

【日本製薬工業協会シンポジウム】 ICH-E9(R1) の理解に役立つ因果推論シンポジウム

セクション 2: 因果推論 IPW/Standardization



2021年3月19日

医薬品評価委員会 データサイエンス部会 KT1

因果推論サブチーム

○アステラス製薬株式会社 寺田 亘

ノバルティス ファーマ株式会社 手島 梨恵

本セッションについて

➤ 本演題は以下を参考に作成しています。

■ Hernán MA, Robins JM (2020). Causal Inference: What If. Boca Raton: Chapman & Hall/CRC. Chapter 2, 12, 13.

□ <https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/>
(last access date 23-MAR-2021)

- 観察研究で治療効果 (average causal effect) を求めたい
 - 交絡が存在する場合は, 交絡を調整して治療効果を推定する必要がある
 - ここではIPWを紹介

- まずはシンプルな状況で説明 (ノンパラ)
 - 交絡に対処する方法として疑似集団を作成

- モデルへの拡張 (モデル)
 - Marginal structural modelを紹介

- SASを用いた解析事例の紹介

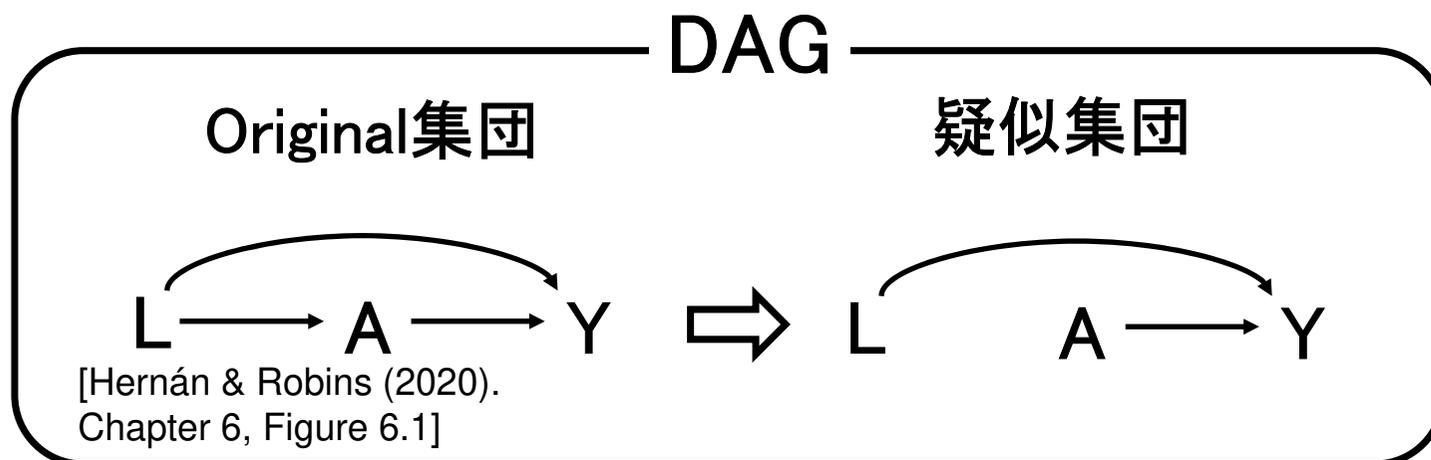
IPW

ノンパラメトリックのIPW

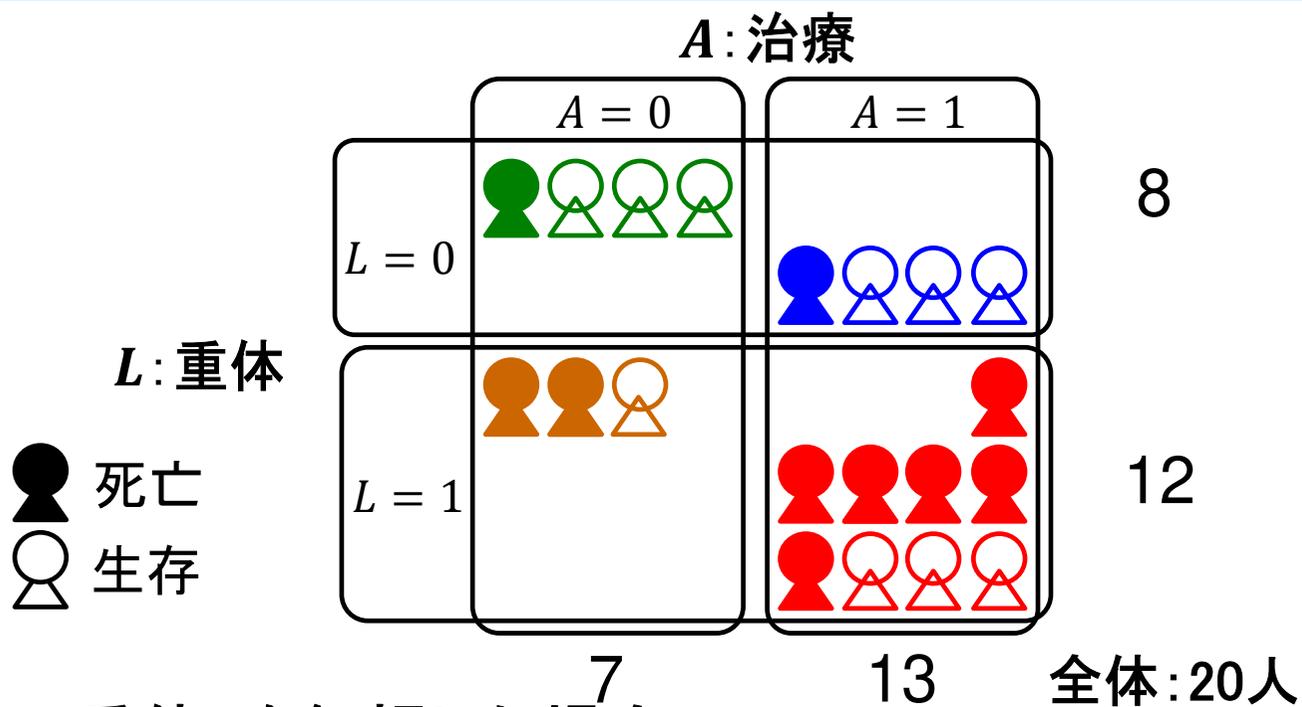
- 観察研究では交絡が起こっていると考えるのが妥当
 - そのまま推定するとバイアスを含む推定値. 何らかの方法で調整したい.
ここでは2つの補正方法を紹介
 - Inverse Probability Weighting (IPW) ⇔ まずはこっち
 - *Standardization*

