北大SPH・統計的因果推論と臨床疫学⑥

構造ネストモデルとg推定

北海道大学　医学統計学
横田　勲

1

今回の内容

2

- 構造ネストモデル
- g推定

2

$L = l$である人での平均因果効果

3

- $E[Y^{a=1}|L=l] - E[Y^{a=0}|L=l]$
  - 平均因果効果は集団全体での効果だったけど そこは目をつむろう
- A, Lの交互作用をすべて含めた 周辺構造モデルをおけば、 通常のアウトカム回帰でパラメータ推定
  - Stabilized weightを用いた場合
    - 重みが全員とも1になる

3

条件付き交換可能性

4

- 予後因子Lが同じ値を持つ集団（層内） では交換可能性が成立
  - $\Pr[Y^{a=1}|A=1, L=1] = \Pr[Y^{a=1}|A=0, L=1]$
    $\Pr[Y^{a=0}|A=0, L=1] = \Pr[Y^{a=0}|A=1, L=1]$
  - $\Pr[Y^{a=1}|A=1, L=0] = \Pr[Y^{a=1}|A=0, L=0]$
    $\Pr[Y^{a=0}|A=0, L=0] = \Pr[Y^{a=0}|A=1, L=0]$
  - $Y^a \perp\!\!\!\perp A|L$ for all a
- No unmeasured confounding
  - 残差交絡 residual confounding がない

4

条件付き交換可能性が成立すれば

5

- 反事実アウトカムは曝露確率を予測しない

$$\Pr[A=1|Y^{a=0}, L] = \Pr[A=1|L]$$

- ロジスティックモデルをおけば、
  $\operatorname{logit} \Pr[A=1|Y^{a=0}, L] = \alpha_0 + \alpha_1 Y^{a=0} + \alpha_2 L$
  - 仮に$Y^{a=0}$を知っていて、モデルが正しければ α_1は期待的にどうなる？

5

$E[\alpha_1] = 0$

6

- 逆転の発想
- $E[\alpha_1] = 0$となるような$Y^{a=0}$が 正しいのではないか

6

構造ネスト平均モデル

7

- 条件付き因果効果に対するモデル
  - $E[Y^a - Y^{a=0}|L] = \beta_1 a$
    - 効果の修飾がない場合の例
  - $E[Y^a - Y^{a=0}|L] = \beta_1 a + \beta_2 aL$
    - 効果の修飾がある場合の例
- 交換可能性の下、
  $E[Y^a - Y^{a=0}|A = a, L] = \beta_1 a + \beta_2 aL$
- 標準的なパラメータ推定法として<u>g推定</u>

7

セミパラメトリックな性質

8

- 潜在アウトカムごとのモデルが不要
  - g-formulaの場合
    $E[Y^a|L] = \beta_0 + \beta_1 a + \beta_2 aL + \beta_3 L$
  - 構造ネスト平均モデル
    $E[Y^a - Y^{a=0}|L] = \beta_1 a + \beta_2 aL$
    - あくまで、<u>因果効果部分だけをモデル化</u>
- 解析時におく仮定を少なくできる
  - モデルの誤特定に頑健

8

乗法構造ネスト平均モデル

9

- 加法モデルだけではなく、
  二値や再発数のような場合

$$\log\left(\frac{E[Y^a|A = a, L]}{E[Y^{a=0}|A = a, L]}\right) = \beta_1 a + \beta_2 aL$$

- 構造ネストロジスティックモデル
  - ただし、繰り返し治療の状況へ一般化が困難

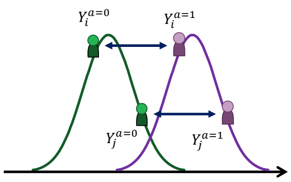
$$\text{logit} \Pr[Y^a|A = a, L] - \text{logit} \Pr[Y^{a=0}|A = a, L] = \beta_1 a + \beta_2 aL$$

9

Rank preservation

10

- 個人間で効果の大きさは共通
  - Sharp causal null hypothesisと対応
- Rank-preserving 構造モデル
  $Y_i^a - Y_i^{a=0} = \psi_1 a + \psi_2 aL_i \quad \text{for all } i$



$Y^{a=0}$	$Y^{a=1}$
3	6
4	7
8	11
5	8
0	3
9	12

10

Rank preservationの仮定をおく

11

- 非現実的な仮定
  - 平均因果効果の定義にも不要な仮定
- 周辺構造モデル、構造平均モデルとも
  rank preservationの仮定は不要
- g推定はrank preservingの有無によらず
  同一の手順
  - rank preservingモデルに対しわかりやすい

