

# 博士論文

共変量情報に基づいた潜在結果変数の

同時確率の識別と推定

Identification and estimation of joint probabilities

of potential outcomes based on covariate

information

横浜国立大学大学院 理工学府

数物・電子情報系理工学専攻 数学教育分野

新垣 隆生

Ryusei Shingaki

学位授与 2024年3月

## 論文要旨

潜在結果変数の同時確率は、(i) それらが識別可能であれば因果リスクも識別可能であるが、因果リスクが識別可能であっても潜在結果変数の同時確率は識別可能であるとは限らない [Pearl, 2009, Tian and Pearl, 2000], (ii) 説明可能 AI (XAI) の開発で重要となる「必要性」「十分性」「必要十分性」の確率的側面を評価することができる [Watson et al., 2021], という意味で因果推論の基本的な要素である。しかし、これらは仮定なしでは識別できないため、潜在結果変数の同時確率を評価するために単調性 [Pearl, 2009, Tian and Pearl, 2000] などの様々な仮定が用いられてきた。

本論文では、共変量の代理変数を用いた新しい識別条件を、観察研究と実験研究のそれぞれで提案する。提案した条件により潜在結果変数の同時確率が識別可能な場合、潜在結果変数の同時確率の推定問題は特異モデルの推定問題に帰着するため、最尤法のような標準的な統計的推定手法では評価することが難しい。そこで、本論文では観察研究の問題に対しては最適化問題に基づいた推定方法を、実験研究の問題に対してはプラグイン推定量に基づいた推定方法を提案する。

本論文は5部構成となっており、第1部「はじめに」、第2部「準備」、第3部「観察研究における代理共変量情報に基づいた潜在結果変数の同時確率の識別と推定」、第4部「実験研究における代理共変量情報に基づいた潜在結果変数の同時確率の識別と推定」、第5部「まとめ」となっている。第3部は、観察研究において共変量の代理変数を活用することで潜在結果変数の同時確率の識別と推定が可能であることを示す。第4部は、実験研究において共変量の代理変数を活用することで潜在結果変数の同時確率の識別と推定が可能であることを示す。

# 目次

1	はじめに	1
2	準備	5
2.1	構造的因果モデル	5
3	観察研究における代理共変量情報に基づいた潜在結果変数の同時確率の識別 と推定	11
3.1	準備	11
3.2	識別	13
3.2.1	Case 1: 因果リスクと一つの代理共変量に基づいた識別	13
3.2.2	Case 2: 因果リスクなしの二つの共変量に基づいた識別	19
3.3	推定	24
3.3.1	Case 1 における推定	24
3.3.2	Case 2 における推定	27
3.4	数値実験	31
3.5	適用例	33
3.6	考察	36
4	実験研究における代理共変量情報に基づいた潜在結果変数の同時確率の識別 と推定	37
4.1	問題設定と記法	37

4.2 識別 . . . . .	39
4.3 推定 . . . . .	48
4.4 数値実験 . . . . .	50
4.5 適用例 . . . . .	52
4.6 考察 . . . . .	54
<b>5 まとめ</b>	<b>57</b>
<b>参考文献</b>	<b>60</b>

# 第 1 部

## はじめに

同じ被験者がある処置を受けたとした場合の結果と受けなかったとした場合の結果を同時に導くことは、因果推論における本質的な問題である。例えば、COVID-19 ワクチンとプラセボを比較した第 III 相ランダム化臨床試験に関する米国食品医薬品局 (FDA) のガイダンスについて考えてみる。FDA のガイダンスでは、COVID-19 ワクチンについて、接種した人の少なくとも 50% において疾患を予防すること、もしくはその重症度を軽減することが認可の期待値として設定されている [Goodman et al., 2020]。また、ワクチンが軽症な患者だけでなくより重症な患者の症状の軽減、入院や死亡を低減するかどうかについても重要とされている。これを評価するには、ワクチンを接種した健常者が（反事実的に）接種しなかったとしたら重症度が上がっていたかどうか、ワクチンを接種しなかった健常者が（反事実的に）その時点で接種していたとしたら重症度が下がっていたかどうか、に対する不確実性を考慮することが必要となる。このような不確実性を考慮するためには、被験者のタイプを “doomed”, “causative”, “preventive”, “immune” の四つに分類し、これら四つの潜在結果変数の同時確率を評価することが有用である [Lash et al., 2021]。例としてあげた第 III 相ランダム化臨床試験において、“doomed” タイプの被験者はワクチンを接種してもプラセボを接種しても発症するためワクチン接種とは関係ない場合を、“causative” タイプの被験者はプラセボを接種された場合にのみ発症する場合を、“preventive” タイプの被験者はワクチンを接種された場合にのみ発症する場合

を, “immune” タイプの被験者はワクチンを接種してもプラセボを接種しても発症しないためワクチン接種とは関係ない場合を表す. 効果的なワクチン接種政策を取るには, ワクチン接種を受けると重症度が低下するが, ワクチン接種を受けないと重症度が低下しない “causative” タイプの被験者を対象とすることが望ましいと考えられる. 一方, “preventive” タイプの被験者の場合, ワクチン接種を受けないと重症度は低下するが, ワクチン接種を受けると重症度は低下しないため, ワクチン接種を受けることの重症度を評価するには “preventive” タイプの被験者である確率が有用であろう.

しかしながら, 各被験者がどのタイプに属するかは, (ランダム化実験であったとしても) 観測データからはわからない. Tian and Pearl [2000] や Kuroki and Cai [2011] は, 潜在結果タイプの確率に基づいて定義される原因の確率 (probabilities of causation) について, 実験研究や観察研究で得られるデータからバウンドを与える方法を示している. また, Tian and Pearl [2000] は単調性 (即ち, 被験者が “preventive” タイプである確率が 0 であること) と因果リスクが識別可能であるという仮定の下で, 原因の確率が識別可能であることを示しているが, これらの仮定が逸脱した場合に潜在結果タイプの確率を識別する方法については議論されていない.

**論文の構成** 本研究では共変量情報を用いた潜在結果変数の同時確率に対する識別と推定について論じる. 本論文は以下の 5 部構成となっている:

**第 1 部** はじめに

**第 2 部** 準備

**第 3 部** 観察研究における代理共変量情報に基づいた潜在結果変数の同時確率の識別と推定

**第 4 部** 実験研究における代理共変量情報に基づいた潜在結果変数の同時確率の識別と推定

## 第5部 まとめ

これらの内容は, Shingaki and Kuroki [2021], Shingaki and Kuroki [2023] において公表されている. 各部の概要は以下の通りである:

**第2部の概要** 第2部では, 第3部や第4部で必要となる構造的因果モデルの概念について, Bareinboim et al. [2022], Plečko and Bareinboim [forthcoming] に倣って説明する.

**第3部の概要** 第3部では, 観察研究における潜在結果変数の同時確率の識別と推定について考える. 既存研究では, 潜在結果変数の同時確率を評価するために, 単調性 [Pearl, 2009, Tian and Pearl, 2000], 潜在結果変数間の独立性 [Robins and Richardson, 2011], 利得等式 [Li and Pearl, 2019] など, 様々な仮定を利用してきた. 既存の識別条件とは異なり, このような仮定なしに潜在的な結果の結合確率を評価するために, 共変量情報を用いた2種類の新しい識別条件を提案する. 多くの観察研究では, 年齢, 性別, 人種などの個体の共変量が観測されており, 個体の共変量を非観測な変数  $U$  の代わりに観測される代理変数とみなすことができる [Kuroki and Pearl, 2014, Miao et al., 2018, Tchetgen et al., forthcoming]. ここでは, 潜在結果変数の同時確率を識別するために, 因果リスクと一つの共変量の代理変数を用いる方法 (定理 3.1) と, 因果リスクを用いずに二つの共変量の代理変数を用いるもの (定理 3.2) を提案する. また, 提案した条件により潜在結果変数の同時確率が識別可能なとき, 潜在結果変数の同時確率の推定問題は特異モデルの推定問題に帰着するため, 最尤法のような標準的な推定方法では評価することができない. この問題を解決するために, 推定量を得るための最適化問題を拡張 Lagrange 法で解くことによる新たな推定方法を提案する.

**第4部の概要** 第4部では, 実験研究における潜在結果変数の同時確率の識別と推定について考える. 観察研究と同様, ランダム化試験でも年齢, 性別, 人種などの

個体の共変量の測定も可能であり [Lui, 2011], 個体の共変量を非観測な変数  $U$  の代わりに観測される代理変数とみなすことができる [Kuroki and Pearl, 2014, Miao et al., 2018, Tchetgen et al., forthcoming]. ここでは, 潜在結果変数の同時確率を識別するために, 観察研究の場合と同様の二つの代理変数を用いるもの (定理 4.1) を提案する. 但し, 第 2 部で述べたような最適化問題に基づいた方法ではなく, 潜在結果変数の同時確率に対する推定量を具体的に構成する.



## 第 2 部

## 準備

### 2.1 構造的因果モデル

Bareinboim et al. [2022], Plečko and Bareinboim [forthcoming] に従い, 本論文で必要となる構造的因果モデルの諸概念について述べる.

**定義 2.1** (構造的因果モデル (Structural Causal Model, SCM) [Pearl, 2009]). 構造的因果モデルとは以下の要素からなる四つ組  $\langle \mathbf{V}, \mathbf{U}, \mathcal{F}, p(\mathbf{u}) \rangle$  のことである:

- (1)  $\mathbf{U}$  はモデルの外側の要因で決まる外生変数からなる集合.
- (2)  $\mathbf{V} = \{V_1, \dots, V_m\}$  はモデルの変数 (即ち,  $\mathbf{U} \cup \mathbf{V}$  の変数) で決まる内生変数からなる集合.
- (3)  $\mathcal{F} = \{f_1, \dots, f_m\}$  は  $V$  を決定する構造的関数  $v_i \leftarrow f_i(\text{pa}(v_i), u_i)$ ,  $i = 1, \dots, m$  からなる集合, ここで,  $\text{pa}(V_i) \subseteq \mathbf{V} \setminus V_i$  と  $U_i \subseteq \mathbf{U}$  は  $f_i$  の引数を表す.
- (4)  $p(\mathbf{u})$  は外生変数  $\mathbf{U}$  の分布.

各構造的因果モデルは, 現象に関与する変数を外生変数 (未観測変数) の集合  $\mathbf{U}$  と内生変数 (観測変数) の集合  $\mathbf{V}$  に分割しているとみなすことができる. 外生変数はモデルの「外側」で決定され, それらの関連する確率分布  $p(\mathbf{u})$  は, 調べている現象の外側の世界を要約したものとみなせる. これらの変数は, ユニットと呼ばれる現象に関与する単位を表し, 調査中の母集団の要素 (例えば, 患者, 学生, 顧客な

ど)に対応する。これらの  $(p(\mathbf{u})$  で表される) ランダム性は内生集合  $\mathbf{V}$  の変動を引き起こす。モデル内部では、各内生変数  $V_i$  の値は外生要因  $U_i$  と内生変数からなる集合  $\text{pa}(v_i)$  の集合を  $V_i$  に写像する因果的な過程  $V_i \leftarrow f_i(\text{pa}(v_i), u_i)$  によって決定される。

**定義 2.2** (部分モデル (Submodel) [Pearl, 2009]).  $\mathcal{M} = \langle \mathbf{V}, \mathbf{U}, \mathcal{F}, p(\mathbf{u}) \rangle$  を構造的因果モデルとし、 $\mathbf{X}$  は  $\mathbf{V}$  の変数からなる集合、 $\mathbf{x}$  は  $\mathbf{X}$  が取り得る特定の値とする。このとき、 $\mathcal{M}$  の部分モデル  $\mathcal{M}_{\mathbf{x}}$  は

$$\mathcal{M}_{\mathbf{x}} = \langle \mathbf{V}, \mathbf{U}, \mathcal{F}_{\mathbf{x}}, p(\mathbf{u}) \rangle \quad (2.1)$$

のことである。ここで、

$$\mathcal{F}_{\mathbf{x}} = \{f_i: V_i \notin \mathbf{X}\} \cup \{\mathbf{X} \leftarrow \mathbf{x}\}. \quad (2.2)$$

SCM  $\mathcal{M}_{\mathbf{x}}$  は、変数  $\mathbf{X}$  に関連する  $\mathcal{F}$  のすべての方程式を、 $\mathbf{X}$  を特定の値  $\mathbf{x}$  に設定した方程式で置き換えることによって、 $\mathcal{M}$  から得られる。

部分モデルの概念に基づいて、潜在結果変数を導入する:

**定義 2.3** (潜在結果変数 (Potential Outcome) [Rubin, 1974, Pearl, 2009]).  $\mathbf{X}$  と  $\mathbf{Y}$  は  $\mathbf{V}$  の変数からなる二つの集合であり、 $\mathbf{u} \in \mathbf{U}$  はユニットであるとする。潜在結果変数  $\mathbf{Y}_{\mathbf{x}}(\mathbf{u})$  は、 $\mathbf{U} = \mathbf{u}$  で評価された方程式からなる集合  $\mathcal{F}_{\mathbf{x}}$  の  $\mathbf{Y}$  に対する解として定義される。即ち、 $\mathbf{Y}_{\mathbf{x}}(\mathbf{u})$  は  $\mathcal{M}$  の部分モデル  $\mathcal{M}_{\mathbf{x}}$  における  $\mathbf{Y}$  の解を表す。

$\mathbf{Y}_{\mathbf{x}}(\mathbf{u})$  は、特定のユニット  $\mathbf{u}$  について、(ひょっとすると観測された事実と反して)  $\mathbf{X}$  が  $\mathbf{x}$  に設定された場合に、変数  $\mathbf{Y}$  が取るであろう値を表す。

各 SCM  $\mathcal{M}$  は、異なるデータ収集モードを表し、本論文で重要な役割を果たす異なるタイプの確率分布を生み出す。まず、意思決定過程に干渉することなく、分析者がただ受動的にデータを収集する基礎となる意思決定システムの状態を表す観測分布を定義する:

**定義 2.4** (観測分布 (Observational Distribution)).  $SCM \mathcal{M} = \langle \mathbf{V}, \mathbf{U}, \mathcal{F}, p(\mathbf{u}) \rangle$  から導かれる, 各  $\mathbf{Y} \subseteq \mathbf{V}$  に対して

$$p(\mathbf{y}) = \sum_{\mathbf{u}} \mathbf{1}(\mathbf{Y}(\mathbf{u}) = \mathbf{y}) p(\mathbf{u}) \quad (2.3)$$

であるような同時確率分布  $p(\mathbf{V})$  を観測分布という. ここで,  $\mathbf{Y}(\mathbf{u})$  は  $U = \mathbf{u}$  で  $\mathcal{F}$  を評価した後の解である.