やりたいこと

L : 交絡因子 (1: 重体, 0: 非重体)
 A : 治療有無 (1: 治療, 0: 治療なし)
 Y : 生死 (1: 死亡, 0: 生存)



- 交絡因子 L と治療 A の間に関連がないような疑似集団を作成し、治療効果を推定する



➤ 重体 L を無視した場合

- 治療群: 死亡確率 $7/13$
- 非治療群: 死亡確率 $3/7$

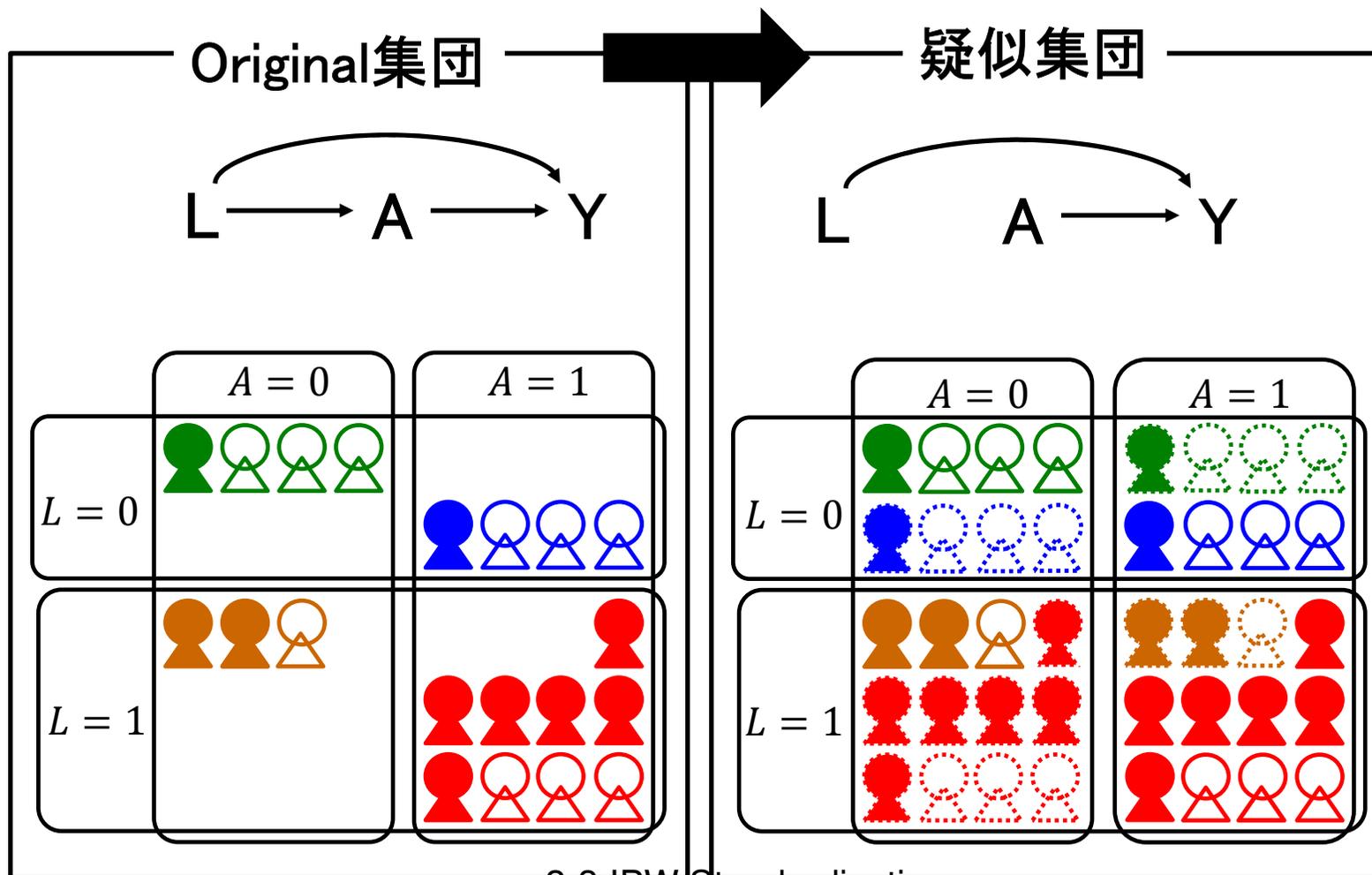
条件付き交換可能性

- $L = l$ の患者において、治療 ($A = 1$) の患者の死亡率と、仮に無治療 ($A = 0$) の患者が治療 ($A = 1$) となった場合の死亡率が同じ