11

g推定①

12

- $E[Y^a - Y^{a=0}|L] = \beta_1 a$
- 加法rank-preservingモデルを仮定
  $Y_i^a - Y_i^{a=0} = \psi_1 a \Leftrightarrow Y_i^{a=0} = Y_i^a - \psi_1 a$
- Consistencyの仮定より
  $Y^{a=0} = Y - \psi_1 A$
  - $Y_i^{a=0}|a = 0 = Y_i,\ Y_i^{a=1}|a = 1 = Y_i$
  - 左辺を$H(\psi_1)$と表記

12

g推定② 13

- ψ_1の候補$\psi_1^\dagger$を適当に設定

$\text{logit Pr}\left[A = 1 | H(\psi_1^\dagger), L\right] = \alpha_0 + \alpha_1 H(\psi_1^\dagger) + \alpha_2 L$

- $\alpha_1 = 0$となる$H(\psi_1^\dagger)$が$Y^{a=0}$
  - そんな$\psi_1^\dagger$をψ_1の推定値$\widehat{\psi_1}$
  - $\psi_1^\dagger$を適当な範囲で動かして$\widehat{\psi_1}$を探索
- 信頼区間の構成
  - $\alpha_1 = 0$に対する検定で有意にならない$\widehat{\psi_1}$

13

Non-rank preserving modelでも 14

- ψ_1は構造ネストモデルのβ_1を一致推定
  - 構造ネストモデルにおける平均が正しく特定
- Rank preservingの仮定は不要
  - 個人間で共通する治療効果
  - $Y^{a=0} \perp\!\!\!\perp A|L$となる$Y^{a=0}$を$H(\psi_1^\dagger)$で探索したにすぎない
    - $\perp\!\!\!\perp$ を$\alpha_1 = 0$として表した

14

感度解析に応用 15

- α_1は条件付き交換可能性の程度を表現
- 通常、未測定の交絡はあるはず
  - α_1がある値である場合の因果効果を調べる
  - 様々なシナリオでのα_1の範囲を調べる

Robins, Rotnitzky, and Scharfstein (1999) book chapter.

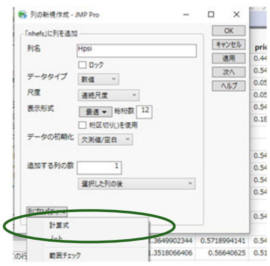
15

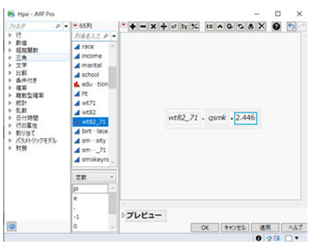
練習 16

- NHEFSデータでg推定を行ってみよう
  - 体重変化量の平均因果差
    - 候補は2.446, 3.446, 4.446
  - 曝露は喫煙習慣
  - 交絡変数
    - 性別、人種、年齢(2次項まで)、教育歴、喫煙本数(2次項まで)、喫煙期間(2次項まで)、運動習慣、活動性、体重(2次項まで)
- データは以下より入手
  - https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/

16

$H(\psi_1^\dagger)$を作成 17





17

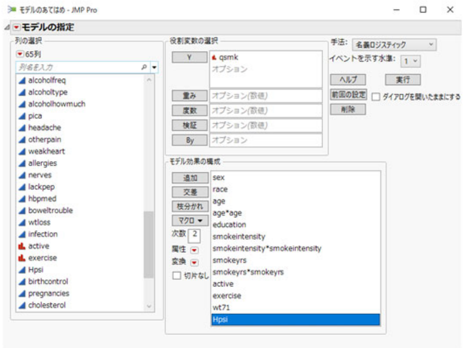
$H(\psi_1^\dagger)$を確認してみよう 18

seqn	qsmk	wt82_71	Hpsi
262	0	-4.0829919	-4.0829919
266	0	0.2270081	0.2270081
419	0	-2.71784531	-2.71784531
420	0	9.86118683	9.86118683
428	1	15.87225712	13.42625712
431	0	-1.81900392	-1.81900392
434	0	0.56215469	0.56215469
443	0	0.11013065	0.11013065
446	1	-18.94095375	-21.38695375
455	0	5.32963261	5.32963261
457	0	0.4507029	0.4507029
596	1	8.27681735	5.83081735
603	0	6.80691978	6.80691978
604	0	3.85963261	3.85963261
605	0	0.90914868	0.90914868
616	0	-13.94473687	-13.94473687
618	1	9.98040972	7.53440972
619	0	6.2407029	6.2407029
620	0	-1.69862242	-1.69862242
804	0	-5.44132114	-5.44132114
806	1	10.42983747	7.98383747
813	0	-8.16454612	-8.16454612