次に, 反事実的な結果変数に関する, もう一つの重要な分布族を定義する.

**定義 2.5** (反事実分布 (Counterfactual Distribution)).  $\mathcal{M} = \langle \mathbf{V}, \mathbf{U}, \mathcal{F}, p(\mathbf{u}) \rangle$  を  $SCM$  とし,  $\mathbf{Y}_1, \dots, \mathbf{Y}_k \subset \mathbf{V}$  と  $\mathbf{X}_1, \dots, \mathbf{X}_k \subset \mathbf{V}$  は観測変数からなる部分集合,  $\mathbf{x}_1, \dots, \mathbf{x}_k$  は  $\mathbf{X}_1, \dots, \mathbf{X}_k$  の特定の値であるとする. また, 変数  $\mathbf{Y}_i$  の  $\mathbf{X}_i = \mathbf{x}_i$  と設定したときの潜在結果変数を  $(\mathbf{Y}_i)_{\mathbf{x}_i}$  で表す.  $SCM \mathcal{M}$  から導かれる

$$p((\mathbf{Y}_1)_{\mathbf{x}_1} = \mathbf{y}_1, \dots, (\mathbf{Y}_k)_{\mathbf{x}_k} = \mathbf{y}_k) = \sum_{\mathbf{u}} \mathbf{1} \left( \bigwedge_{i=1}^k (\mathbf{Y}_i)_{\mathbf{x}_i}(\mathbf{u}) = \mathbf{y}_i \right) p(\mathbf{u}) \quad (2.4)$$

であるような反事実事象  $(\mathbf{Y}_1)_{\mathbf{x}_1}, \dots, (\mathbf{Y}_k)_{\mathbf{x}_k}$  にわたる同時分布族を反事実分布という.

式 (2.4) は次のように解釈できる:

- (1) 各変数集合  $\mathbf{Y}_1, \dots, \mathbf{Y}_k$  に関する各添字集合  $\mathbf{X}_1, \dots, \mathbf{X}_k$  に対して,  $\mathcal{F}_{\mathbf{x}_1}, \dots, \mathcal{F}_{\mathbf{x}_k}$  を生成する適切な定数で対応するメカニズムを置き換え, 部分モデル  $\mathcal{M}_{\mathbf{x}_1}, \dots, \mathcal{M}_{\mathbf{x}_k}$  を作る.
- (2) 各ユニット  $U = \mathbf{u}$  に対して, 観測変数の潜在結果変数を得るために修正されたメカニズム  $\mathcal{F}_{\mathbf{x}_1}, \dots, \mathcal{F}_{\mathbf{x}_k}$  を評価する.
- (3) 確率  $P(U = \mathbf{u})$  が反事実変数にわたる事象と一致する, 即ち,  $(\mathbf{Y}_1)_{\mathbf{x}_1} = \mathbf{y}_1, \dots, (\mathbf{Y}_k)_{\mathbf{x}_k} = \mathbf{y}_k$  となるように各ユニット  $U = \mathbf{u}$  に対して足し上げられる.

本論文では主に  $\mathbf{Y}$  と  $\mathbf{X}$  がスカラーのときを考えるため,  $\mathbf{Y}$  と  $\mathbf{X}$  を改めて  $Y$  と  $X$  と書き,  $x, x' \in \{x_0, x_1\}$  と  $y, y' \in \{y_0, y_1\}$  に対して,

$$p(Y_x = y) = \sum_{\mathbf{u}} \mathbf{1}(Y_x(\mathbf{u}) = y) p(\mathbf{u}),$$

$$p(Y_x = y, Y_{x'} = y') = \sum_{\mathbf{u}} \mathbf{1}(Y_x(\mathbf{u}) = y \wedge Y_{x'}(\mathbf{u}) = y') p(\mathbf{u}) \text{ for } x \neq x'.$$

といった量について考えていく.  $x$  と  $x'$  が両立しないのであれば  $Y_x$  と  $Y_{x'}$  を同時に測定することはできず, 「 $X = x$  であれば  $Y$  は  $y$  になり,  $X = x'$  であれば  $Y$  は  $y'$  になる」という同時の主張に確率を帰属させることは無意味であるとして, 反事実を同時分布する確率変数として扱うことに対する異論もある [Dawid, 2000].  $Y_x$  と  $Y_{x'}$  を  $U$  上の標準的な確率空間に支配される二つの異なる部分モデルにおける  $Y$  の解の観点から定義することで, 矛盾する同時の主張を  $U$  の空間における通常の事象として解釈することで, このような異論にある程度反論することができる.

また, 実際の観測値を条件とする反事実の確率を自然に定義することもできる. 例えば, 事象  $X = x$  が事象  $Y = y$  の「原因であった」確率は,  $X = x$  と  $Y = y$  が実際に発生したとすると,  $X$  が  $x$  でなかったら  $Y$  は  $y$  と等しくなかったという確率と解釈できる. このような確率は, 先のモデルでうまく定義されており, 必要性の確率 (Probability of Necessity, PN)  $p(Y_{x_0} = y_0 | X = x_1, Y = y_1)$  は

$$\begin{aligned} p(Y_{x_0} = y_0 | X = x_1, Y = y_1) &= \frac{p(Y_{x_0} = y_0, X = x_1, Y = y_1)}{p(X = x_1, Y = y_1)} \\ &= \sum_{\mathbf{u}} \mathbf{1}(Y_{x_0}(\mathbf{u}) = y_0) p(\mathbf{u} | X = x_1, Y = y_1), \end{aligned} \quad (2.5)$$

十分性の確率 (Probability of Sufficiency, PS)  $p(Y_{x_1} = y_1 | X = x_0, Y = y_0)$  は

$$\begin{aligned} p(Y_{x_1} = y_1 | X = x_0, Y = y_0) &= \frac{p(Y_{x_0} = y_0, X = x_1, Y = y_1)}{p(X = x_1, Y = y_1)} \\ &= \sum_{\mathbf{u}} \mathbf{1}(Y_{x_1}(\mathbf{u}) = y_1) p(\mathbf{u} | X = x_0, Y = y_0) \end{aligned} \quad (2.6)$$

のように評価できる. また, 必要十分性の確率 (Probability of Necessity and Sufficiency, PNS) は  $p(Y_{x_1} = y_1, Y_{x_0} = y_0)$  と定義される [Pearl, 2009].

SCM はうまく定義されているが, 通常, 観測可能ではない. 基礎となる SCM に関する仮定を符号化するために, 次に定義する因果ダイアグラムが用いられる.

**定義 2.6** (因果ダイアグラム (Causal Diagram) [Pearl, 2009]).  $\mathcal{M} = \langle \mathbf{V}, \mathbf{U}, \mathcal{F}, P(\mathbf{u}) \rangle$  は SCM であるとする. 次のようなグラフ  $\mathcal{G}$  を ( $\mathcal{M}$  の) 因果ダイアグラムという:

- (1) すべての内生変数  $V_i \in \mathbf{V}$  に対して頂点がある.
- (2)  $V_i$  が  $f_j \in \mathcal{F}$  の引数として現れるとき, 有向辺  $V_i \rightarrow V_j$  がある.
- (3)  $U_i, U_j \in \mathbf{U}$  が相関を持つ, もしくは対応する関数  $f_i, f_j$  がある  $U_{ij} \in \mathbf{U}$  を引数として共有するとき, 対応する双方向辺  $V_i \leftrightarrow V_j$  がある.

$V_j$  が  $V_i$  の値に基づいて決定されるときはいつでも, 内生変数  $V_i$  から  $V_j$  への有向辺が存在する. 同様に,  $V_i$  と  $V_j$  の間の双方向辺の存在は,  $V_i$  と  $V_j$  の両方が値を得る方法に影響する, 共有された, 観測されていない情報があることを示す. SCM がすべての構造的メカニズム  $\mathcal{F}$  と外生変数  $P(\mathbf{u})$  に関する明示的な情報を含むのに対して, 因果ダイアグラムは  $\mathcal{F}$  の関数への入力としてどの関数の引数が用いられる可能性があるかといった情報のみを符号化する. 有向辺  $V_i \rightarrow V_j$  の存在は,  $V_j$  のメカニズムが変数  $V_i$  に基づいて決定される可能性を符号化するが, 必要性は符号化しない. 例えば,  $f_j$  は  $V_i$  の値を考慮しないように決めるかもしれない. 一方, この仮定はダイアグラムに存在する辺ではなく辺がないことに符号化され, 辺がないことは一方の変数が他方の引数ではないことを意味する.

$\mathbf{Y}_x(\mathbf{u})$  は次の二つの性質を満たす:

**有効性 (effectiveness)** 任意の二つの排反な変数集合  $\mathbf{Y}$  と  $\mathbf{W}$  に対して,

$$\mathbf{Y}_{yw}(\mathbf{u}) = \mathbf{y} \tag{2.7}$$

言い換えれば,  $\mathbf{W}$  の変数を  $w$  に設定しても,  $\mathbf{Y}$  を  $\mathbf{y}$  に設定すれば  $\mathbf{Y}$  には何の影響もないことを意味する.

**構成性 (composition)** 任意の二つの排反な変数集合  $\mathbf{X}$  と  $\mathbf{W}$ , および任意の変数集合  $\mathbf{Y}$  に対して,

$$\mathbf{W}_x(\mathbf{u}) = w \implies \mathbf{Y}_{xw}(\mathbf{u}) = \mathbf{Y}_x(\mathbf{u}) \tag{2.8}$$

言い換えれば,  $\mathbf{X}$  を  $x$  に設定すれば,  $\mathbf{W}$  の変数を ( $x$  の下で) 到達するのと同じ値  $\mathbf{W}$  に設定しても,  $\mathbf{Y}$  には何の影響もないはずであることを意味する.

$M$ が再帰的である（即ち、 $G$ が非巡回である）ときはいつでも、有効性と構成性は完全である [Galles and Pearl, 1998]. 即ち、構造的因果モデルの意味論から導かれる反事実のすべての性質は、有効性と構成性を繰り返し適用することによって導き出すことができる。さらに、 $\mathbf{Y}_\emptyset(\mathbf{u}) = \mathbf{Y}(\mathbf{u})$ と定義することで、構成性から以下の性質が直ちに従う:

**系 2.1** (一致性 (consistency)). 任意の二つの排反な変数集合  $\mathbf{Y}$  と  $\mathbf{X}$  に対して,

$$\mathbf{X}(\mathbf{u}) = \mathbf{x} \implies \mathbf{Y}(\mathbf{u}) = \mathbf{Y}_x(\mathbf{u}). \quad (2.9)$$

一致性は、あるユニット  $\mathbf{u}$  において変数  $\mathbf{X}$  が値  $\mathbf{x}$  であることを観測し、介入して  $\mathbf{X}$  を同じ値  $\mathbf{x}$  に設定した場合、結果変数  $\mathbf{Y}$  にいかなる変化も期待すべきではないということを意味する。この性質は、第3部と第4部の結果を示すのに多用される。

## 第 3 部

# 観察研究における代理共変量情報に基づいた潜在結果変数の同時確率の識別と推定

第 3 部では，観察研究の枠組みにおいて潜在結果変数の同時確率の識別および推定問題について考える．3.1 節では，3.2.1 節では，共変量の一つの代理変数に加えて因果リスクが既知である場合の，潜在結果変数の同時確率の識別問題について考える．また，3.2.2 節では，因果リスクなしに共変量の一つの代理変数が観測できている場合の，潜在結果変数の同時確率の識別問題について考える．3.2 節の結果に基づいて，3.3 節では潜在結果変数の同時確率に対する新たな推定量を提案する．最後に，数値実験と適用例を通して提案した推定量の有用性について検証する．第 3 部の内容は Shingaki and Kuroki [2021] で公表されている．

### 3.1 準備

図 3.1 で示されるような，曝露（例えば喫煙）の結果を対照的な曝露（例えば非喫煙）の結果と比較することを目的とした観察研究について考える． $x \in \{x_0, x_1\}$  と  $y \in \{y_0, y_1\}$  に対して， $p(X = x, Y = y) = p(x, y)$  は  $(X, Y) = (x, y)$  の同時確率， $p(Y = y | X = x) = p(y | x)$  は  $Y = y$  の  $X = x$  が与えられた下での条件付き確率，

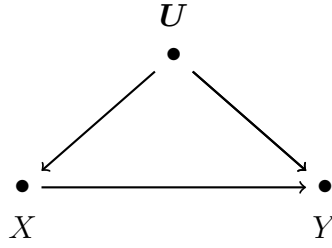


図 3.1: 第 3 部で考える因果グラフ.