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

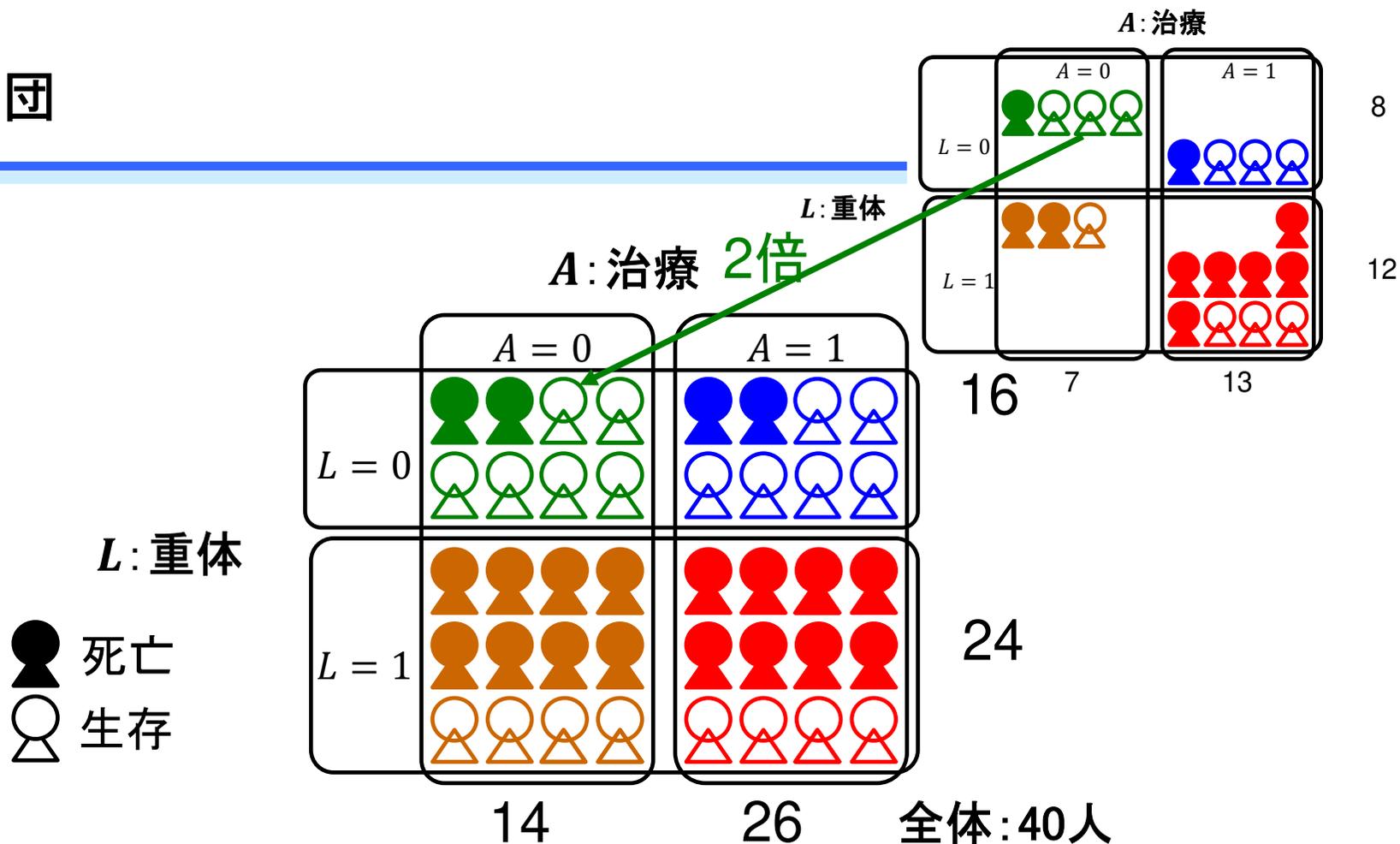
条件付き交換可能性: $Y^a \perp\!\!\!\perp A | L$ for all a

