる...

以上の議論から，

- 「因果関係と相関関係は違う」だとか，「因果関係を立証できるのは実験研究だけである」といった低級な考えは誤り
- 因果関係を指し示す程度がやや高い研究と，やや低い研究がある
- その間には一定の線引きを行うことができるのではなく，様々な観点から得られた証拠全体から総合的に判断していくべき

という考えを持つべきだとわかる．

Hume の帰納法への批判　＝「得られたデータから母集団一般について議論ができるのか」

$\rightarrow$「因果関係の検証を行う際には，対象が本来の母集団からの代表性を確保するように十分留意しよう」

Hill のガイドライン

当該研究がどのような要件を持っていれば，より因果関係に近い関係を示したと言えるの？
疫学分野ではHill の因果関係判定のガイドラインが因果推論の指標として利用されている．

事象 A が事象 B の原因である，もしくは，変数 A の値の高低が変数 B に因果的な影響を与えていると判断するために，Hill は 9 つの基準を与えている．(次の項に列挙する)

Hillのガイドライン

Hillのガイドライン

1. 相関関係の強さ-AとBの間に強い相関関係がある．
2. 相関関係の一致性-相関関係は様々な状況で一致している
3. 相関関係の特異性-BとA以外に原因として想定される変数の相関，AとB以外の結果変数の相関が高くない
4. 時間的な先行性-AはBに時間的に先行する
5. 量・反応関係の成立-Aの値が大きくなると，Bも単調に変化する
6. 妥当性-AがBの原因という因果関係が知見に基づいて尤もらしい
7. 先行知見との整合性-先行研究や知見と齟齬がない
8. 実験による知見-動物実験等での実験研究による証拠がある
9. 他の知見との類似性-既に確立している別の因果関係と類似した関係・構造をしている．

Hill のガイドライン続き

Hill のガイドラインの基準をなるべく多く，かつ強く満たすほど，「得られた相関関係が因果関係である可能性が高い」と考えられる．

このガイドラインは，疫学のみならず，社会科学における因果関係の立証においても極めて有用である．

因果推論における母集団への代表性という観点

心理学や教育学，行動経済学や社会科学などで行われている人間を対象とした実験研究において，対象は均質であるという斉一性原理の仮定を置くことは難しい (= 代表値は母集団を表現していると判断できない！)...

調査研究では,

1. なるべく多くの共変量について，同時に調査時に測定し，中間変数でないことを吟味した上で調整を行う.
2. 解析の際にはなるべく仮定の少ない方法を利用する (さらに感度分析などの方法論も積極的に利用する．)
3. 母集団に対する代表性を担保する.

といった努力を行った結果，それでも因果効果が大きいとされた場合には，十分因果関係に近い関係を導いたと考えるべき.

References

- 高井, 星野, 野間 (2016)「調査観察データ解析の実際 欠測データの統計科学-医学と社会科学への応用-」
- 星野 (2009)「調査観察データの統計科学-因果推論・選択バイアス・データ融合-」

Thank you for your attention!!

Any Questions?

メモに使ってね!!