$p(X = x) = p(x)$  は  $X = x$  の周辺確率を表すものとする. また,  $x \in \{x_0, x_1\}$  に対して, ユニット  $\mathbf{u}$  は潜在結果変数  $Y_x(\mathbf{u})$  を持ち, これはユニット  $\mathbf{u}$  について  $X$  が  $x$  であったと仮定したときに  $Y$  が取ったであろう値を表すものとする.  $Y_x(i) = y$  は「ユニット  $\mathbf{u}$  に対して  $X$  を実験的に  $x$  と設定したときに,  $Y$  が値  $y$  を取る」もしくは反事実的な文「ユニット  $\mathbf{u}$  に対して  $X$  が  $x$  であったとすると  $Y$  が  $y$  であったであろう」ということを意味する. 潜在結果変数  $Y_x$  は  $X$  が  $x$  であるときのみ観測される, 即ち, ユニット  $\mathbf{u}$  に対して

$$X(\mathbf{u}) = x \implies Y_x(\mathbf{u}) = Y \quad (3.1)$$

が成り立つ. この性質は 2.1 節で述べたように一貫性 (consistency) と呼ばれる [Robins, 1989, Pearl, 2009].

ここでは, SUTVA (stable unit treatment value assumption) を仮定する [Imbens and Rubin, 2015]. これは次のように要約される:

- (i) 任意の被験者の処置の状態は他の被験者の結果に影響しない (no interference),
- (ii) すべての被験者の処置は比較可能 (no variation in treatment).

つまり,  $n$  人の被験者が関心のある母集団からのランダムサンプルとみなせるとき,  $Y_x(\mathbf{u})$  は確率変数  $Y_x$  の実現値とみなせる.

$X = x$  の  $Y = y$  に対する因果リスク (causal risk) は  $p(Y_x = y)$  と定義され,  $X = x_0$  と比較した  $X = x_1$  の因果リスク差 (causal risk difference) は  $p(Y_{x_1} =$



$y_1) - p(Y_{x_0} = y_1)$  と定義される．ここで，識別可能であるとは， $p(Y_x = y)$  のような因果的な量が観測変数の同時確率から一致推定できることを意味する．

## 3.2 識別

### 3.2.1 Case 1: 因果リスクと一つの代理共変量に基づいた識別

Pearl [2009] に従い， $U$  は  $X$  と  $Y$  に影響し得るあらゆる共変量であるとし，図 3.2a で示される有向グラフに基づいて潜在結果変数の同時確率を評価する問題について考える．ここで，図 3.2a の共変量  $Z$  は  $U$  の代理変数として測定されている．図 3.2a において， $X$  から  $Y$  への有向辺  $X \rightarrow Y$  は  $X$  が  $Y$  への効果を持ち得ることを示し， $Y$  から  $X$  への有向辺  $Y \rightarrow X$  がないことは  $Y$  が  $X$  の原因とはなり得ないことを示している． $U$  から  $Y$  への  $X$  を通じた有向パス  $U \rightarrow X \rightarrow Y$  は， $U$  のいくつかの要素が  $X$  で媒介されて  $Y$  に影響するかもしれないことを意味する．図 3.2a は次のデータ生成過程を視覚的に表している：

$$Y = g_y(X, U, \epsilon_y), \quad X = g_x(U, \epsilon_x), \quad Z = g_z(U, \epsilon_z). \quad (3.2)$$

ここで， $\epsilon_x, \epsilon_y, \epsilon_z$  は独立な攪乱項 (random disturbances) であり， $U$  とも独立である．2.1 節で述べたように，式 (3.2) のような構造方程式モデルがデータ生成過程を表すのに使われるとき，図 3.2a で示される対応するグラフは因果ダイアグラム (causal diagram) と呼ばれる．

図 3.2a のような状況は，Pearl [2010] や Kuroki and Pearl [2014] でも議論されているが，ここではこれらの研究と異なり， $U$  は被験者の処置への反応の仕方に影響するあらゆるすべての共変量を含むことができるので， $U$  が与えられた下で  $\{X, Y\}$  と独立な共変量  $Z$  および  $\{X, U\}$  が与えられた下で  $Z$  と独立な共変量  $W$  の存在を仮定できる場面は多いと考えられる．このとき， $U \cup \{\epsilon_x, \epsilon_y, \epsilon_z\}$  の複雑な構造に依らず， $U$  の  $Y$  への影響は，二つの二値変数  $X$  と  $Y$  の間の四つの関数に制限され， $U \cup \{\epsilon_y, \epsilon_z, \epsilon_w\}$  が取る値はこれらの四つの関数の一つを選択することに対応

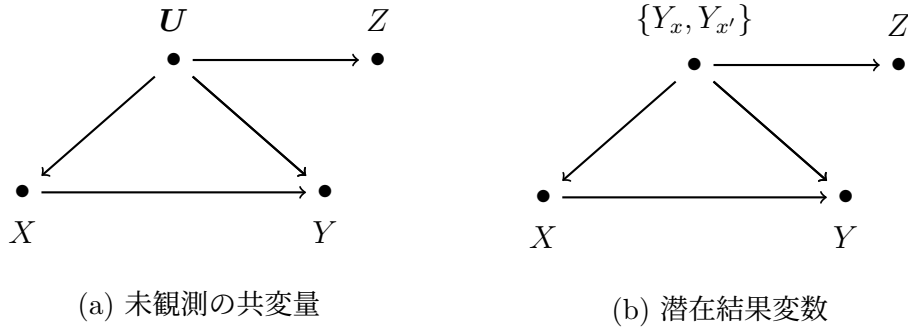


図 3.2: 3.2.1 節のグラフ表現.

する [Pearl, 2009]. Rothman et al. [2008, p. 59] に従い,  $U \cup \{\epsilon_x, \epsilon_y, \epsilon_z\}$  の状態を次の四つの潜在結果タイプに分割する:

$u_1 = (Y_{x_1} = y_1, Y_{x_0} = y_1)$  “doomed” 曝露の有無に依らずに疾患が生じる状況.

$u_2 = (Y_{x_1} = y_1, Y_{x_0} = y_0)$  “causative” 疾患が生じるときかつそのときのみ被験者が曝露される状況.

$u_3 = (Y_{x_1} = y_0, Y_{x_0} = y_1)$  “preventive” 疾患が生じるときかつそのときのみ被験者が曝露されない状況.

$u_4 = (Y_{x_1} = y_0, Y_{x_0} = y_0)$  “immune” 曝露の有無に依らず疾患が生じない状況.

この分割に従い, 値  $u \in \{u_1, u_2, u_3, u_4\}$  を取る変数として  $U$  を再定義する. このとき, Pearl [2009, ch.8] によるノンコンプライアンスの問題の設定と同様に, 図 3.2a を図 3.2b として描き直し,  $X, Y, Z$  の同時確率  $p(x, y, z)$  を図 3.2b に従い,

$$p(x, y, z) = \sum_{i=1}^4 p(y | x, z, u_i) p(x | u_i) p(z | u_i) p(u_i)$$

と分解できると仮定する. このとき, この分解は  $U$  が与えられた下で  $X$  が  $Z$  と条件付き独立であることを意味し,  $X \perp\!\!\!\perp Z | U$  と表す. さらに, 一致性から

$$p(x_1, y_1, z, u_i) = p(x_1, z, u_i), \quad p(x_1, y_0, z, u_i) = 0, \quad i = 1, 2,$$

$$p(x_1, y_0, z, u_i) = p(x_1, z, u_i), \quad p(x_1, y_1, z, u_i) = 0, \quad i = 3, 4,$$

$$p(x_0, y_1, z, u_i) = p(x_0, z, u_i), \quad p(x_0, y_0, z, u_i) = 0, \quad i = 1, 3,$$

$$p(x_0, y_0, z, u_i) = p(x_0, z, u_i), \quad p(x_0, y_1, z, u_i) = 0, \quad i = 2, 4$$

となる. よって, 例えば  $p(x_1, y_1, z)$ ,  $p(z)$ ,  $p(Y_{x_1} = y_1, z)$  は

$$p(x_1, y_1, z) = \sum_{i=1}^2 p(x_1 | u_i) p(z | u_i) p(u_i), \quad p(z) = \sum_{i=1}^4 p(z | u_i) p(u_i),$$

$$p(Y_{x_1} = y_1, z) = p(Y_{x_1} = y_1 | z) p(z)$$

と書き直すことができる. このとき,  $Z \in \{z_1, z_2, \dots, z_k\}$  ( $k \geq 4$ ) であり,

$$P = \begin{pmatrix} 1 & p(z_1) & p(z_2) & p(z_3) \\ p(x_1, y_1) & p(x_1, y_1, z_1) & p(x_1, y_1, z_2) & p(x_1, y_1, z_3) \\ p(x_1, y_0) & p(x_1, y_0, z_1) & p(x_1, y_0, z_2) & p(x_1, y_0, z_3) \\ p(x_0, y_1) & p(x_0, y_1, z_1) & p(x_0, y_1, z_2) & p(x_0, y_1, z_3) \end{pmatrix}, \quad (3.3)$$

$$Q = \begin{pmatrix} 1 & p(z_1) & p(z_2) & p(z_3) \\ p(Y_{x_1} = y_1) & p(Y_{x_1} = y_1, z_1) & p(Y_{x_1} = y_1, z_2) & p(Y_{x_1} = y_1, z_3) \\ p(Y_{x_0} = y_1) & p(Y_{x_0} = y_1, z_1) & p(Y_{x_0} = y_1, z_2) & p(Y_{x_0} = y_1, z_3) \end{pmatrix}, \quad (3.4)$$

と置くと, 次の定理を得る:

**定理 3.1.**  $Z$  は  $k \geq 4$  個の値  $z_1, z_2, \dots, z_k$  を取る変数であるとする. 次の条件が満たされるとき,  $p(Y_{x_1} = y, Y_{x_0} = y')$  と  $p(x | Y_{x_1} = y, Y_{x_0} = y')$  は識別可能である ( $y, y' \in \{y_1, y_0\}; x \in \{x_1, x_0\}$ ):

(1)  $z \in \{z_1, \dots, z_4\}$  に対して,  $p(Y_{x_1} = y_1 | z)$ ,  $p(Y_{x_0} = y_1 | z)$ ,  $p(x_1, y_1, z)$ ,  $p(x_0, y_1, z)$ ,

$p(x_1, y_0, z)$ ,  $p(x_0, y_0, z)$  が得られる.

(2)  $\{X, Y\} \perp\!\!\!\perp Z | U$  が成り立つ.

(3)  $P$  は可逆であり,  $QP^{-1}$  は四つの異なる非ゼロの列ベクトルからなる.

証明. 定理 3.1 の条件 (1) と条件 (2) から, 一致性により,  $z \in \{z_1, \dots, z_4\}$  に対して

$$\begin{aligned}
p(x_1, y_1, z) &= \sum_{i=1,2} p(x_1 | u_i) p(z | u_i) p(u_i), & p(x_0, y_1, z) &= \sum_{i=1,3} p(x_0 | u_i) p(z | u_i) p(u_i), \\
p(x_1, y_0, z) &= \sum_{i=3,4} p(x_1 | u_i) p(z | u_i) p(u_i), & p(x_0, y_0, z) &= \sum_{i=2,4} p(x_0 | u_i) p(z | u_i) p(u_i), \\
p(x_1, y_1) &= \sum_{i=1,2} p(x_1 | u_i) p(u_i), & p(x_0, y_1) &= \sum_{i=1,3} p(x_0 | u_i) p(u_i), \\
p(x_1, y_0) &= \sum_{i=3,4} p(x_1 | u_i) p(u_i), & p(x_0, y_0) &= \sum_{i=2,4} p(x_0 | u_i) p(u_i), \\
p(y_{x_1}, z) &= p(y_{x_1} | z) p(z) = \sum_{i=1,2} p(z | u_i) p(u_i), \\
p(y_{x_0}, z) &= p(y_{x_0} | z) p(z) = \sum_{i=1,3} p(z | u_i) p(u_i), \\
p(z) &= \sum_{i=1}^4 p(z | u_i) p(u_i), & p(y_{x_1}) &= \sum_z p(y_{x_1}, z), & p(y_{x_0}) &= \sum_z p(y_{x_0}, z)
\end{aligned}$$

を得る. よって,

$$\begin{aligned}
P &= \begin{pmatrix} 1 & p(z_1) & p(z_2) & p(z_3) \\ p(x_1, y_1) & p(x_1, y_1, z_1) & p(x_1, y_1, z_2) & p(x_1, y_1, z_3) \\ p(x_1, y_0) & p(x_1, y_0, z_1) & p(x_1, y_0, z_2) & p(x_1, y_0, z_3) \\ p(x_0, y_1) & p(x_0, y_1, z_1) & p(x_0, y_1, z_2) & p(x_0, y_1, z_3) \end{pmatrix}, \\
Q &= \begin{pmatrix} 1 & p(z_1) & p(z_2) & p(z_3) \\ p(y_{x_1}) & p(y_{x_1}, z_1) & p(y_{x_1}, z_2) & p(y_{x_1}, z_3) \\ p(y_{x_0}) & p(y_{x_0}, z_1) & p(y_{x_0}, z_2) & p(y_{x_0}, z_3) \end{pmatrix}, \\
R &= \begin{pmatrix} 1 & p(x_1 | u_1) & 0 & p(x_0 | u_1) \\ 1 & p(x_1 | u_2) & 0 & 0 \\ 1 & 0 & p(x_1 | u_3) & p(x_0 | u_3) \\ 1 & 0 & p(x_1 | u_4) & 0 \end{pmatrix},
\end{aligned}$$

$$S = \begin{pmatrix} 1 & p(z_1 | u_1) & p(z_2 | u_1) & p(z_3 | u_1) \\ 1 & p(z_1 | u_2) & p(z_2 | u_2) & p(z_3 | u_2) \\ 1 & p(z_1 | u_3) & p(z_2 | u_3) & p(z_3 | u_3) \\ 1 & p(z_1 | u_4) & p(z_2 | u_4) & p(z_3 | u_4) \end{pmatrix},$$

$$M = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}, \quad \Delta = \begin{pmatrix} p(u_1) & 0 & 0 & 0 \\ 0 & p(u_2) & 0 & 0 \\ 0 & 0 & p(u_3) & 0 \\ 0 & 0 & 0 & p(u_4) \end{pmatrix},$$