Original集団から疑似集団



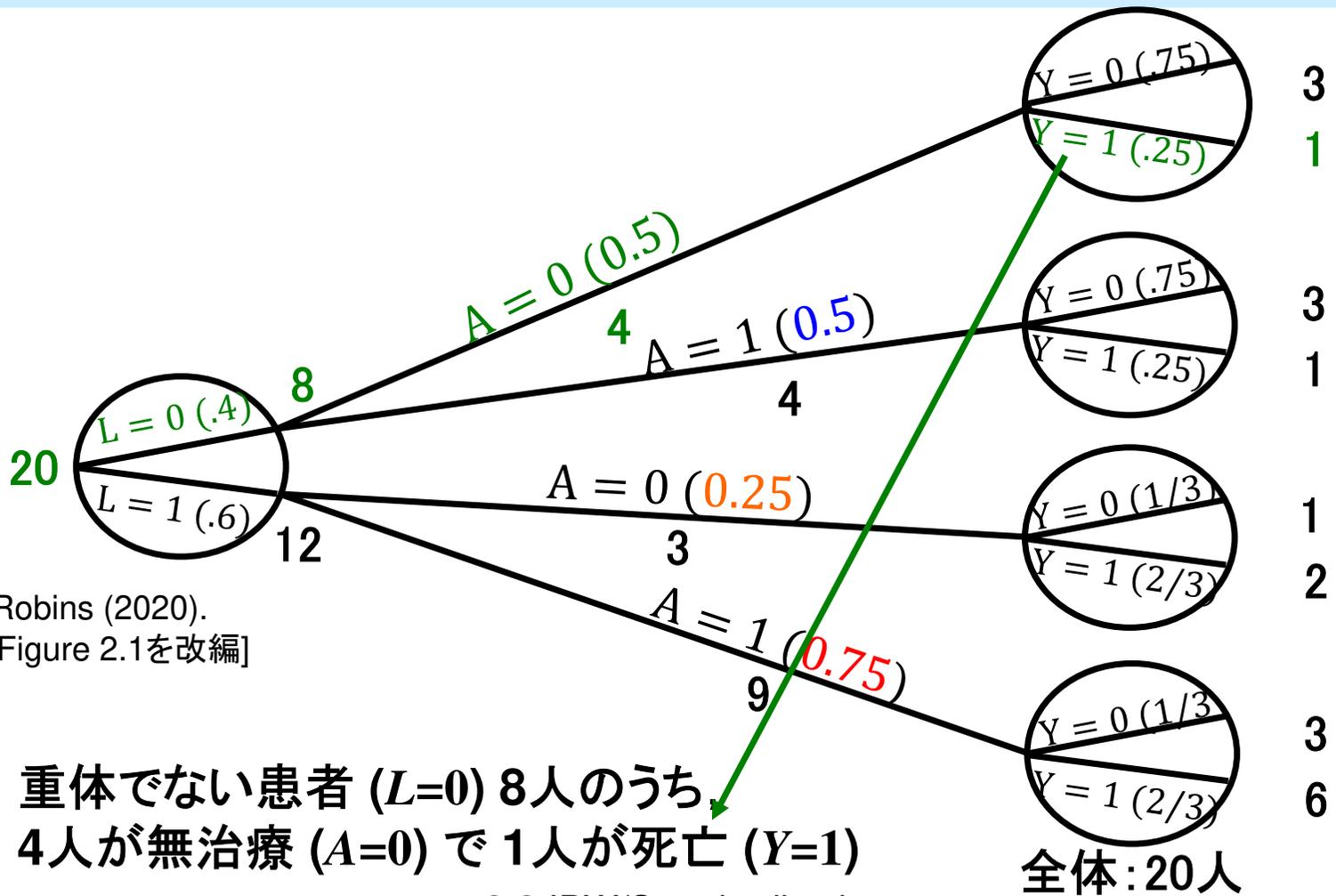
2-3 IPW/Standardization

疑似集団



- Original集団に条件付き交換可能性の成立を仮定
 - 重体Lが治療Aに影響を与えない

Original集団をTreeで表示



[Hernán & Robins (2020).
Chapter 2, Figure 2.1を改編]

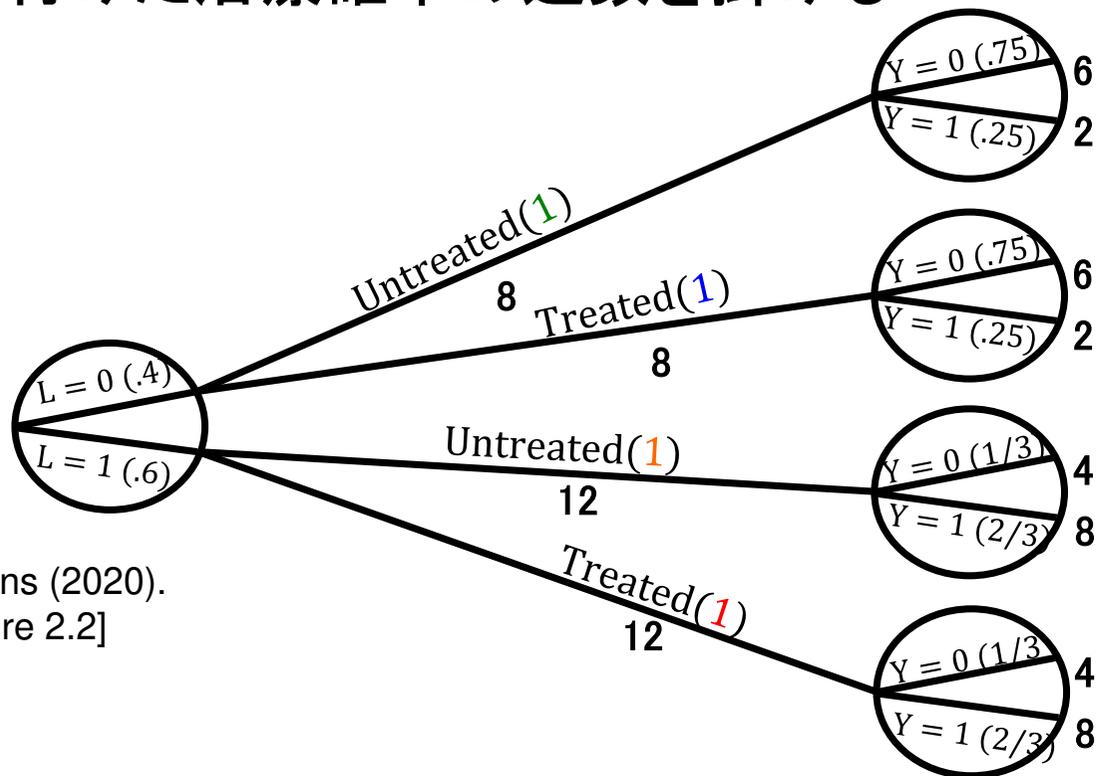
重体でない患者 ($L=0$) 8人のうち
4人が無治療 ($A=0$) で1人が死亡 ($Y=1$)

2-3 IPW/Standardization

全体: 20人

疑似集団をTreeで表示

➤ L で条件付けた治療確率の逆数を掛ける



$$W^A = 1/f(A|L)$$

$$\begin{aligned} 1/.5 &= 2 \\ 1/.5 &= 2 \end{aligned}$$

$$1/\Pr[A = 0|L = 0]$$

$$1/.5 = 2$$

$$1/.5 = 2$$

$$1/\Pr[A = 1|L = 0]$$

$$1/.25 = 4$$

$$1/.25 = 4$$

$$1/\Pr[A = 0|L = 1]$$

$$1/.75 = 1.33$$

$$1/.75 = 1.33$$

$$1/\Pr[A = 1|L = 1]$$

[Hernán & Robins (2020).
Chapter 2, Figure 2.2]

