数理統計 補助資料
～統計的因果推論に必要な仮定・条件～

2023年度2学期:　月曜1限, 水曜3限

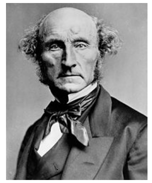
担当教員:　石垣　司

人間の推論・知識に必要な仮定

自然の斉一性：過去, 現在, 未来において同じ条件下では自然は同じ法則に従う, という仮定(公理)

- 統計的因果推論のみではなく, 人間の知性・推論全般に影響を与える仮定
- 自然の斉一性は(現在の人類の知識では)証明不可能
- 人間の科学的推論(より正確には帰納的推論)の多くは自然の斉一性という証明不可能な仮定の上で成立している

David Hume
1711-76



J. S. Mill
1806-73

統計的因果推論でも多くの仮定が必要

Rubinの因果効果を再考する

Rubinの因果効果と推定量

- y_1 と y_0 は処置を行った/行わなかった場合の結果変数
- $y_{i,1}$ と $y_{i,0}$ は主体 i で観測される標本(どちらかが反実仮想で観測不可)

因果効果(平均処置効果)の定義：$E[y_1 - y_0]$

推定量：$\hat{E}[y_1 - y_0] = \frac{1}{N_1}\sum_{i=1}^{N} z_i y_{i,1} - \frac{1}{N_0}\sum_{i=1}^{N}(1 - z_i) y_{i,0}$

定義は反実仮想を含むが, 推定量は観測データのみから計算可能

(復習)ランダム割り当ての場合は既学習

- $\hat{E}[y_1 - y_0]$ をRubinの因果効果の推定量とみなすためには, いくつかの重要な仮定が必要

#メモ　本日の話は面白くはないかもしれないが、統計的因果推論を行う上で重要な内容。

SUTVA

Stable Unit Treatment Value Assumption (SUTVA)

- Rubin流の反実仮想アプローチによるすべての統計的因果推論で必要となる仮定

  次の1と2の両方をまとめてSUTVAとよぶ

1. No Interference between units **の仮定**(Rubin 1980, Cox 1985)
   - どの主体 i の潜在的な結果変数の値も, i 以外の主体が処置を受けた結果で変化しない
2. No versions of treatments **の仮定**(Rubin 1980, Neyman 1935)
   - 各主体について異なる潜在的な結果変数をもたらす異なる処置の形態やバージョンは無い

SUTVA ~ No interference between unitsの仮定 #1

仮定を満たす例：(非感染性の)**頭痛薬の因果効果**

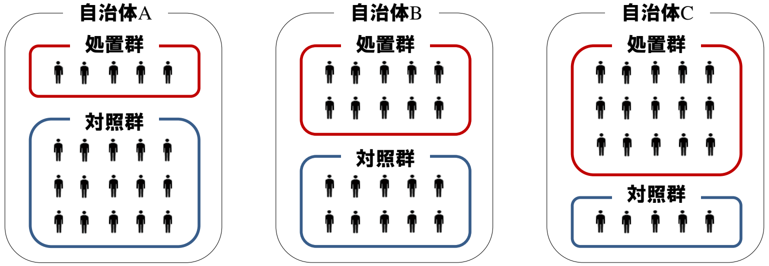
– 集団内のある主体の非感染性の頭痛の発生や痛みの程度は，他の主体が頭痛薬を飲んだ/飲まないという処置から影響を受けない

自治体Aでも自治体Cでも各主体の頭痛の起こりやすさなどは同様

仮定を満たさない例：感染症ワクチンの因果効果

– 集団内のワクチン接種率は，各主体の感染のしやすさに影響を与える

もしも，ワクチンに感染症の感染予防効果がある場合，自治体Aの方が自治体Cよりも各個人の感染率は高くなる



SUTVA ~ No interference between unitsの仮定 #2

仮定を満たす例：比較的大きな集団(多くの家庭が知り合い同士ではない集団という意味)**で商品が十分に市場に存在する場合のマーケティング施策と商品購買の因果効果の分析**

– ある家庭の商品の購買/非購買は，他の家庭へのマーケティング施策からは影響を受けない

仮定を満たさない例：市場に商品が十分に存在せず，かつ，マーケティング施策に効果がある場合

– マーケティング施策を受けた家庭で商品の購買が続いて品薄になることで，その他の家庭での購買意欲が変化する

– 波及効果・スピルオーバー効果などがあると仮定が満たされないケースあり

仮定を満たさない例：集団内の多くの家庭が知り合い同士でマーケティング施策に関する情報の共有がある場合

SUTVA ~ No versions of treatmentsの仮定 #1

例1：運動と健康の因果効果

– 仮定を満たす例：毎朝30分ほど時速7km/h程度のウォーキングを実施した

– 仮定を満たさない例：毎朝軽度のウォーキングを実施した

軽度の運動の意味が人によって異なる。処置のレベルが不明確

例2：ダイレクトメールと商品購買

– 仮定を満たす例：同じダイレクトメールを処置群へ送付した

– 仮定を満たさない例：デザインの異なるダイレクトメールを処置群へランダムに送付した

処置(ダイレクトメール)の効果とデザインの効果が区別できない

一方，"どのデザインが効果が高いか？"というデザインの効果を見るための実験計画であればOK

SUTVA ~ No versions of treatmentsの仮定 #2

No versions of treatments **の仮定から導出される性質**

– "処置群で処置を受けた場合の結果変数の期待値は，実際に観測される標本の期待値に等しい" & "対照群で処置を受けなかった場合の結果変数の期待値は，実際に観測される標本の期待値に等しい"