と置くと,

$$P = R^\top \Delta S, \quad Q = M \Delta S, \quad (3.5)$$

を得る. ここで, “ $\top$ ” は転置を表す. 定理 3.1 の条件 (3) から  $P$  は可逆なので, 式 (3.5) から,  $R$  は線形方程式

$$QP^{-1}R^\top = M. \quad (3.6)$$

の解として得られる.  $i = 1, \dots, 4$  に対して,  $a_i$  と  $m_i$  はそれぞれ  $3 \times 4$  行列  $QP^{-1}$  と  $M$  の第  $i$  列ベクトル, 即ち,  $QP^{-1} = (a_1; a_2; a_3; a_4)$  と  $M = (m_1; m_2; m_3; m_4)$  であるとする. 式 (3.6) から,

$$QP^{-1}R^\top = (a_1; a_2; a_3; a_4)R^\top = M = (m_1; m_2; m_3; m_4),$$

または同じことであるが

$$\begin{aligned} a_1 + p(x_1 | u_1)a_2 + \{1 - p(x_1 | u_1)\}a_4 &= m_1, & a_1 + p(x_1 | u_2)a_2 &= m_2, \\ a_1 + p(x_1 | u_3)a_3 + \{1 - p(x_1 | u_3)\}a_4 &= m_3, & a_1 + p(x_1 | u_4)a_3 &= m_4, \end{aligned}$$

であるので,

$$\begin{aligned} p(x_1 | u_1)(a_2 - a_4) &= m_1 - a_1 - a_4, & p(x_1 | u_2)a_2 &= m_2 - a_1, \\ p(x_1 | u_3)(a_3 - a_4) &= m_3 - a_1 - a_4, & p(x_1 | u_4)a_3 &= m_4 - a_1 \end{aligned}$$

を得る. この式から, 定理 3.1 の条件 (3) の下で,  $j = 1, \dots, 4$  に対して  $a_j$  と  $m_j$  は識別可能であるので,  $p(x_1 | u_j)$  と  $p(x_0 | u_j) = 1 - p(x_1 | u_j)$  は識別可能である. さらに, 式 (3.5) から  $(R^\top)^{-1}P = \Delta S$  であり,  $\Delta S$  の第 1 列は  $(p(u_1), \dots, p(u_4))^\top$  である. よって,  $P$  と  $R$  の両方が識別可能であるので,  $(R^\top)^{-1}P$  と  $\Delta S$  の第 1 列を比較することで,  $p(u_1), \dots, p(u_4)$  が識別可能であることがわかる.  $\square$

定理 3.1 から,  $i = 1, \dots, 4$  に対して

$$p(x_1, u_i) = p(x_1 | u_i)p(u_i), \quad p(x_0, u_i) = p(x_0 | u_i)p(u_i)$$

は  $R, \Delta, P$  から識別可能であるので, 2.1 節で述べた PN, PS, PNS は識別可能となる.

同時線形方程式 (3.6) が解を持たないとき, 少なくとも一つの条件が成り立たないことから, 定理 3.1 の仮定からの逸脱は検証可能であることに注意する. また, 定理 3.1 の条件 (1) は観測できない分布が既知であることを要請しているように見えるかもしれないが, 観察研究における因果リスクの (バックドア基準やフロントドア基準 [Pearl, 2009] のような) 識別条件やパイロットスタディなどの別の調査により識別できていればよい. 例えば,  $p(x, y, z, w)$  に関して,  $\{Z, W\}$  が  $(X, Y)$  についてバックドア基準を満たすとき,  $p(Y_x = y | z)$  は一つの観測分布  $p(x, y, z, w)$  から識別可能であり,

$$p(Y_x = y | z) = \sum_w p(y | x, z, w)p(w | z) \quad (3.7)$$

で与えられる. 条件 (3) はすべての潜在結果変数タイプが存在し, 単調性 (即ち,  $p(u_3) = 0$ ) が成り立たないことを意味する. 単調性の仮定が成り立たないとき, PN, PS, PNS といった原因の確率は共変量  $Z$  なしでは識別可能ではない [Tian and Pearl, 2000]. 定理 3.1 は, 潜在結果変数と関連する一つの共変量を観測し, 因果リスクが手に入るとき, 潜在結果変数の同時確率が構築できることを示している. このことは, 単調性の仮定なしに潜在結果変数の同時確率を復元できることを意味する.

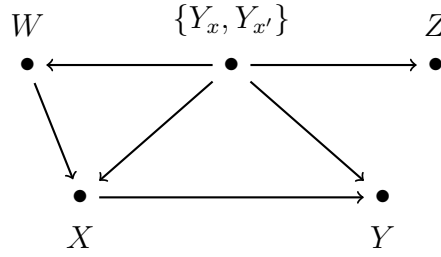


図 3.3: 3.2.2 節のグラフ表現

### 3.2.2 Case 2: 因果リスクなしの二つの共変量に基づいた識別

ここでは因果リスクについての事前知識を用いることなく、一定の条件の下で、潜在結果変数の同時確率は共変量のみを用いて評価できることを示す。これを見るために、図 3.2b に観測変数  $W$  を追加して得られる図 3.3 の因果ダイアグラムについて考える。  $X, Y, Z, W$  の正の同時確率について考えるとき、  $Z, W$  が取る値  $z, w$  に対して、一貫性から、

$$\begin{aligned}
 p(x_1, y_1, z, w, u_i) &= p(x_1, z, w, u_i), & p(x_1, y_0, z, w, u_i) &= 0, & i &= 1, 2, \\
 p(x_1, y_0, z, w, u_i) &= p(x_1, z, w, u_i), & p(x_1, y_1, z, w, u_i) &= 0, & i &= 3, 4, \\
 p(x_0, y_1, z, w, u_i) &= p(x_0, z, w, u_i), & p(x_0, y_0, z, w, u_i) &= 0, & i &= 1, 3, \\
 p(x_0, y_0, z, w, u_i) &= p(x_0, z, w, u_i), & p(x_0, y_1, z, w, u_i) &= 0, & i &= 2, 4,
 \end{aligned}$$

を得る。よって、例えば、  $p(y_1, z, w | x_1)$  と  $p(z, w | x_1)$  は

$$\begin{aligned}
 p(y_1, z, w | x_1) &= \sum_{i=1}^2 p(w | x_1, u_i) p(z | u_i) p(u_i | x_1), \\
 p(z, w | x_1) &= \sum_{i=1}^4 p(w | x_1, u_i) p(z | u_i) p(u_i | x_1).
 \end{aligned}$$

と書ける.  $Z$  は  $k \geq 4$  個の値  $z_1, z_2, \dots, z_k$  を,  $W$  は  $l \geq 4$  個の値  $w_1, w_2, \dots, w_l$  を取るとき,

$$P_{x_1} = \begin{pmatrix} 1 & p(z_1 | x_1) & p(z_2 | x_1) & p(z_3 | x_1) \\ p(w_1 | x_1) & p(z_1, w_1 | x_1) & p(z_2, w_1 | x_1) & p(z_3, w_1 | x_1) \\ p(w_2 | x_1) & p(z_1, w_2 | x_1) & p(z_2, w_2 | x_1) & p(z_3, w_2 | x_1) \\ p(w_3 | x_1) & p(z_1, w_3 | x_1) & p(z_2, w_3 | x_1) & p(z_3, w_3 | x_1) \end{pmatrix}, \quad (3.8)$$

$$P_{x_0} = \begin{pmatrix} 1 & p(z_1 | x_0) & p(z_2 | x_0) & p(z_3 | x_0) \\ p(w_1 | x_0) & p(z_1, w_1 | x_0) & p(z_2, w_1 | x_0) & p(z_3, w_1 | x_0) \\ p(w_2 | x_0) & p(z_1, w_2 | x_0) & p(z_2, w_2 | x_0) & p(z_3, w_2 | x_0) \\ p(w_3 | x_0) & p(z_1, w_3 | x_0) & p(z_2, w_3 | x_0) & p(z_3, w_3 | x_0) \end{pmatrix}, \quad (3.9)$$

と置くと, 次の定理を得ることができる:

**定理 3.2.**  $Z$  は  $k \geq 4$  個の値  $z_1, z_2, \dots, z_k$  を,  $W$  は  $l \geq 4$  個の値  $w_1, w_2, \dots, w_l$  を取るとき, 次の条件が満たされれば,  $y, y' \in \{y_1, y_0\}$ ,  $x \in \{x_1, x_0\}$  に対して,  $p(Y_{x_1} = y, Y_{x_0} = y')$  と  $p(Y_{x_1} = y, Y_{x_0} = y' | x)$  は識別可能である.

(1)  $z \in \{z_1, z_2, \dots, z_k\}$  と  $w \in \{w_1, w_2, \dots, w_l\}$  に対して,  $p(x_1, y_1, z, w)$ ,  $p(x_1, y_0, z, w)$ ,  $p(x_0, y_1, z, w)$ ,  $p(x_0, y_0, z, w)$  が得られる.

(2)  $W \perp\!\!\!\perp Z | \{X, U\}$  と  $X \perp\!\!\!\perp Z | U$  がともに成り立つ.

(3)  $P_{x_1}$  と  $P_{x_0}$  がともに可逆である.

**証明.** 定理 3.2 の条件 (1) と (2) から, 一致性により,  $z \in \{z_1, z_2, \dots, z_k\}$  と  $w \in \{w_1, w_2, \dots, w_l\}$  に対して,

$$p(y_1, z, w | x_1) = \sum_{i=1,2} p(w | x_1, u_i) p(z | u_i) p(u_i | x_1),$$

$$p(y_1, z, w | x_0) = \sum_{i=1,3} p(w | x_0, u_i) p(z | u_i) p(u_i | x_0),$$



$$\begin{aligned}
p(y_0, z, w | x_1) &= \sum_{i=3,4} p(w | x_1, u_i) p(z | u_i) p(u_i | x_1), \\
p(y_0, z, w | x_0) &= \sum_{i=2,4} p(w | x_0, u_i) p(z | u_i) p(u_i | x_0), \\
p(z, w | x_1) &= \sum_{i=1}^4 p(w | x_1, u_i) p(z | u_i) p(u_i | x_1), \\
p(z, w | x_0) &= \sum_{i=1}^4 p(w | x_0, u_i) p(z | u_i) p(u_i | x_0), \\
p(y_1 | x_1) &= \sum_{i=1,2} p(u_i | x_1), \quad p(y_1 | x_0) = \sum_{i=1,3} p(u_i | x_0)
\end{aligned}$$

を得る.  $x \in \{x_1, x_0\}$  に対して,

$$\begin{aligned}
P_x &= \begin{pmatrix} 1 & p(z_1 | x) & p(z_2 | x) & p(z_3 | x) \\ p(w_1 | x) & p(z_1, w_1 | x) & p(z_2, w_1 | x) & p(z_3, w_1 | x) \\ p(w_2 | x) & p(z_1, w_2 | x) & p(z_2, w_2 | x) & p(z_3, w_2 | x) \\ p(w_3 | x) & p(z_1, w_3 | x) & p(z_2, w_3 | x) & p(z_3, w_3 | x) \end{pmatrix}, \\
Q_x &= \begin{pmatrix} p(y_1 | x) & p(y_1, z_1 | x) & p(y_1, z_2 | x) & p(y_1, z_3 | x) \\ p(y_1, w_1 | x) & p(y_1, z_1, w_1 | x) & p(y_1, z_2, w_1 | x) & p(y_1, z_3, w_1 | x) \\ p(y_1, w_2 | x) & p(y_1, z_1, w_2 | x) & p(y_1, z_2, w_2 | x) & p(y_1, z_3, w_2 | x) \\ p(y_1, w_3 | x) & p(y_1, z_1, w_3 | x) & p(y_1, z_2, w_3 | x) & p(y_1, z_3, w_3 | x) \end{pmatrix}, \\
R_x &= \begin{pmatrix} 1 & p(w_1 | x, u_1) & p(w_2 | x, u_1) & p(w_3 | x, u_1) \\ 1 & p(w_1 | x, u_2) & p(w_2 | x, u_2) & p(w_3 | x, u_2) \\ 1 & p(w_1 | x, u_3) & p(w_2 | x, u_3) & p(w_3 | x, u_3) \\ 1 & p(w_1 | x, u_4) & p(w_2 | x, u_4) & p(w_3 | x, u_4) \end{pmatrix}, \\
\Delta_x &= \begin{pmatrix} p(u_1 | x) & 0 & 0 & 0 \\ 0 & p(u_2 | x) & 0 & 0 \\ 0 & 0 & p(u_3 | x) & 0 \\ 0 & 0 & 0 & p(u_4 | x) \end{pmatrix},
\end{aligned}$$

$$M_{x_1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad M_{x_0} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$S = \begin{pmatrix} 1 & p(z_1|u_1) & p(z_2|u_1) & p(z_3|u_1) \\ 1 & p(z_1|u_2) & p(z_2|u_2) & p(z_3|u_2) \\ 1 & p(z_1|u_3) & p(z_2|u_3) & p(z_3|u_3) \\ 1 & p(z_1|u_4) & p(z_2|u_4) & p(z_3|u_4) \end{pmatrix},$$