20人全員が「治療あり」「治療なし」両方に存在
 L によらず、割付確率が一定

全体: 40人

2-3 IPW/Standardization

疑似集団の作成方法: 重み W^A を掛ける

- 重み: $W^A = 1/f(A|L)$ を掛ける
 - 交絡因子が与えられた下での, 観察された治療確率
 - 治療 $A = a$ の平均: $E \left[\frac{I(A=a)Y}{f(A|L)} \right]$
 - 推定値が不安定になりがち ($f(A|L)$ が小さい場合, 重みが大きくなる)

➤ $SW^A = f(A)/f(A|L)$

- 重みは W^A に限らない (ポイントは, 交絡を調整した疑似集団を作成すること)
- 分子に何かの重みを掛ける (L によらない A の関数). 例えば, 各群の治療確率
- 重みの幅が狭まり, 推定値は W^A より安定

⇒ このような重みを **Stabilized weight** と呼ぶ

IPW

モデルへ拡張

➤ 今までの話

- 比較的シンプルな状況を想定

- 治療: 2群

- 交絡要因: 1つで2水準

- ノンパラメトリックで対応可能

➤ 現実にはもう少し複雑な状況が多い

- 治療群: 多群, 連続値

- 交絡要因: 複数, 連続量の要因, 交互作用など

- ノンパラメトリックを拡張し, モデルで対応

- Marginal structural model (MSM) を用いる (詳細は次ページ以降)

- モデル: $E[Y^a] = \beta_0 + \beta_1 a$ を考える
 - $E[Y^{a=0}] = \beta_0$
 - $E[Y^{a=1}] = \beta_0 + \beta_1$
 - $E[Y^{a=1}] - E[Y^{a=0}] = \beta_1$: average causal effect
- Marginal structural model (MSM) と呼ぶ
 - “Marginal”: (潜在アウトカムの)周辺分布に対するモデル
 - “Structural”: 潜在アウトカムに対するモデル

➤ IPWによる重み付き最小二乗法を用いて推定

- IPWを用いて疑似集団を作成
- 疑似集団に対して $E[Y|A] = \theta_0 + \theta_1 A$ とモデルをおく

⇒ $\hat{\theta}_0$ および $\hat{\theta}_1$ は $\sum_i \hat{W}_i [Y_i - (\theta_0 + \theta_1 A_i)]^2$ を最小にする θ_0, θ_1 である

➤ 重み付き最小二乗推定量 $\hat{\theta}_1$ により, MSMのパラメータ β_1 を不偏に推定できる

IPW

SASを用いた解析事例の解析

MSM (重み: $SW^A = f(A)/f(A|L)$) の推定

➤ 重みの推定 (分母) : ロジスティック回帰より

```
proc logistic data=データセット名 descending;  
  ods exclude ClassLevelInfo ModelAnova Association FitStatistics GlobalTests;  
  class exercise active education;  
  model qsmk = sex race age age*age education smokeintensity  
             smokeintensity*smokeintensity smokeyrs smokeyrs*smokeyrs  
             exercise active wt71 wt71*wt71;  
  output out=データセット名 p=変数名 (禁煙確率の推定値);  
run;
```

➤ 重みの推定 (分子) : ロジスティック回帰より

```
proc logistic data=データセット名 descending;  
  ods exclude ClassLevelInfo ModelAnova Association FitStatistics GlobalTests;  
  class exercise active education;  
  model qsmk = ;  
  output out=データセット名 p=変数名 (禁煙確率の推定値);  
run;
```

MSM (重み: $SW^A = f(A)/f(A|L)$) の推定

➤ 重み: $SW^A = \text{分子} / \text{分母}$ を算出

```
data データセット名;  
  merge データセット名(重み:分母) データセット名(重み:分子);  
  by seqn;  
  if qsmk=1 then 重み= (重み:分子) / (重み:分母);  
  else if qsmk=0 then 重み= (1-重み:分子) / (1-重み:分母);  
run;
```

➤ 重み付き最小二乗法で推定

```
proc genmod data= データセット名;  
  class seqn;  
  weight 重み;  
  model wt82_71= qsmk;  
  repeated subject=seqn / type=ind;  
run;
```

パラメータ	推定値	標準誤差	95% 信頼限界	
Intercept	1.7800	0.2247	1.3395	2.2204
qsmk	3.4405	0.5255	2.4106	4.4705

CAUSALTRTプロシジャを用いた推定

```
proc causaltrt data=データセット名;  
  class exercise active education qsmk/desc;  
  psmodel qsmk= sex race age age*age education smokeintensity  
    smokeintensity*smokeintensity smokeyrs smokeyrs*smokeyrs  
    exercise active wt71 wt71*wt71 ;  
  model wt82_71;  
  bootstrap seed=1776;  
run;
```

Parameter	Treatment Level	Estimate	Robust Std Err	Bootstrap Std Err	Wald 95% Confidence Limits		Bootstrap Bias Corrected 95% Confidence Limits	
POM	1	5.2205	0.4449	0.4443	4.3486	6.0925	4.3928	6.1297
POM	0	1.7800	0.2181	0.2181	1.3525	2.2075	1.3381	2.2103
ATE		3.4405	0.4871	0.4903	2.4859	4.3952	2.5535	4.4594

まとめ (IPWパート)

- 観察研究で交絡に対処する方法の1つとしてIPWを紹介
- ノンパラメトリックな状況で、治療確率の逆数を用いた重みから疑似集団を作成することにより交絡に対処する方法を紹介
- 潜在アウトカムにモデルを仮定するMSMを紹介