18

曝露に対するロジスティック回帰

19



$H(\psi_1^\dagger)$部分のカイ二乗値

20

- 0から大きく離れている

パラメータ推定値

項	推定値	標準誤差	カイ2乗	p値(Prob>ChiSq)
切片	-2.3100196	0.4879048	22.42	<.0001*
sex	-0.4649867	0.1491273	9.72	0.0018*
race	-0.8399326	0.2098389	16.02	<.0001*
age	0.05132957	0.0106881	23.06	<.0001*
(age-43.6596)*(age-43.6596)	-0.0006898	0.0005391	1.64	0.2007
education[1]	-0.106111	0.1370728	0.60	0.4389
education[2]	-0.1549343	0.1277212	1.47	0.2251
education[3]	-0.036988	0.1028653	0.13	0.7192
education[4]	-0.0726983	0.1890661	0.15	0.7006
smokeintensity	-0.0346566	0.0059669	33.73	<.0001*
(smokeintensity-20.5255)*(smokeintensity-20.5255)	0.00106796	0.0002882	13.74	0.0002*
smokeyrs	-0.0323062	0.0105226	9.43	0.0021*
(smokeyrs-24.5881)*(smokeyrs-24.5881)	0.00087482	0.0004638	3.56	0.0593
active[0]	-0.0779601	0.0961843	0.66	0.4176
active[1]	-0.0313123	0.0933987	0.11	0.7374
exercise[0]	-0.247063	0.1137551	4.72	0.0299*
exercise[1]	0.10510462	0.0871802	1.45	0.2280
wt71	0.00817104	0.0044833	3.32	0.0684
Hpsi	0.01854688	0.0081571	5.17	0.0230*

推定値は次の対数オッズに対するものです： 1/0

因果差の候補を変えて探索

21

パラメータ推定値

項	因果差 2.446 カイ2乗	因果差 2.446 p値(Prob>ChiSq)	因果差 3.446 カイ2乗	因果差 3.446 p値(Prob>ChiSq)	因果差 4.446 カイ2乗	因果差 4.446 p値(Prob>ChiSq)
切片	22.42	<.0001*	16.62	<.0001*	11.38	0.0007*
sex	9.72	0.0018*	11.00	0.0009*	12.57	0.0004*
race	16.02	<.0001*	15.83	<.0001*	15.51	<.0001*
age	23.06	<.0001*	20.48	<.0001*	17.83	<.0001*
(age-43.6596)*(age-43.6596)	1.64	0.2007	2.26	0.1325	2.96	0.0856
education[1]	0.60	0.4389	0.72	0.3948	0.87	0.3498
education[2]	1.47	0.2251	1.38	0.2397	1.31	0.2516
education[3]	0.13	0.7192	0.12	0.7306	0.11	0.7429
education[4]	0.15	0.7006	0.09	0.7689	0.03	0.8527
smokeintensity	33.73	<.0001*	33.02	<.0001*	32.13	<.0001*
(smokeintensity-20.5255)*(smokeintensity-20.52	13.74	0.0002*	13.52	0.0002*	13.16	0.0003*
smokeyrs	9.43	0.0021*	9.00	0.0027*	8.51	0.0035*
(smokeyrs-24.5881)*(smokeyrs-24.5881)	3.56	0.0593	3.30	0.0694	2.98	0.0844
active[0]	0.66	0.4176	0.51	0.4756	0.37	0.5423
active[1]	0.11	0.7374	0.18	0.6733	0.26	0.6110
exercise[0]	4.72	0.0299*	4.86	0.0275*	4.92	0.0265*
exercise[1]	1.45	0.2280	1.49	0.2221	1.52	0.2170
wt71	3.32	0.0684	[illegible]	[illegible]	0.76	0.3831
Hpsi	5.17	0.0230*	0.00	0.9961	5.22	0.0224*

推定値は次の対数オッズに対するものです： 1/0

1時点のみの曝露(治療)の場合

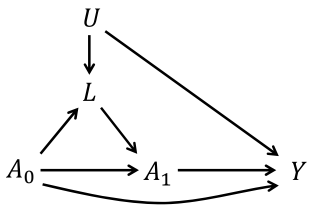
22

- g推定は使わずともIPW法や標準化で十分
- 繰り返し治療の場合
  - 過去の治療、共変量の履歴を考慮したモデル
    - "nested"モデルと呼ばれる所以

時間依存性交絡

23

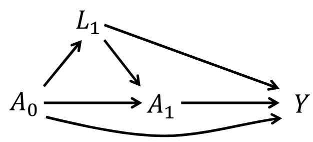
- シンプルな状況として以下の因果DAG
  - A_0とA_1の効果を知りたい
    - Lを条件付けない場合
    - Lを条件付ける場合