今まで学んできた統計学では，多くの場合に上の関係が成立する前提で話が展開された。それは標本がi.i.d.(独立同一分布)から発生すると仮定しているため

標本がi.i.d.であれば上記仮定は満たされる。よって，i.i.d.の仮定の方が上記の仮定よりも強い仮定

SUTVAの数理的記述 #1

変数の表記法

- $Z_i \in \{0,1\}$: **主体 i の割当変数**(確率変数)
- $\boldsymbol{Z} = [Z_1, \dots, Z_N]^T$: **全主体の割当変数ベクトル**
- $D_i \in \{0,1\}$: **主体 i が処置を受けたかどうかの変数**(確率変数)
- $\boldsymbol{D} = [D_1, \dots, D_N]^T$: **全主体の処置/非処置の確率変数ベクトル**
- $D_i(\boldsymbol{Z}) \in \{0,1\}$: **$\boldsymbol{Z}$ (つまり, 全主体 $i = 1, \dots, N$ についての割り当ての情報)が与えられた時の主体 i が処置を受けたかどうかの変数**(確率変数)
- $Y_i(\boldsymbol{Z}, \boldsymbol{D})$: **$\boldsymbol{Z}$ と $\boldsymbol{D}$ (つまり, 全主体 $i = 1, \dots, N$ についての割り当てと処置を受けたかどうかの情報)が与えられた時の主体 i の結果変数**(確率変数)

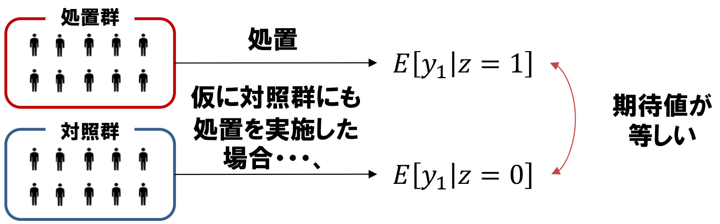
SUTVAの数理的記述 #2

SUTVA (Angrist, Imbens & Rubin 1996)

- (a) $Z_i = Z_i' \Rightarrow D_i(\boldsymbol{Z}) = D_i(\boldsymbol{Z}')$
  主体 i の割り当て(Z_i)が決まったら, 他の主体の割り当て(Z)が何であれ, 主体 i が処置をうけるかどうか($D_i(\boldsymbol{Z})$)は変わらない
- (b) $Z_i = Z_i'$ **かつ** $D_i = D_i' \Rightarrow Y_i(\boldsymbol{Z}, \boldsymbol{D}) = Y_i(\boldsymbol{Z}', \boldsymbol{D}')$
  主体 i の割り当て(Z_i)と処置をうけるかどうか(D_i)が決まったら, 他の主体の割り当てと処置が何であれ, 主体 i の結果変数($Y_i(\boldsymbol{Z}, \boldsymbol{D})$)は変わらない
- **仮定**(a),(b)**を満たすならば**, $Y_i(Z_i, D_i) = Y_i(\boldsymbol{Z}, \boldsymbol{D})$
  主体 i の割り当て(Z_i)と処置をうけるかどうか(D_i)が決まったら, 他の主体の割り当てと処置が何であれ, 主体 i の結果変数($Y_i(\boldsymbol{Z}, \boldsymbol{D})$)の値は主体 i に施された割り当てと処理(Z_i, D_i)のみで決まる

交換可能性 #1

処置群/対照群への割り当てを交換しても結果変数の期待値は等しい, という性質



- **交換可能性の数理的記述**

$$E[y_1|z=1] = E[y_1|z=0]$$
$$\Rightarrow E[y_1|z=1] = E[y_1|z=0] = E[y_1]$$

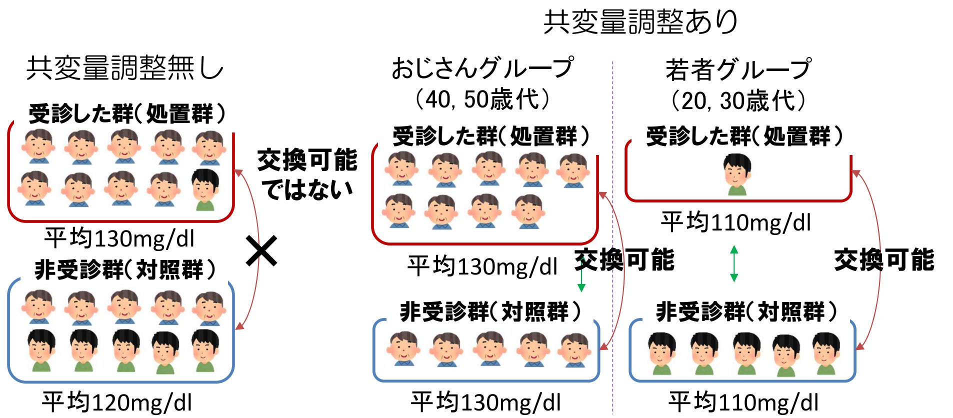
処置群に割り当てられた場合も対照群に割り当てられた場合も処置の効果の期待値は同じ ⇒ 全体での処置の効果の期待値も同じ