- Standardizationとは
- Standardizationによる治療効果の推定
- SASを用いたStandardizationの解析
- IPWとStandardizationの比較

Standardization

Standardizationとは

Standardization とは

一般に治療効果の推定は交絡因子の調整をせずに行うとバイアスが生じる。

Standardizationとは、以下のように交絡因子を調整することで治療効果を推定する手法である：

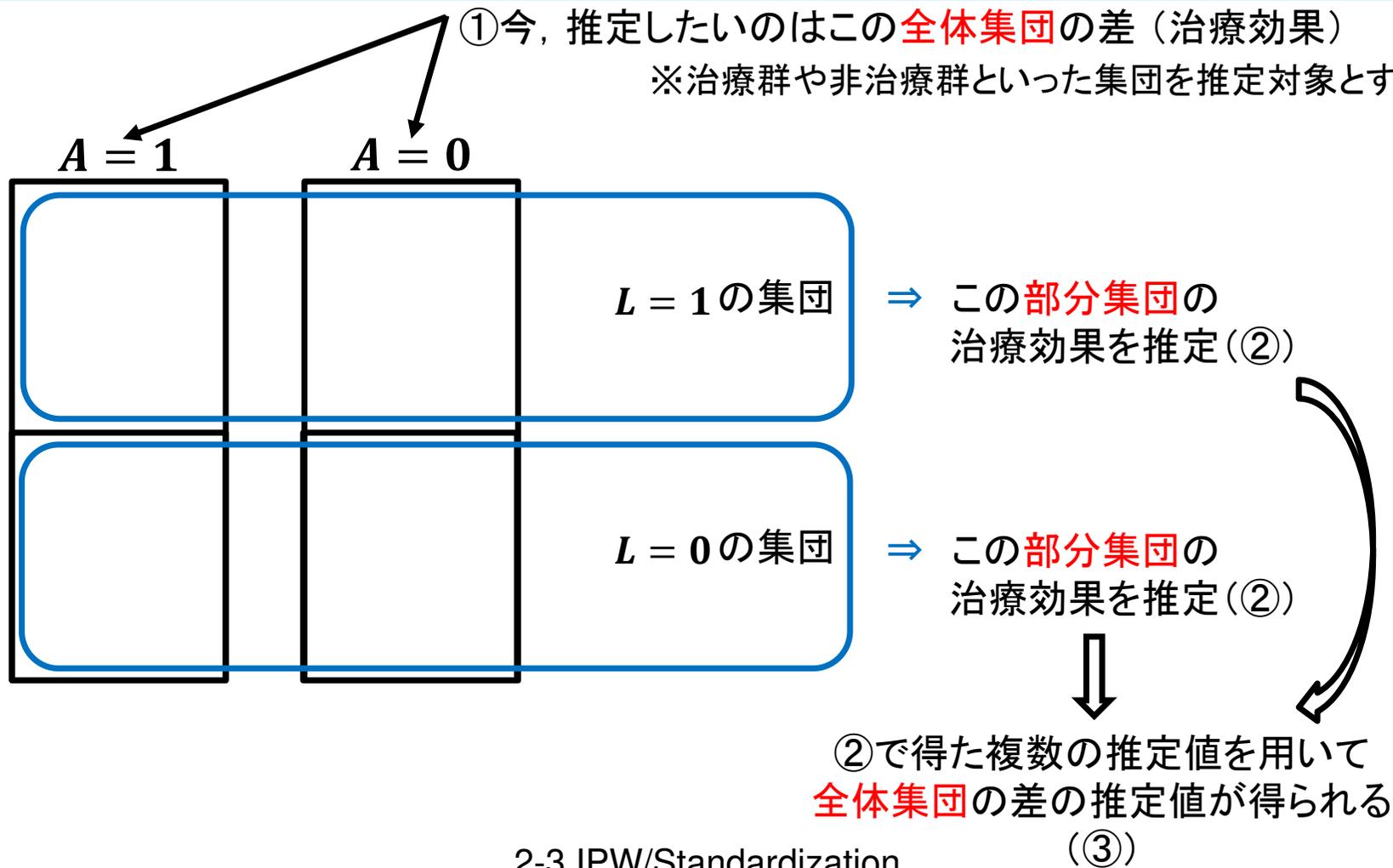
まず交絡因子を条件付けて、その交絡因子の影響を除去した状況における推定量を構成する。

その後、条件付けた交絡因子ごとに得られた推定量を統合することで全体の治療効果を推定する。

Standardization

Standardizationによる治療効果の推定

シンプルな状況を考える



シンプルな状況を考える

なぜ“Standardization”なのかを, 2値の交絡因子が1個で, アウトカムも2値といったシンプルな例を用いて紹介する.

➤ $N(= N_1 + N_0)$ 人を交絡因子の値で層に分ける

$L = 1$ の層	$Y = 1$	$Y = 0$	
$A = 1$	x_1	$n_1 - x_1$	n_1
$A = 0$	y_1	$m_1 - y_1$	m_1
			N_1

$L = 0$ の層	$Y = 1$	$Y = 0$	
$A = 1$	x_0	$n_0 - x_0$	n_0
$A = 0$	y_0	$m_0 - y_0$	m_0
			N_0

シンプルな状況を考える

知りたいこと: N 人全員が治療を受けたときのアウトカムと,
 N 人全員が治療を受けなかったときのアウトカムの差

$L = 1$ の層	$Y = 1$	$Y = 0$	
$A = 1$	x_1	$n_1 - x_1$	n_1
$A = 0$	y_1	$m_1 - y_1$	m_1
			N_1

層内では交絡は起きていないと考えられるので、ひとまずもし N_1 人全員が治療を受けた(受けなかった)ときに期待されるアウトカムは計算できるのでは？

簡単な例

つまり,

$L = 1$ の層で(N_1 人に対して)は, $N_1 \times \frac{x_1}{n_1}$ 人と $N_1 \times \frac{y_1}{m_1}$ 人が各群で

期待される $Y = 1$ の数で,

$L = 0$ の層で(N_0 人に対して)は, $N_0 \times \frac{x_0}{n_0}$ 人と $N_0 \times \frac{y_0}{m_0}$ 人が各群で

期待される $Y = 1$ の数である.