治療が2時点

24

- 以下の因果DAG



- 2つの式をもつ構造ネスト平均モデル

$$E\left[Y^{a_0,a_1=0} - Y^{a_0=0,a_1=0}\right] = \beta_0 a_0$$

$$E\left[Y^{a_0,a_1} - Y^{a_0,a_1=0} \middle| L_1^{a_0} = l_1, A_0 = a_0\right]$$

$$= a_1(\beta_{11} + \beta_{12} l_1 + \beta_{13} a_0 + \beta_{14} a_0 l_1)$$

“nested”の意味

- 時点0にて受けた治療の効果

$E[Y^{a_0,a_1=0} - Y^{a_0=0,a_1=0}] = \beta_0 a_0$

  - 以降の治療内容による影響を受けない

- 時点1にて受けた治療の効果

$E\left[Y^{a_0,a_1} - Y^{a_0,a_1=0} \middle| L_1^{a_0} = l_1, A_0 = a_0\right]$

$= a_1(\beta_{11} + \beta_{12} l_1 + \beta_{13} a_0 + \beta_{14} a_0 l_1)$

25

a_1の効果

$E[Y^{a_0,a_1=0} - Y^{a_0=0,a_1=0}] = \beta_0 a_0$

$E\left[Y^{a_0,a_1} - Y^{a_0,a_1=0} \middle| L_1^{a_0} = l_1, A_0 = a_0\right]$

$= a_1(\beta_{11} + \beta_{12} l_1 + \beta_{13} a_0 + \beta_{14} a_0 l_1)$

- $a_0 = 0$である場合のa_1の効果
  - $L_1^{a_0=0} = 0$である個人：β_{11}
  - $L_1^{a_0=0} = 1$である個人：$\beta_{11} + \beta_{12}$
- $a_0 = 1$である場合のa_1の効果
  - $L_1^{a_0=1} = 0$である個人：$\beta_{11} + \beta_{13}$
  - $L_1^{a_0=1} = 1$である個人：$\beta_{11} + \beta_{13} + \beta_{14}$

26

各個人に対するRPSNM

- $Y_i^{a_0,a_1=0} = Y_i^{a_0=0,a_1=0} + \psi_0 a_0$
- $Y_i^{a_0,a_1} = Y_i^{a_0,a_1=0} + \psi_{11} a_1 + \psi_{12} a_1 L_{1,i}^{a_0}$
  $+\psi_{13} a_1 a_0 + \psi_{14} a_1 a_0 L_{1,i}^{a_0}$
- 2式をあわせて

$Y_i^{a_0=0,a_1=0}$

$= Y^{a_0,a_1} - \psi_0 a_0$

$-\left(\psi_{11} a_1 + \psi_{12} a_1 L_{1,i}^{a_0} + \psi_{13} a_1 a_0 + \psi_{14} a_1 a_0 L_{1,i}^{a_0}\right)$

27

$H(\psi^\dagger)$を作成

- $H(\psi^\dagger)$

$= Y^{a_0=0,a_1=0}$

$= Y - \psi_0 A_0 - (\psi_{11} A_1 + \psi_{12} A_1 L_1 + \psi_{13} A_1 A_0 + \psi_{14} A_1 A_0 L_1)$

28

曝露確率のモデル

- 逐次交換可能性
  - 交換可能性を繰り返し曝露に拡張したもの
  - $Y^{a_0=0,a_1=0} \perp\!\!\!\perp A_1 | L_1, A_0$
  - $Y^{a_0=0,a_1=0} \perp\!\!\!\perp A_0$
- 曝露確率のロジスティックモデル
  - 赤字の係数が0となる$\psi^\dagger$を探す

$\mathrm{logit}\ \Pr\left[A_0 = 1 | H(\psi^\dagger)\right] = \alpha_{00} + \alpha_{01} H(\psi^\dagger)$

$\mathrm{logit}\ \Pr\left[A_1 = 1 | H(\psi^\dagger), L_1, A_0\right]$

$= \alpha_{10} + (\alpha_{11} + \alpha_{12} A_0 + \alpha_{13} L_1 + \alpha_{14} L_1 A_0) H(\psi^\dagger) + \alpha_{15} L_1 + \alpha_{16} A_0$

29

まとめ

- 構造ネストモデル
  - 条件付きの効果に対するモデリング
- g推定
  - 条件付き交換可能を関連なしと読み替えて妥当な推定値を探索
  - 繰り返し曝露の状況で有力
    - 時間依存性交絡への対処

30