と置くと,  $x \in \{x_1, x_0\}$  に対して,

$$P_x = R_x^\top \Delta_x S, \quad Q_x = R_x^\top \Delta_x M_x S \quad (3.10)$$

を得る. 定理 3.2 の条件 6 から,  $P_{x_1}$  と  $P_{x_0}$  は可逆なので,  $x \in \{x_1, x_0\}$  に対して,

$$P_x^{-1} Q_x = S^{-1} M_x S \quad (3.11)$$

となり, これは固有値として 0 と 1 を持つ.  $E_{x_1}$  と  $E_{x_0}$  を  $4 \times 4$  行列

$$E_{x_1} = \begin{pmatrix} e_{11} & e_{12} & 0 & 0 \\ e_{21} & e_{22} & 0 & 0 \\ 0 & 0 & e_{33} & e_{34} \\ 0 & 0 & e_{43} & e_{44} \end{pmatrix}, \quad E_{x_0} = \begin{pmatrix} e'_{11} & 0 & e'_{13} & 0 \\ 0 & e'_{22} & 0 & e'_{24} \\ e'_{31} & 0 & e'_{33} & 0 \\ 0 & e'_{42} & 0 & e'_{44} \end{pmatrix}$$

と置くと,  $P_{x_1}^{-1} Q_{x_1}$  と  $P_{x_0}^{-1} Q_{x_0}$  の固有ベクトルからなる行列は, それぞれ  $A_{x_1} = S^{-1} E_{x_1}$  と  $A_{x_0} = S^{-1} E_{x_0}$  のように書くことができる. よって,

$$A_{x_1} E_{x_1}^{-1} = S^{-1} = A_{x_0} E_{x_0}^{-1}$$

となる.  $A_{x_1}^{-1} = (a^{ij})$  かつ  $A_{x_1}^{-1} A_{x_0} = (b^{ij})$  と置き,  $E_{x_1} A_{x_1}^{-1} A_{x_0} = E_{x_0}$  と  $E_{x_1} A_{x_1}^{-1} = S$

の第1列を並べることで,

$$\begin{aligned} \begin{pmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{pmatrix} \begin{pmatrix} a^{11} & b^{11} & b^{12} \\ a^{21} & b^{21} & b^{22} \end{pmatrix} &= \begin{pmatrix} 1 & e'_{11} & 0 \\ 1 & 0 & e'_{22} \end{pmatrix}, \\ \begin{pmatrix} e_{33} & e_{34} \\ e_{43} & e_{44} \end{pmatrix} \begin{pmatrix} a^{31} & b^{31} & b^{32} \\ a^{41} & b^{41} & b^{42} \end{pmatrix} &= \begin{pmatrix} 1 & e'_{31} & 0 \\ 1 & 0 & e'_{42} \end{pmatrix} \end{aligned} \quad (3.12)$$

を得る. 式 (3.12) から,

$$\begin{aligned} \begin{pmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{pmatrix} &= \begin{pmatrix} 1 & e'_{11} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a^{11} & b^{11} \\ a^{21} & b^{21} \end{pmatrix}^{-1}, \\ \begin{pmatrix} e_{33} & e_{34} \\ e_{43} & e_{44} \end{pmatrix} &= \begin{pmatrix} 1 & e'_{31} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a^{31} & b^{31} \\ a^{41} & b^{41} \end{pmatrix}^{-1}. \end{aligned} \quad (3.13)$$

の各第2行を見ることで,  $e_{21}, e_{22}, e_{43}, e_{44}$  は識別可能であることがわかる. 同様に,

$$\begin{aligned} \begin{pmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 1 & e'_{22} \end{pmatrix} \begin{pmatrix} a^{11} & b^{12} \\ a^{21} & b^{22} \end{pmatrix}^{-1}, \\ \begin{pmatrix} e_{33} & e_{34} \\ e_{43} & e_{44} \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 1 & e'_{42} \end{pmatrix} \begin{pmatrix} a^{31} & b^{32} \\ a^{41} & b^{42} \end{pmatrix}^{-1}. \end{aligned} \quad (3.14)$$

の各第1行を見ることで,  $e_{11}, e_{12}, e_{33}, e_{34}$  は識別可能であることがわかる. よって,  $E_{x_1}$  は識別可能であり,  $E_{x_1} A_{x_1}^{-1} A_{x_0} = E_{x_0}$  から  $E_{x_0}$  が,  $E_{x_1} A_{x_1}^{-1} = S$  から  $S$  が識別可能となる.

このとき,  $P_{x_1}$  と  $S$  はともに識別可能であり,  $R_{x_1}^\top \Delta_{x_1}$  の第1行は  $(p(u_1 | x_1), p(u_2 | x_1), p(u_3 | x_1), p(u_4 | x_1))$  と表せるので,  $P_{x_1} S^{-1}$  の第1行と  $R_{x_1}^\top \Delta_{x_1}$  の第1行を比較することで,  $p(u_1 | x_1), p(u_2 | x_1), p(u_3 | x_1), p(u_4 | x_1)$  は識別可能となる. 同様に,  $P_{x_0} S^{-1}$  の第1行と  $R_{x_0}^\top \Delta_{x_0}$  の第1行を比較することで,  $p(u_1 | x_0), p(u_2 | x_0), p(u_3 | x_0), p(u_4 | x_0)$  も識別可能となる. よって,  $i = 1, 2, 3, 4$  に対して,

$$p(u_i) = p(u_i | x_1)p(x_1) + p(u_i | x_0)p(x_0),$$

も識別可能である. □

未観測変数からなる確率は Kuroki and Pearl [2014] では完全に識別されないが、定理 3.2 は潜在結果変数の同時確率  $p(u_1), p(u_2), p(u_3), p(u_4)$  が識別可能であり、因果リスクも識別可能であることを示す。例えば、

$$p(Y_{x_1} = y_1) = p(u_1) + p(u_2), \quad p(Y_{x_0} = y_1) = p(u_1) + p(u_3) \quad (3.15)$$

であるから PN, PS, PNS も識別可能であることがわかる。

### 3.3 推定

潜在結果変数の同時確率が提案した定理 3.1 もしくは定理 3.2 の条件を通して識別可能であるとき、これらの定理の証明からわかるように、潜在結果変数の同時確率の推定問題は通常、特異モデルの推定問題に帰着されるため、最尤法のような標準的な統計的推定手法で評価することは難しい。そこでここでは、拡張 Lagrange 法 [Bertsekas, 1996] に基づいた新たな推定方法を提案する。

#### 3.3.1 Case 1 における推定

$\{(X_i, Y_i, Z_i)\}_{i=1}^n$  は図 3.2 におけるデータ生成過程で生成した観測値であるとする。さらに  $\{Z, W\}$  が  $(X, Y)$  に関するバックドア基準を満たすような  $W$  を観測し、観測値を  $\{W_i\}_{i=1}^n$  とする。  $x \in \{x_1, x_0\}$ ,  $y \in \{y_1, y_0\}$ ,  $z \in \{z_1, z_2, \dots, z_k\}$  に対して、  $p(z|x)$ ,  $p(y|x)$ ,  $p(w|z)$ ,  $p(y, z|x)$ ,  $p(y|x, z, w)$  の最尤推定量をそれぞれ  $\hat{p}(z|x)$ ,  $\hat{p}(y|x)$ ,  $\hat{p}(w|z)$ ,  $\hat{p}(y, z|x)$ ,  $\hat{p}(y|x, z, w)$  とする。  $p(Y_x = y|z)$  は識別可能であり、

$$\hat{p}(Y_x = y|z) = \sum_w \hat{p}(y|x, z, w)\hat{p}(w|z), \quad (3.16)$$

で与えられるので,  $p(y_x), p(y_x, z)$  も識別可能である. これらの推定量を  $\hat{p}(y_x), \hat{p}(y_x, z)$  と表す. このとき,  $P$  と  $Q$  のプラグイン推定量を

$$\hat{P} = \begin{pmatrix} 1 & \hat{p}(z_1) & \hat{p}(z_2) & \hat{p}(z_3) \\ \hat{p}(x_1, y_1) & \hat{p}(x_1, y_1, z_1) & \hat{p}(x_1, y_1, z_2) & \hat{p}(x_1, y_1, z_3) \\ \hat{p}(x_1, y_0) & \hat{p}(x_1, y_0, z_1) & \hat{p}(x_1, y_0, z_2) & \hat{p}(x_1, y_0, z_3) \\ \hat{p}(x_0, y_1) & \hat{p}(x_0, y_1, z_1) & \hat{p}(x_0, y_1, z_2) & \hat{p}(x_0, y_1, z_3) \end{pmatrix}, \quad (3.17)$$

$$\hat{Q} = \begin{pmatrix} 1 & \hat{p}(z_1) & \hat{p}(z_2) & \hat{p}(z_3) \\ \hat{p}(y_{x_1}) & \hat{p}(y_{x_1}, z_1) & \hat{p}(y_{x_1}, z_2) & \hat{p}(y_{x_1}, z_3) \\ \hat{p}(y_{x_0}) & \hat{p}(y_{x_0}, z_1) & \hat{p}(y_{x_0}, z_2) & \hat{p}(y_{x_0}, z_3) \end{pmatrix} \quad (3.18)$$

とする. 定理 3.1 の証明から,  $P$  と  $Q$  が与えられたとき, 識別可能な行列  $R$  は

$$QP^{-1}R^\top = M$$

を満たす. このことは,  $R$  が次の最小化問題の解であることを意味する:

$$\underset{\Theta \in \mathcal{T}}{\text{minimize}} \quad \frac{1}{2} \|QP^{-1}\Theta^\top - M\|_F^2 \quad (3.19)$$

$$\text{subject to } 0 \leq (\Theta^\top)^{-1}P\mathbf{e}_1 \leq 1, \quad \mathbf{1}^\top(\Theta^\top)^{-1}P\mathbf{e}_1 = 1, \quad \theta_{12} + \theta_{14} = \theta_{33} + \theta_{34} = 1. \quad (3.20)$$

ここで,  $\mathbf{e}_1 = (1, 0, 0, 0)^\top$ ,  $\mathbf{e}_2 = (0, 1, 0, 0)^\top$ ,  $\mathbf{e}_3 = (0, 0, 1, 0)^\top$ ,  $\mathbf{e}_4 = (0, 0, 0, 1)^\top$ ,  $\mathbf{1} = (1, 1, 1, 1)^\top$ ,

$$\mathcal{T} := \left\{ \begin{pmatrix} 1 & \theta_{12} & 0 & \theta_{14} \\ 1 & \theta_{22} & 0 & 0 \\ 1 & 0 & \theta_{33} & \theta_{34} \\ 1 & 0 & \theta_{43} & 0 \end{pmatrix} \in GL_4(\mathbb{R}) : 0 \leq \theta_{12}, \theta_{22}, \theta_{33}, \theta_{43}, \theta_{14}, \theta_{34} \leq 1 \right\}$$

であり,  $GL_4(\mathbb{R})$  は  $\mathbb{R}$  の成分を持つ可逆な  $4 \times 4$  行列からなる一般線形群である. また, 式 (3.20) の不等号は成分ごとの不等号を表し, 式 (3.20) の第 2 式は  $(\Theta^\top)^{-1}P$  の第 1 列は  $(p(u_1), \dots, p(u_4))$  と一致するような条件である.

そこで、 $P$ と $Q$ をそれぞれ $\hat{P}$ と $\hat{Q}$ に置き換えることで、次の最小化問題の解として $R$ を推定することを提案する:

$$\underset{\Theta \in \mathcal{T}}{\text{minimize}} \quad \frac{1}{2} \|\hat{Q}\hat{P}^{-1}\Theta^\top - M\|_F^2 \quad (3.21)$$

$$\text{subject to } 0 \leq (\Theta^\top)^{-1}\hat{P}\mathbf{e}_1 \leq 1, \quad \mathbf{1}^\top(\Theta^\top)^{-1}\hat{P}\mathbf{e}_1 = 1, \quad \theta_{12} + \theta_{14} = \theta_{33} + \theta_{34} = 1. \quad (3.22)$$

この最小化問題を解くために Bertsekas [1996] に従って、拡張 Lagrange 関数を

$$\begin{aligned} L(\Theta; \hat{P}, \hat{Q}, \mu, \mu_1, \mu_2, \boldsymbol{\lambda}) &= \frac{1}{2} \|\hat{Q}\hat{P}^{-1}\Theta^\top - M\|_F^2 + \mu \left( \mathbf{1}^\top(\Theta^\top)^{-1}\hat{P}\mathbf{e}_1 - 1 \right) + \frac{\rho}{2} \left( \mathbf{1}^\top(\Theta^\top)^{-1}\hat{P}\mathbf{e}_1 - 1 \right)^2 \\ &\quad + \mu_1 (\theta_{12} + \theta_{14} - 1) + \frac{\rho}{2} (\theta_{12} + \theta_{14} - 1)^2 + \mu_2 (\theta_{33} + \theta_{34} - 1) + \frac{\rho}{2} (\theta_{33} + \theta_{34} - 1)^2 \\ &\quad + \sum_{j=1}^4 f_j(\Theta, \boldsymbol{\lambda}, \rho) \end{aligned} \quad (3.23)$$