以上より, $N(= N_1 + N_0)$ 人に対しては, $N_1 \times \frac{x_1}{n_1} + N_0 \times \frac{x_0}{n_0}$ 人と

$N_1 \times \frac{y_1}{m_1} + N_0 \times \frac{y_0}{m_0}$ 人が各群で期待される $Y = 1$ の数になる.

簡単な例

よって因果効果は $\frac{1}{N} \sum_l N_l \times \frac{x_l}{n_l} - \frac{1}{N} \sum_l N_l \times \frac{y_l}{m_l}$ で推定される.

式変形すると, $\sum_l \frac{N_l}{N} \times \left(\frac{x_l}{n_l} - \frac{y_l}{m_l} \right)$ と表現できる.

⇒ 各層内の **部分集団** における治療効果を推定

➤ ここまで述べた方法の問題点

- 調整すべき交絡因子が多いと、各部分集団のサイズが小さくなり、推定が不安定になってしまう
- 交絡因子が連続変数のとき、カテゴリー化により情報が失われてしまう

そのため、部分集団に対する治療効果の推定のために回帰モデルを当てはめるなどといった工夫がなされる(次スライド).

Standardizationによる治療効果の推定量:

$$\hat{E}[Y^{a=1}] - \hat{E}[Y^{a=0}] = \sum_l (\hat{E}[Y|A = 1, L = l] - \hat{E}[Y|A = 0, L = l]) \times \Pr[L = l]$$

各層 l に対して $E[Y|A = a, L = l]$ を推定し, $\Pr[L = l]$ で重み付け平均することで治療効果を推定.

⇒ 本発表では $E[Y|A = a, L = l]$ には, 回帰モデルを当てはめ, パラメトリックに推定することを考える

Standardizationによる推定量は

$$\begin{aligned}\hat{E}[Y^a] &= \sum_l \hat{E}[Y|A = a, L = l] \times \Pr[L = l] \\ &= \frac{1}{n} \sum_{i=1}^n \hat{E}[Y|A = a, L_i]\end{aligned}$$

と表現でき, $\Pr[L = l]$ にモデルの仮定を置くことなく推定できる.

Standardization

SASを用いたStandardizationの解析

SASを用いた Standardization の解析

本節では、禁煙の例に対してStandardizationを用いて治療効果を推定する。

Standardizationの推定量：

$$\frac{1}{n} \sum_{i=1}^n \hat{E}[Y|A = a, L_i]$$

線形モデルを仮定して推定し、平均をとる。

⇒プログラムと出力結果を次に示す

SASを用いた Standardization の解析

```

proc genmod data = nhefs;
class exercise active education;
model wt82_71 = qsmk sex race age age*age education
smokeintensity smokeintensity*smokeintensity smokeyrs
smokeyrs*smokeyrs exercise active wt71 wt71*wt71 qsmk*smokeintensity;
output out = predicted_mean p = meanY ;
run;
quit;

```

最大尤度パラメータ推定値の分析

※出力は一部

パラメータ	自由度	推定値	標準誤差	Wald 95% 信頼限界	カイ 2 乗	Pr > ChiSq
Intercept	1	-1.6906	4.3594	-10.2348 6.8536	0.15	0.6982
qsmk	1	2.5596	0.8037	0.9844 4.1348	10.14	0.0014
sex	1	-1.4303	0.4658	-2.3432 -0.5173	9.43	0.0021
race	1	0.5601	0.5780	-0.5727 1.6929	0.94	0.3325
age	1	0.3596	0.1622	0.0417 0.6776	4.91	0.0266
age*age	1	-0.0061	0.0017	-0.0095 -0.0027	12.66	0.0004
education	1	0.1950	0.7364	-1.2483 1.6383	0.07	0.7912
education	2	0.9854	0.6965	-0.3797 2.3505	2.00	0.1571
education	3	0.7513	0.6297	-0.4828 1.9854	1.42	0.2328
education	4	1.6865	0.8658	-0.0104 3.3835	3.79	0.0514
education	5	0.0000	0.0000	0.0000 0.0000	.	.
smokeintensity	1	0.0491	0.0514	-0.0516 0.1498	0.91	0.3389
smokeinte#smokeinten	1	-0.0010	0.0009	-0.0028 0.0008	1.13	0.2877

2-3 IPW/Standardization

以上より,
Standardizationを用いた治療効果の推定値は

$$\hat{E}[Y^{a=1}] - \hat{E}[Y^{a=0}] = 5.18 - 1.66 = 3.5$$

95%信頼区間をブートストラップ法を用いて算出した結果,
禁煙により3.5kg (95%CI: 2.6~4.5kg) の体重増加と推定された。

IPW/Standardization

IPWとStandardizationの比較

治療効果の推定手法であるIPWとStandardizationについて紹介し、禁煙の例では両手法とも推定値がほとんど同じ値を示した。

実際、条件付き交換可能性とpositivityが成立し、かつ仮定したモデルが正しいときに両手法は同等である。

(疑問) では、治療効果を推定したいときにはどちらを使用すればよい？

今, どこにモデルを仮定しているのか着目すると, 異なる部分にモデルの仮定をしていると分かる.