ランダム割り当ての場合は交換可能性が満たされる

交換可能性 #2

共変量があると明らかに交換可能性を満たさない

- **共変量調整とは, 両群間での交換可能性を確保する方法**
  例: 前々回授業の人間ドックとコレステロール値

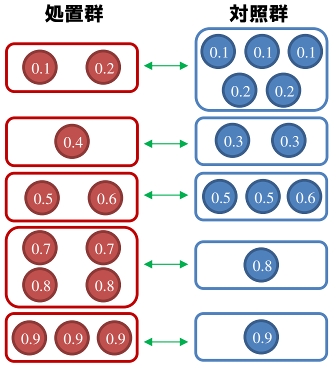


Positivity

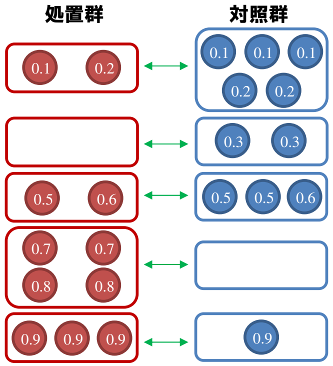
処置群/対照群に割り当てられた主体の数はゼロより大きい, という性質

– 例：共変量調整のための傾向スコアを用いた層別分析

Positivityを満たす層別分析



Positivityを満たさない層別分析



Rubinの因果効果の推定量の導出

因果効果の定義から平均処置効果の推定量の導出

– 交換可能性を満たすならば,

$$E[y_1 - y_0] = E[y_1] - E[y_0] = E[y_1|z=1] - E[y_0|z=0]$$

– SUTVAを満たすならば,

$$E[y_1|z=1] - E[y_0|z=0] = E\left[y_{i,1}|z=1\right] - E\left[y_{i,0}|z=0\right]$$

– よって, 交換可能性とSUTVAとPositivityを満たすとき,

$$\hat{E}[y_1 - y_0] = \frac{1}{N_1}\sum_{i=1}^{N} z_i y_{i,1} - \frac{1}{N_0}\sum_{i=1}^{N}(1 - z_i) y_{i,0}$$

#メモ　つまり、本日の授業内容は、何でもかんでもデータをとってきて因果推論を行うことができるとは限らない、ということを言っている。様々な仮定を満たしていると考えられる場合にのみ、正しい因果効果の推定が可能となる。卒論などで因果推論手法の利用を考えている人は十分に注意してほしい

満たすべき仮定・条件の整理

どんな場合でも必要

SUTVAとPositivityの仮定を満たす

ランダム割り当て可能

$\Rightarrow \hat{E}[y_1 - y_0] = \frac{1}{N_1}\sum_{i=1}^{N} z_i y_{i,1} - \frac{1}{N_0}\sum_{i=1}^{N}(1 - z_i) y_{i,0}$

ランダム割り当てではない ⇒ 交換可能性を満たす

$\Rightarrow \hat{E}[y_1 - y_0] = \frac{1}{N_1}\sum_{i=1}^{N} z_i y_{i,1} - \frac{1}{N_0}\sum_{i=1}^{N}(1 - z_i) y_{i,0}$

共変量がある ⇒ 共変量調整で平均での独立性を満たす

$\Rightarrow \hat{E}[y_1 - y_0] = \sum_{k=1}^{K} \frac{N_k}{N} \hat{y}^{(k)}$ （マッチングや層別分析）

パネルデータがあり介入前後の因果効果を知りたい ⇒ 平行トレンドの仮定を満たす ⇒ 共通ショックの仮定を満たす

$\Rightarrow \widehat{DD} = \frac{1}{N_1}\sum_{i=1}^{N} z_i (y_{i,t=1} - y_{i,t=0}) - \frac{1}{N_0}\sum_{i=1}^{N}(1 - z_i)(y_{i,t=1} - y_{i,t=0})$

データ分析結果の外的妥当性

外的妥当性

– 分析結果を実験対象以外の対象に応用できる度合い

例：（ノーベル経済学賞受賞者らの研究）ケニアの小学校でのランダム化比較試験の事例から得られた知見が日本の小学校に当てはまるか？

科学分野と得られる知見の外的妥当性の強さ

– 高：物理法則

同じ条件で実験をするとほぼ同じ結果が地球上のどこででも得られる

– 中：生物学的メカニズム

物理的・化学的法則が作用するメカニズムは全人類で共通。一方, 人種・性別・所属・属性・遺伝的性質などでその効果が異なる

– 低：社会的傾向、心理的傾向

所属集団などで大きく異なる

経済学・経営学などの社会科学はここ