と表す。ここで、 $\mu$ ,  $\mu_1$ ,  $\mu_2$ , and  $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_4)^\top$  は Lagrange 乗数であり,

$$\begin{aligned} f_j(\Theta, \boldsymbol{\lambda}, \rho) &= \min_{0 \leq \mathbf{e}_j^\top(\Theta^\top)^{-1}\hat{P}\mathbf{e}_1 - v_j \leq 1} \left( \lambda_j v_j + \frac{\rho}{2} v_j^2 \right) \\ &= \begin{cases} \lambda_j (\mathbf{e}_j^\top(\Theta^\top)^{-1}\hat{P}\mathbf{e}_1 - 1) + \frac{\rho}{2} \left| \mathbf{e}_j^\top(\Theta^\top)^{-1}\hat{P}\mathbf{e}_1 - 1 \right|^2 & \text{if } \lambda_j + \rho(\mathbf{e}_j^\top(\Theta^\top)^{-1}\hat{P}\mathbf{e}_1 - 1) > 1, \\ \lambda_j \mathbf{e}_j^\top(\Theta^\top)^{-1}\hat{P}\mathbf{e}_1 + \frac{\rho}{2} \left| \mathbf{e}_j^\top(\Theta^\top)^{-1}\hat{P}\mathbf{e}_1 \right|^2 & \text{if } \lambda_j + \rho \mathbf{e}_j^\top(\Theta^\top)^{-1}\hat{P}\mathbf{e}_1 < 1, \\ -\frac{\lambda_j^2}{2\rho} & \text{otherwise} \end{cases} \end{aligned} \quad (3.24)$$

である。Lagrange 乗数に対する反復計算は

$$\mu^{(t+1)} = \mu^{(t)} + \rho(\mathbf{1}^\top(\Theta^\top)^{-1}\hat{P}\mathbf{e}_1 - 1), \quad (3.25)$$

$$\mu_1^{(t+1)} = \mu_1^{(t)} + \rho(\theta_{12} + \theta_{14} - 1), \quad (3.26)$$

$$\mu_2^{(t+1)} = \mu_2^{(t)} + \rho(\theta_{33} + \theta_{34} - 1), \quad (3.27)$$

$$\lambda_j^{(t+1)} = \begin{cases} \lambda_j^{(t)} + \rho(\mathbf{e}_j^\top (\Theta^\top)^{-1} \widehat{P} \mathbf{e}_1 - 1) & \text{if } \lambda_j^{(t)} + \rho(\mathbf{e}_j^\top (\Theta^\top)^{-1} \widehat{P} \mathbf{e}_1 - 1) > 1, \\ \lambda_j^{(t)} + \rho \mathbf{e}_j^\top (\Theta^\top)^{-1} \widehat{P} \mathbf{e}_1 & \text{if } \lambda_j^{(t)} + \rho \mathbf{e}_j^\top (\Theta^\top)^{-1} \widehat{P} \mathbf{e}_1 < 1, \\ 0 & \text{otherwise} \end{cases} \quad (3.28)$$

で与えられる。このとき、 $R$ の推定量は次の推定方程式の解 $\widehat{\Theta}$ で与えられる:

$$\frac{\partial}{\partial \Theta} L(\Theta; \widehat{P}, \widehat{Q}, \mu, \mu_1, \mu_2, \boldsymbol{\lambda}) = 0. \quad (3.29)$$

Algorithm 1 は勾配降下法による拡張 Lagrange 法と更新則に基づいた最適化問題の解を与えるアルゴリズムである。  $\alpha$  は  $t$  番目の反復における一定のステップサイズ、  $T$  は反復回数、  $\Theta^{(0)}$  は初期値である。推定量  $\widehat{R}$  を最適化問題 (3.21) の解として得れば、  $\mathbf{u} = (p(u_1), \dots, p(u_4))^\top$  の推定量は

$$\widehat{\mathbf{u}} = (\widehat{R}^\top)^{-1} \widehat{P} \mathbf{e}_1$$

で与えられる。例えば、PNS は  $\mathbf{u}$  の第 2 成分であるので、PNS を  $\widehat{\mathbf{u}}$  の第 2 成分として推定でき、同様に、因果リスク差を  $\widehat{\mathbf{u}}$  の第 2 成分と第 3 成分の差として推定できる。

### 3.3.2 Case 2 における推定

$\{(X_i, Y_i, Z_i, W_i)\}_{i=1}^n$  は図 3.3 におけるデータ生成過程からの観測値であるとする。  $x \in \{x_1, x_0\}$ ,  $y \in \{y_1, y_0\}$ ,  $z \in \{z_1, z_2, \dots, z_k\}$ ,  $w \in \{w_1, w_2, \dots, w_l\}$  に対して、  $p(z|x)$ ,  $p(w|x)$ ,  $p(z, w|x)$ ,  $p(y|x)$ ,  $p(y, z|x)$ ,  $p(y, w|x)$ ,  $p(y, z, w|x)$  の最尤推定量をそれぞれ  $\widehat{p}(z|x)$ ,  $\widehat{p}(w|x)$ ,  $\widehat{p}(z, w|x)$ ,  $\widehat{p}(y|x)$ ,  $\widehat{p}(y, z|x)$ ,  $\widehat{p}(y, w|x)$ ,  $\widehat{p}(y, z, w|x)$  とする。このとき、  $x \in \{x_1, x_0\}$  に対して、  $P_x$  と  $Q_x$  のプラグイン推定量を

$$\widehat{P}_x = \begin{pmatrix} 1 & \widehat{p}(z_1|x) & \widehat{p}(z_2|x) & \widehat{p}(z_3|x) \\ \widehat{p}(w_1|x) & \widehat{p}(z_1, w_1|x) & \widehat{p}(z_2, w_1|x) & \widehat{p}(z_3, w_1|x) \\ \widehat{p}(w_2|x) & \widehat{p}(z_1, w_2|x) & \widehat{p}(z_2, w_2|x) & \widehat{p}(z_3, w_2|x) \\ \widehat{p}(w_3|x) & \widehat{p}(z_1, w_3|x) & \widehat{p}(z_2, w_3|x) & \widehat{p}(z_3, w_3|x) \end{pmatrix}, \quad (3.30)$$

---

**Algorithm 1** Case 1における  $\mathbf{u} = (p(u_1), \dots, p(u_4))^\top$  の推定.

---

**Input:**  $\{(X_i, Y_i, Z_i, W_i)\}_{i=1}^n, \alpha, T, \rho, \Theta^{(0)}$

**Output:**  $\hat{\mathbf{u}}$

- 1:  $\mu^{(0)} \leftarrow 0, \mu_1^{(0)} \leftarrow 0, \mu_2^{(0)} \leftarrow 0, \boldsymbol{\lambda}^{(0)} \leftarrow 0$
  - 2: 式 (3.16) を用いて  $\hat{p}(y_x), \hat{p}(y_x, z_1), \hat{p}(y_x, z_2), \hat{p}(y_x, z_3)$  を計算
  - 3: 観測データ  $\{(X_i, Y_i, Z_i)\}_{i=1}^n$  を用いて  $\hat{P}$  と  $\hat{Q}$  を計算
  - 4: **for**  $t = 0, 1, \dots, T - 1$  **do**
  - 5:  $\theta_{12}^{(t)} \leftarrow \max\{\min\{\theta_{12}^{(t)}, 1\}, 0\}, \theta_{22}^{(t)} \leftarrow \max\{\min\{\theta_{22}^{(t)}, 1\}, 0\},$   
 $\theta_{33}^{(t)} \leftarrow \max\{\min\{\theta_{33}^{(t)}, 1\}, 0\}, \theta_{43}^{(t)} \leftarrow \max\{\min\{\theta_{43}^{(t)}, 1\}, 0\},$   
 $\theta_{14}^{(t)} \leftarrow \max\{\min\{\theta_{14}^{(t)}, 1\}, 0\}, \theta_{34}^{(t)} \leftarrow \max\{\min\{\theta_{34}^{(t)}, 1\}, 0\}$
  - 6:  $\Theta^{(t+1)} \leftarrow \Theta^{(t)} - \alpha \frac{\partial}{\partial \Theta} L(\Theta; \hat{P}, \hat{Q}, \mu^{(t)}, \boldsymbol{\lambda}^{(t)})|_{\Theta=\Theta^{(t)}}$
  - 7:  $\mu^{(t+1)} \leftarrow \mu^{(t)} + \rho(\mathbf{1}^\top (\Theta^\top)^{-1} \hat{P} \mathbf{e}_1 - 1)$
  - 8:  $\mu_1^{(t+1)} \leftarrow \mu_1^{(t)} + \rho(\hat{\theta}_{12} + \hat{\theta}_{14} - 1)$
  - 9:  $\mu_2^{(t+1)} \leftarrow \mu_2^{(t)} + \rho(\hat{\theta}_{33} + \hat{\theta}_{34} - 1)$
  - 10: **for**  $j = 1, 2, 3, 4$  **do**
  - 11: **if**  $\lambda_j^{(t)} + \rho(\mathbf{e}_j^\top (\Theta^{(t+1)\top})^{-1} \hat{P} \mathbf{e}_1 - 1) > 1$  **then**
  - 12:  $\lambda_j^{(t+1)} \leftarrow \lambda_j^{(t)} + \rho(\mathbf{e}_j^\top (\Theta^{(t+1)\top})^{-1} \hat{P} \mathbf{e}_1 - 1)$
  - 13: **else if**  $\lambda_j^{(t)} + \rho \mathbf{e}_j^\top (\Theta^{(t+1)\top})^{-1} \hat{P} \mathbf{e}_1 < 1$  **then**
  - 14:  $\lambda_{x_{1,j}}^{(t+1)} \leftarrow \lambda_j^{(t)} + \rho \mathbf{e}_j^\top (\Theta^{(t+1)\top})^{-1} \hat{P} \mathbf{e}_1$
  - 15: **else**
  - 16:  $\lambda_j^{(t+1)} \leftarrow 0; \lambda_j^{(t+1)} \leftarrow 0$
  - 17: **end if**
  - 18: **end for**
  - 19: **end for**
  - 20:  $\hat{R} \leftarrow \Theta^{(T)}$
  - 21:  $\hat{\mathbf{u}} \leftarrow (\hat{R}^\top)^{-1} \hat{P} \mathbf{e}_1$
-



$$\widehat{Q}_x = \begin{pmatrix} \widehat{p}(y|x) & \widehat{p}(y, z_1|x) & \widehat{p}(y, z_2|x) & \widehat{p}(y, z_3|x) \\ \widehat{p}(y, w_1|x) & \widehat{p}(y, z_1, w_1|x) & \widehat{p}(y, z_2, w_1|x) & \widehat{p}(y, z_3, w_1|x) \\ \widehat{p}(y, w_2|x) & \widehat{p}(y, z_1, w_2|x) & \widehat{p}(y, z_2, w_2|x) & \widehat{p}(y, z_3, w_2|x) \\ \widehat{p}(y, w_3|x) & \widehat{p}(y, z_1, w_3|x) & \widehat{p}(y, z_2, w_3|x) & \widehat{p}(y, z_3, w_3|x) \end{pmatrix} \quad (3.31)$$

と表す. 定理 3.2 の証明から,  $x \in \{x_1, x_0\}$  に対して  $P_x$  と  $Q_x$  が与えられた下で, 識別可能な行列  $S$  は

$$P_x = R_x^\top \Delta_x S, \quad Q_x = R_x^\top \Delta_x M_x S \quad (3.32)$$

を満たすので,

$$SP_x^{-1}Q_x = M_x S$$

を得る. このことは  $S$  が最小化問題

$$\underset{\Theta \in \mathcal{T}}{\text{minimize}} \frac{1}{2} \|\Theta P_{x_1}^{-1} Q_{x_1} - M_{x_1} \Theta\|_F^2 + \frac{1}{2} \|\Theta P_{x_0}^{-1} Q_{x_0} - M_{x_0} \Theta\|_F^2 \quad (3.33)$$

$$\text{subject to } 0 \leq (P_{x_1} \Theta^{-1})^\top \mathbf{e}_1 \leq 1, \quad \mathbf{1}^\top (P_{x_1} \Theta^{-1})^\top \mathbf{e}_1 = 1, \quad (3.34)$$

$$0 \leq (P_{x_0} \Theta^{-1})^\top \mathbf{e}_1 \leq 1, \quad \mathbf{1}^\top (P_{x_0} \Theta^{-1})^\top \mathbf{e}_1 = 1, \quad (3.35)$$

の解であることを意味するので,  $P_x$  と  $Q_x$  を  $\widehat{P}_x$  と  $\widehat{Q}_x$  に置き換えた次の最小化問題の解として,  $S$  を推定することを提案する:

$$\underset{\Theta \in \mathcal{T}}{\text{minimize}} \frac{1}{2} \|\Theta \widehat{P}_{x_1}^{-1} \widehat{Q}_{x_1} - M_{x_1} \Theta\|_F^2 + \frac{1}{2} \|\Theta \widehat{P}_{x_0}^{-1} \widehat{Q}_{x_0} - M_{x_0} \Theta\|_F^2 \quad (3.36)$$

$$\text{subject to } 0 \leq (\widehat{P}_{x_1} \Theta^{-1})^\top \mathbf{e}_1 \leq 1, \quad \mathbf{1}^\top (\widehat{P}_{x_1} \Theta^{-1})^\top \mathbf{e}_1 = 1, \quad (3.37)$$

$$0 \leq (\widehat{P}_{x_0} \Theta^{-1})^\top \mathbf{e}_1 \leq 1, \quad \mathbf{1}^\top (\widehat{P}_{x_0} \Theta^{-1})^\top \mathbf{e}_1 = 1. \quad (3.38)$$

ここで

$$\mathcal{T} := \left\{ \Theta = (\theta_{ij}) \in GL_4(\mathbb{R}) : \theta_{i1} = 1, \sum_{j=2}^4 \theta_{ij} \leq 1 \text{ for } i = 1, \dots, 4 \text{ and } 0 \leq \Theta \leq 1 \right\}.$$

式 (3.34), (3.35) は,  $x \in \{x_0, x_1\}$  に対して  $P_x \Theta^{-1}$  の第 1 行が  $(p(u_1|x), p(u_2|x), p(u_3|x), p(u_4|x))$  と一致するような条件である.