➤ IPW: $\Pr[A = a|L]$

⇒治療モデル

➤ Standardization: $\Pr[Y|A = a, L = l]$

⇒アウトカムモデル

ある程度のモデルの誤特定は現実的に避けられない問題である.

⇒モデルの誤特定はバイアスを導く

しかし, 治療モデルの誤特定とアウトカムモデルの誤特定は効果の推定において, 一般に同じ大きさ, 方向のバイアスにはならない.

⇒IPWとStandardizationのモデルの誤特定による各点推定値への影響は一般に異なる.

そのため、IPWとStandardizationの推定値が大きく異なる場合、少なくともいずれかが深刻なモデルの誤特定をしていると考えられる。

以上の観点から、
使用可能であればモデルの妥当性の確認のために両手法とも推定に用いるのがよく、他の推定手法としてDoubly Robust推定量を考えることもできる。

BACK UP

- 拡張した重み: Stabilized weightの特徴
効果修飾因子を考慮したMSM
IPWとStandardizationの同等性の証明
Standardizationスライドの等式確認

拡張した重み: Stabilized weightの特徴

➤ $SW^A = f(A)/f(A|L)$

■ 重みの期待値は1 (疑似集団は20人)

■ 重みの幅が狭まり, 推定値は W^A より安定

■ **Stabilized weight**

□ $W^A = 1/f(A|L)$ (non-stabilized weightの拡張: 分子を1とした場合)

□ 点推定値: W^A と SW^A のいずれを利用しても同じ

□ 95%CI: 治療が時間依存性または連続量の場合はstabilized weightの方が狭くなる

- 治療効果が水準によって異なることが疑われ、各水準（例：性別）での効果に興味がある場合
 - 効果修飾因子 V を考慮可能
 - $E[Y^a|V] = \beta_0 + \beta_1 a + \beta_2 Va + \beta_3 V$
 - 解釈: $\beta_2 \neq 0 \rightarrow$ 性別による効果修飾因子あり

➤ 推定方法

- 重み: $SW^A(V) = f(A|V)/f(A|L)$ を用いて疑似集団を作成
- 注意事項: 交絡因子 L に効果修飾因子 V を含める
 - より安定した推定値となる (95%CIが狭くなる)
 - ✓ 直観的な説明: 分母・分子共に効果修飾因子 V が入っているため
- 効果修飾因子 V の水準毎に交換可能性が保たれる必要あり

- 期待値の定義より以下が成り立つ.

$$\begin{aligned} E \left[\frac{I(A = a)Y}{f[A|L]} \right] &= \sum_l \frac{1}{f[a|l]} \{ E[Y|A = a, L = l] f[a|l] \Pr[L = l] \} \\ &= \sum_l \{ E[Y|A = a, L = l] \Pr[L = l] \} \end{aligned}$$

- 条件付き交換可能性 ($Y^a \perp\!\!\!\perp A|L$) を仮定し, 2つの異なる証明を通じて, 平均が同じことを証明する (連続量の場合, 和を積分に置き換える).
- Standardized mean が counterfactual mean と同値であることの証明

$$\begin{aligned} \underline{E[Y^a]} &= \sum_l E[Y^a|L = l] \Pr[L = l] && (E[Y^a] = E[E[Y^a|L]]) \\ &= \sum_l E[Y^a|A = a, L = l] \Pr[L = l] && (\text{Exchangeability}) \\ &= \underline{\sum_l E[Y|A = a, L = l] \Pr[L = l]} && (\text{Consistency}) \\ &\quad \text{Standardized mean} \\ &\quad \text{Counterfactual mean} \end{aligned}$$

IPWとStandardizationの同等性の証明

(IPWとStandardizationの同等性の証明)

- IP weighting meanがcounterfactual meanと同値であることの証明

$$\begin{aligned} \underbrace{E \left[\frac{I(A=a)}{f[A|L]} Y \right]}_{\text{IP weighting mean}} &= E \left[\frac{I(A=a)}{f[A|L]} Y^a \right] && \text{(Consistency)} \\ &= E \left\{ E \left[\frac{I(A=a)}{f[A|L]} Y^a \mid L \right] \right\} && (E[Y] = E[E[Y|L]]) \\ &= E \left\{ E \left[\frac{I(A=a)}{f[A|L]} \mid L \right] E[Y^a | L] \right\} && \text{(Conditional exchangeability)} \\ &= E \{ E[Y^a | L] \} && \left(E \left[\frac{I(A=a)}{f[a|L]} \mid L \right] = 1 \right) \\ &= \underbrace{E[Y^a]}_{\text{Counterfactual mean}} && (E[E[Y|L]] = E[Y]) \end{aligned}$$

$$\sum_l \hat{E}[Y|A = a, L = l] \times \Pr[L = l] = \frac{1}{n} \sum_{i=1}^n \hat{E}[Y|A = a, L_i]$$

を確認する.

記号を定義:

集団人数を n 人とし, L は K カテゴリーで $l_1 \dots l_K$ とする.

L のそれぞれの層は $n_1 \dots n_K$ 人 ($n_1 + \dots + n_K = n$)いるとする.

$$\begin{aligned} & \sum_l \hat{E}[Y|A = a, L = l] \times \Pr[L = l] \\ &= \hat{E}[Y|A = a, L = l_1] \Pr[L = l_1] + \cdots + \hat{E}[Y|A = a, L = l_K] \Pr[L = l_K] \\ &= \frac{1}{n} \left\{ n_1 \cdot \hat{E}[Y|A = a, L = l_1] + \cdots + n_K \cdot \hat{E}[Y|A = a, L = l_K] \right\} \\ & \quad L = l_1 \text{の層内にいる} n_1 \text{人の回帰の足し合わせ} \\ &= \frac{1}{n} \sum_{i=1}^n \hat{E}[Y|A = a, L_i] \end{aligned}$$