Case 2 における推定と同様に、この最小化問題を解くために Bertsekas [1996] に従って、拡張 Lagrange 関数を

$$\begin{aligned}
& L(\Theta; \widehat{P}_{x_1}, \widehat{P}_{x_0}, \widehat{Q}_{x_1}, \widehat{Q}_{x_0}, \mu, \boldsymbol{\lambda}_{x_1}, \boldsymbol{\lambda}_{x_0}) \\
&= \sum_{x \in \{x_1, x_0\}} \frac{1}{2} \|\Theta \widehat{P}_x^{-1} \widehat{Q}_x - M_x \Theta\|_F^2 + \sum_{x \in \{x_1, x_0\}} \mu \left( \mathbf{1}^\top (\widehat{P}_x \Theta^{-1})^\top \mathbf{e}_1 - 1 \right) \\
&+ \sum_{x \in \{x_1, x_0\}} \frac{\rho}{2} \left( \mathbf{1}^\top (\widehat{P}_x \Theta^{-1})^\top \mathbf{e}_1 - 1 \right)^2 + \sum_{x \in \{x_1, x_0\}} \sum_{j=1}^4 f_{x,j}(\Theta, \boldsymbol{\lambda}_x, \rho) \quad (3.39)
\end{aligned}$$

とする。ここで、 $\mu$  と  $\boldsymbol{\lambda}_x = (\lambda_{x,1}, \lambda_{x,2}, \lambda_{x,3}, \lambda_{x,4})^\top$  は Lagrange 乗数であり、 $x \in \{x_1, x_0\}$  に対して

$$\begin{aligned}
f_{x,j}(\Theta, \boldsymbol{\lambda}_x, \rho) &= \min_{0 \leq \mathbf{e}_j^\top (\widehat{P}_x \Theta^{-1})^\top \mathbf{e}_1 - v_j \leq 1} \left( \lambda_{x,j} v_j + \frac{\rho}{2} v_j^2 \right) \\
&= \begin{cases} \lambda_{x,j} (\mathbf{e}_j^\top (\widehat{P}_x \Theta^{-1})^\top \mathbf{e}_1 - 1) + \frac{\rho}{2} \left| \mathbf{e}_j^\top (\widehat{P}_x \Theta^{-1})^\top \mathbf{e}_1 - 1 \right|^2 & \text{if } \lambda_{x,j} + \rho (\mathbf{e}_j^\top (\widehat{P}_x \Theta^{-1})^\top \mathbf{e}_1 - 1) > 1, \\ \lambda_{x,j} (\mathbf{e}_j^\top (\widehat{P}_x \Theta^{-1})^\top \mathbf{e}_1) + \frac{\rho}{2} \left| \mathbf{e}_j^\top (\widehat{P}_x \Theta^{-1})^\top \mathbf{e}_1 \right|^2 & \text{if } \lambda_{x,j} + \rho (\mathbf{e}_j^\top (\widehat{P}_x \Theta^{-1})^\top \mathbf{e}_1) < 1, \\ -\frac{\lambda_{x,j}^2}{2\rho} & \text{otherwise} \end{cases} \quad (3.40)
\end{aligned}$$

である。Lagrange 乗数の反復計算は

$$\begin{aligned}
\mu^{(t+1)} &= \mu^{(t)} + \rho \left( \mathbf{1}^\top (\widehat{P}_x \Theta^{-1})^\top \mathbf{e}_1 - 1 \right), \quad (3.41) \\
\lambda_{x,j}^{(t+1)} &= \begin{cases} \lambda_{x,j}^{(t)} + \rho (\mathbf{e}_j^\top (\widehat{P}_x \Theta^{-1})^\top \mathbf{e}_1 - 1) & \text{if } \lambda_{x,j}^{(t)} + \rho (\mathbf{e}_j^\top (\widehat{P}_x \Theta^{-1})^\top \mathbf{e}_1 - 1) > 1, \\ \lambda_{x,j}^{(t)} + \rho \mathbf{e}_j^\top (\widehat{P}_x \Theta^{-1})^\top \mathbf{e}_1 & \text{if } \lambda_{x,j}^{(t)} + \rho \mathbf{e}_j^\top (\widehat{P}_x \Theta^{-1})^\top \mathbf{e}_1 < 1, \\ 0 & \text{otherwise} \end{cases} \quad (3.42)
\end{aligned}$$

で与えられる。このとき、 $S$  の推定量の候補は次の推定方程式の解  $\widehat{\Theta}$  で与えられる:

$$\frac{\partial}{\partial \Theta} L(\Theta; \widehat{P}_{x_1}, \widehat{P}_{x_0}, \widehat{Q}_{x_1}, \widehat{Q}_{x_0}, \mu, \boldsymbol{\lambda}_{x_1}, \boldsymbol{\lambda}_{x_0}) = 0. \quad (3.43)$$

Algorithm 2 は拡張 Lagrange 法と勾配降下法による更新則に基づいて、最適化問題 3.43 の解を与えるアルゴリズムである。 $\alpha$  は  $t$  番目の反復における一定のステップサイズ、 $T$  は反復回数、 $\Theta^{(0)}$  は初期値である。直ちにわかるように、推定方程式 (3.43) の解  $\hat{\Theta}$  に対して、 $\Pi$  を置換行列としたとき、 $\Pi\hat{\Theta}$  も同じ推定方程式の解となる。したがって、ここでは最小の損失を達成するように行を置換した行列  $\Pi\Theta$  を  $S$  の推定量に採用することとする。推定量  $\hat{\Theta}$  を最適化問題 (3.36) の解として得れば、 $\mathbf{u} = (p(u_1), p(u_2), p(u_3), p(u_4))^\top$  の推定量は

$$\hat{\mathbf{u}} = \left( \frac{1}{n} \sum_{i=1}^n \mathbf{1}\{X_i = x_1\} \right) (\hat{P}_{x_1} \hat{\Theta}^{-1})^\top \mathbf{e}_1 + \left( \frac{1}{n} \sum_{i=1}^n \mathbf{1}\{X_i = x_0\} \right) (\hat{P}_{x_0} \hat{\Theta}^{-1})^\top \mathbf{e}_1$$

で与えられる。

### 3.4 数値実験

本節では、PNS  $p(u_2)$  と因果リスク差  $p(u_2) - p(u_3)$  を通して前節で提案した推定法の性質を説明するための数値実験を行う。簡単のため、 $X, Y, Z, W, U$  は離散変数であるとし、 $(X, Y, Z, W, U)$  の同時確率が表 3.1 に従って与えられるような図 3.3 で示した因果ダイアグラムについて考える。 $(X, Y, Z, W)$  は観測できるが  $U$  は観測できないような状況の下で、 $p(u_2)$  と  $p(u_2) - p(u_3)$  に対する提案した推定量  $\hat{p}(u_2)$  と  $\hat{p}(u_2) - \hat{p}(u_3)$  の性質をサンプルサイズ  $n = 100, 200, 1000, 5000$  の設定の数値実験で調べる。このとき、 $p(u_2)$  と  $p(u_2) - p(u_3)$  の真の値はそれぞれ  $1/4$  とゼロであり、 $\hat{p}(u_2)$  と  $\hat{p}(u_2) - \hat{p}(u_3)$  の平均はそれぞれ  $1/4$  とゼロに近くなることが期待される。表 4.2 と図 4.4 は与えられたサンプルサイズ  $n$  における 1000 回の複製に対する  $\hat{p}(u_2)$  と  $\hat{p}(u_2) - \hat{p}(u_3)$  の基本統計量と箱型図を示す。図 4.4 の水平線は  $p(u_2)$  と  $p(u_2) - p(u_3)$  の真の値を表す。表 4.2 より、 $\hat{p}(u_2)$  と  $\hat{p}(u_2) - \hat{p}(u_3)$  の標本平均は真値に近く、サンプルサイズが大きいほど標本標準偏差は小さくなる。したがって、提案する推定法は  $p(u_2)$  と  $p(u_2) - p(u_3)$  の一致推定値を与える様子が見て取れる。図 4.4 から、 $\hat{p}(u_2)$  と  $\hat{p}(u_2) - \hat{p}(u_3)$  の四分位範囲は狭く、サンプルサイズが大きくても真

---

**Algorithm 2** Case 2における  $\mathbf{u} = (p(u_1), p(u_2), p(u_3), p(u_4))^\top$  の推定.

---

**Input:**  $\{(X_i, Y_i, Z_i, W_i)\}_{i=1}^n, \alpha, T, \rho, \Theta^{(0)}$

**Output:**  $\hat{\mathbf{u}}$

- 1:  $\mu^{(0)} \leftarrow 0, \boldsymbol{\lambda}_{x_1}^{(0)} \leftarrow 0, \boldsymbol{\lambda}_{x_0}^{(0)} \leftarrow 0$
  - 2: 観測データ  $\{(X_i, Y_i, Z_i, W_i)\}_{i=1}^n$  を用いて  $\hat{P}_{x_1}, \hat{P}_{x_0}, \hat{Q}_{x_1}, \hat{Q}_{x_0}$  を計算
  - 3: **for**  $t = 0, 1, \dots, T - 1$  **do**
  - 4:   **for**  $i = 1, 2, 3, 4$  **do**
  - 5:     **for**  $j = 2, 3, 4$  **do**
  - 6:        $\theta_{ij}^{(t)} \leftarrow \max\{\max\{\min\{\theta_{ij}^{(t)}, 1\}, 0\}, 1 - \sum_{j=2}^{j-1} \theta_{ij}^{(t)}\}$
  - 7:     **end for**
  - 8:   **end for**
  - 9:    $\Theta^{(t+1)} \leftarrow \Theta^{(t)} - \alpha \frac{\partial}{\partial \Theta} L(\Theta; \hat{P}_{x_1}, \hat{P}_{x_0}, \hat{Q}_{x_1}, \hat{Q}_{x_0}, \mu^{(t)}, \boldsymbol{\lambda}_{x_1}^{(t)}, \boldsymbol{\lambda}_{x_0}^{(t)}) \Big|_{\Theta=\Theta^{(t)}}$
  - 10:    $\mu^{(t+1)} = \mu^{(t)} + \rho(\mathbf{1}^\top (\hat{P}_x(\Theta^{(t+1)})^{-1}))^\top \mathbf{e}_1 - 1)$
  - 11:   **for**  $j = 1, 2, 3, 4$  **do**
  - 12:     **if**  $\lambda_{x_1,j}^{(t)} + \rho(\mathbf{e}_j^\top (\hat{P}_{x_1}(\Theta^{(t+1)})^{-1}))^\top \mathbf{e}_1 - 1) > 1$  **then**
  - 13:        $\lambda_{x_1,j}^{(t+1)} \leftarrow \lambda_{x_1,j}^{(t)} + \rho(\mathbf{e}_j^\top (\hat{P}_{x_1}(\Theta^{(t+1)})^{-1}))^\top \mathbf{e}_1 - 1)$
  - 14:        $\lambda_{x_0,j}^{(t+1)} \leftarrow \lambda_{x_0,j}^{(t)} + \rho(\mathbf{e}_j^\top (\hat{P}_{x_0}(\Theta^{(t+1)})^{-1}))^\top \mathbf{e}_1 - 1)$
  - 15:     **else if**  $\lambda_{x,j}^{(t)} + \rho(\mathbf{e}_j^\top (\hat{P}_x(\Theta^{(t+1)})^{-1}))^\top \mathbf{e}_1 < 1$  **then**
  - 16:        $\lambda_{x_1,j}^{(t+1)} \leftarrow \lambda_{x_1,j}^{(t)} + \rho(\mathbf{e}_j^\top (\hat{P}_{x_1}(\Theta^{(t+1)})^{-1}))^\top \mathbf{e}_1; \lambda_{x_0,j}^{(t+1)} \leftarrow \lambda_{x_0,j}^{(t)} + \rho(\mathbf{e}_j^\top (\hat{P}_{x_0}(\Theta^{(t+1)})^{-1}))^\top \mathbf{e}_1$
  - 17:     **else**
  - 18:        $\lambda_{x_1,j}^{(t+1)} \leftarrow 0; \lambda_{x_0,j}^{(t+1)} \leftarrow 0$
  - 19:     **end if**
  - 20:   **end for**
  - 21: **end for**
  - 22:  $\Pi^* \leftarrow \arg \min_{\Pi \Theta} \left\{ \|(\Pi \Theta^{(T)}) \hat{P}_{x_1}^{-1} \hat{Q}_{x_1} - M_{x_1}(\Pi \Theta^{(T)})\|_F^2 + \|(\Pi \Theta^{(T)}) \hat{P}_{x_0}^{-1} \hat{Q}_{x_0} - M_{x_0}(\Pi \Theta^{(T)})\|_F^2 \right\}$
  - 23:  $\hat{\Theta} \leftarrow \Pi^* \Theta^{(T)}$
  - 24:  $\hat{\mathbf{u}} \leftarrow (\frac{1}{n} \sum_{i=1}^n \mathbf{1}\{X_i = x_1\}) (\hat{P}_{x_1} \hat{\Theta}^{-1})^\top \mathbf{e}_1 + (\frac{1}{n} \sum_{i=1}^n \mathbf{1}\{X_i = x_0\}) (\hat{P}_{x_0} \hat{\Theta}^{-1})^\top \mathbf{e}_1$
-

表 3.1: シミュレーションにおける条件付き確率表.

	(a) $p(Z U)$				(b) $p(W U)$				(c) $p(Y = 1 X, U)$	
	$Z = 1$	$Z = 2$	$Z = 3$	$Z = 4$	$W = 1$	$W = 2$	$W = 3$	$W = 4$	$X = 1$	$X = 0$
$U = 1$	7/10	1/10	1/10	1/10	7/10	1/10	1/10	1/10	1	1
$U = 2$	1/10	7/10	1/10	1/10	1/10	7/10	1/10	1/10	1	0
$U = 3$	1/10	1/10	7/10	1/10	1/10	1/10	7/10	1/10	0	1
$U = 4$	1/10	1/10	1/10	7/10	1/10	1/10	1/10	7/10	0	0

	(d) $p(X = 1 W, U)$				(e) $p(U)$
	$W = 1$	$W = 2$	$W = 3$	$W = 4$	
$U = 1$	21/46	18/43	18/43	18/43	1/4
$U = 2$	9/34	21/71	9/34	9/34	1/4
$U = 3$	9/34	9/34	21/71	9/34	1/4
$U = 4$	9/34	9/34	9/34	21/71	1/4

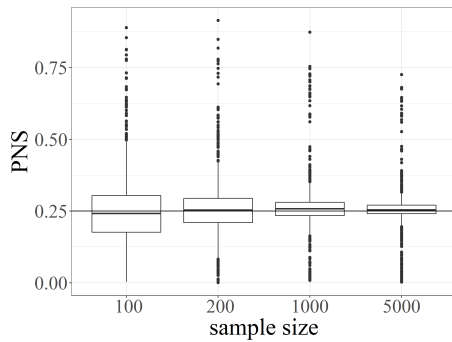
の値を含むことが見て取れる. 一方,  $\hat{p}(u_2) - \hat{p}(u_3)$  は対称分布であり漸近正規性が成り立つ様子がわかるが,  $\hat{p}(u_2)$  は成り立つ様子を見て取ることが難しいかもしれない. これは,  $p(u_2)$  の真値が比較的ゼロに近く, 図 4.4 は有限標本サイズに対する  $\hat{p}(u_2)$  のゼロ打ち切り分布として与えられるからである. また, 各サンプルサイズにおいて外れ値が多いことがわかるが, 観測データから定理 3.2 の条件 (3) が成り立たないような場合に発生することに注意する.

### 3.5 適用例

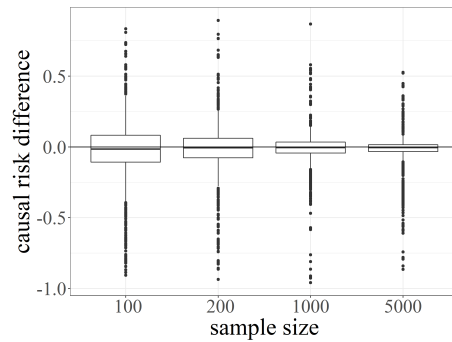
LaLonde [1986] によって報告され, Dehejia and Wahba [1999] によって再分析されたデータセットを通して, 第 3 部の結果を説明する. LaLonde [1986] の研究の目的は, 職業訓練プログラムである NSW (National Supported Work) の実証実験が訓

表 3.2: 提案手法に基づいた推定値の基本統計量.

	(a) $\hat{p}(u_2)$				(b) $\hat{p}(u_2) - \hat{p}(u_3)$			
	$n = 100$	$n = 200$	$n = 1000$	$n = 5000$	$n = 100$	$n = 200$	$n = 1000$	$n = 5000$
Minimum	0.003	0.001	0.008	0.002	-0.906	-0.936	-0.957	-0.864
1st Quantile	0.176	0.210	0.234	0.241	-0.107	-0.076	-0.043	-0.033
Median	0.242	0.253	0.258	0.254	-0.015	-0.006	-0.005	-0.005
Mean	0.251	0.259	0.264	0.255	-0.026	-0.017	-0.013	-0.025
3rd	0.305	0.294	0.281	0.271	0.083	0.061	0.034	0.016
Maximum	0.890	0.915	0.874	0.726	0.833	0.893	0.867	0.527
s.e.	0.126	0.103	0.082	0.072	0.228	0.189	0.139	0.129



(a) PNS に対する箱型図.



(b) 因果リスク差に対する箱型図.

図 3.4: 提案手法に基づいた推定値に対する箱型図.

練生の収入に及ぼす効果をフィールド実験において評価することである。LaLonde [1986] によれば、個人を治療群（参加）と対照群（不参加）に無作為に割り付けたが、その無作為割り当てはうまくいかなかったようである。ここで用いられるデータセットは Dehejia のウェブページ<sup>1</sup>から得られる。サンプルサイズは 445 で、興味のある変数は以下の通りである：

$X$ : 職業訓練プログラムに出席するかどうか ( $x_1$ : 参加;  $x_0$ : 不参加)

$Y$ : 1975 年と 1978 年の間の収入が増えたかどうか ( $y_1$ : 収入増;  $y_0$ : 収入減) ,

<sup>1</sup><https://users.nber.org/~rdehejia/nswdata2.html>

Z: 結婚しているかどうかと高校を卒業したかどうか ( $z_1$ : 1975年に非ゼロの収入かつ結婚している;  $z_2$ : 1975年に非ゼロの収入かつ未婚;  $z_3$ : 1975年に収入ゼロかつ結婚している;  $z_4$ : 1975年に収入ゼロかつ未婚)

W: 年齢区分 ( $w_1$ : 20歳未満;  $w_2$ : 20歳以上27歳未満;  $w_3$ : 27歳以上35歳未満;  $w_4$ : 35歳以上)

この状況の下で、本研究のデータ生成過程は図 3.3 で符号化されると仮定する。

このデータセットにおいて、 $P_{x_1}, P_{x_0}, Q_{x_1}, Q_{x_0}$  の標本推定値は

$$\hat{P}_{x_1} = \begin{pmatrix} 1.000 & 0.141 & 0.568 & 0.049 \\ 0.205 & 0.011 & 0.189 & 0.000 \\ 0.416 & 0.054 & 0.195 & 0.022 \\ 0.254 & 0.038 & 0.135 & 0.022 \end{pmatrix}, \quad \hat{P}_{x_0} = \begin{pmatrix} 1.000 & 0.115 & 0.719 & 0.038 \\ 0.262 & 0.008 & 0.242 & 0.000 \\ 0.404 & 0.050 & 0.288 & 0.008 \\ 0.242 & 0.054 & 0.112 & 0.023 \end{pmatrix},$$

$$\hat{Q}_{x_1} = \begin{pmatrix} 0.670 & 0.103 & 0.351 & 0.032 \\ 0.146 & 0.011 & 0.130 & 0.000 \\ 0.276 & 0.038 & 0.119 & 0.016 \\ 0.162 & 0.032 & 0.076 & 0.011 \end{pmatrix}, \quad \hat{Q}_{x_0} = \begin{pmatrix} 0.581 & 0.046 & 0.446 & 0.023 \\ 0.177 & 0.008 & 0.169 & 0.000 \\ 0.223 & 0.019 & 0.177 & 0.004 \\ 0.135 & 0.015 & 0.062 & 0.015 \end{pmatrix},$$

で与えられる。これらの式から、定理 3.2 の条件 (1) と (3) が成り立つことは合理的であると考えられる。このとき、定理 3.2 の条件 (2) が成り立つという仮定の下で、 $p(u_1), p(u_2), p(u_3), p(u_4)$  は提案手法を通じて、

$$\hat{p}(u_1) = 0.289, \quad \hat{p}(u_2) = 0.297, \quad \hat{p}(u_3) = 0.236, \quad \hat{p}(u_4) = 0.177,$$

のように推定される。これらの確率から、PNS  $p(u_2)$  と因果リスク差  $p(u_2) - p(u_3)$  はそれぞれ  $\hat{p}(u_2) = 0.297$  と  $\hat{p}(u_2) - \hat{p}(u_3) = 0.061$  のように推定される。表 3.3 は PNS と因果リスク差の推定値を 95%信頼区間付きで示す。ここで、95%信頼区間を導出するために、推定値の 1,000 個のブートストラップ複製の 2.5%分位点と 97.5%分位点を用いた<sup>2</sup>。

<sup>2</sup>正確にいうと、可逆行列  $\hat{P}_{x_1}$  と  $\hat{P}_{x_0}$  が得られた 991 個のブートストラップ複製のみを用いた。

表 3.3: NSW データセットにおける PNS と因果リスク差の推定値.

	Estimate (95%CI)
PNS	0.297 (0.041, 0.704)
causal risk difference	0.061 (-0.612, 0.615)

### 3.6 考察

潜在結果変数の同時確率はユニット選択問題 [Li and Pearl, 2019], 社会プログラムの影響評価の問題 [Heckman et al., 1997], 処置効果の非遵守の問題 [Angrist et al., 1996, Balke and Pearl, 1997], 主要層別化法 [Frangakis and Rubin, 2002], 自然な直接効果と間接効果の識別問題 [Pearl, 2001], 機械学習モデルの説明可能性 [Galhotra et al., 2021, Kommiya Mothilal et al., 2021, Watson et al., 2021] など様々な場面で現れる. さらに, 潜在結果変数の同時確率が識別可能であれば因果リスクも識別可能であるが, 逆は真ではない. そのため, これらの確率の識別および推定問題は因果推論における重要なトピックとなっている. 第 3 部では, 観察研究においてこの問題を解決するために, (i) 共変量情報を用いた 2 種類の新しい識別条件と, (ii) 拡張 Lagrange 法に基づく新しい統計的推定法を提案した. 既存の単調性のような識別条件が成立しないノンパラメトリックなデータ生成過程において, これらの確率を識別するために共変量情報がいかに有用であるかについて, これまであまり議論されてこなかった. 従って, 本論文の第 3 部の結果は, ノンパラメトリックな因果モデルの下での因果推論において, 解決可能な識別および推定問題の範囲を広げるものである.



## 第 4 部

# 実験研究における代理共変量情報に基づいた潜在結果変数の同時確率の識別と推定

第 4 部では，実験研究の枠組みにおいて共変量情報を利用することにより潜在結果変数の同時確率を識別するための新たな条件を与える．この条件は，第 3 部で用いられている共変量情報よりも少ない共変量情報の下で，潜在結果タイプの確率に対する一致推定量を与えることを可能にする．

最初に，4.1 節で問題設定について述べ，4.2 節で，潜在結果変数の同時確率の識別可能性について述べる．4.2 節の結果に基づいて，4.3 節では潜在結果変数の同時確率に対する新たな推定量を提案する．最後に，数値実験と適用例を通して提案した推定量の有用性について検証する．第 4 部の内容は第 3 部の実験研究版に相当し，Shingaki and Kuroki [2023] で公表されている．

### 4.1 問題設定と記法

図 4.1 で示されるような，実験的な処置（例えば COVID-19 のワクチン）の結果を対照的な処置（例えば プラセボ）の結果と比較することを目的としたランダム化実験について考える．図 4.1 において，変数  $X, Y$  と変数からなる集合  $U$  が与えられた下

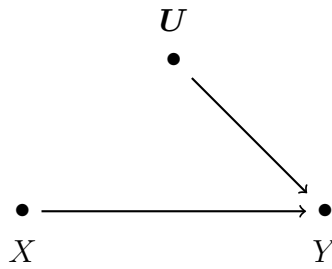


図 4.1: ランダム化実験における因果的な従属関係のグラフ表現

で,  $X$  から  $Y$  への有向辺  $X \rightarrow Y$  は  $X$  が  $Y$  への効果を持ち得ることを示し,  $Y$  から  $X$  への有向辺  $Y \rightarrow X$  がないことは  $Y$  が  $X$  の原因とはなり得ないことを示している.  $U$  から  $Y$  への有向辺  $U \rightarrow Y$  は,  $U$  のいくつかの要素が  $Y$  への効果を持ち得ることを示し,  $Y$  から  $U$  への有向辺  $Y \rightarrow U$  がないことは,  $Y$  が  $U$  のいかなる要素の原因とはなり得ないことを示している.  $X$  と  $U$  の間の辺 ( $X \rightarrow U$ ,  $X \leftarrow U$ ,  $X \leftrightarrow U$ ) がないことは,  $X$  と  $U$  の間に因果関係や関連性を表す関係がないことを意味する.

図 4.1 において,  $X$  と  $Y$  はそれぞれ観測可能な二値の処置変数と結果変数であると仮定する.  $x \in \{x_0, x_1\}$  は変数  $X$  が取る値であり,  $x_0$  は対照となる処置 (例えばワクチン接種),  $x_1$  は実験的な処置 (例えばワクチン接種) を,  $y \in \{y_0, y_1\}$  は変数  $Y$  が取る値であり,  $y_0$  は疾患あり,  $y_1$  は疾患なしを表す.  $U$  は共変量ベクトルと呼ばれ,  $X$  や  $Y$  によって影響されることがない (観測変数と未観測変数の両方を含む) すべての変数からなる集合を表す. 観察研究では  $U$  のいくつかの要素が  $X$  と  $Y$  に効果を持ち得るが, 図 4.1 のようなランダム化実験では,  $X$  がランダム化されているため,  $U$  のいかなる要素も  $X$  と関連しないことに注意しておく.

第 4 部の興味は実験研究において潜在結果変数の同時確率を評価することである. 前節で述べた四つのタイプ (“doomed”, “causative”, “preventive”, “immune”) は,  $(Y_{x_0}, Y_{x_1})$  を通してそれぞれ

$$\begin{aligned}
 u_1 &= (Y_{x_0} = y_0, Y_{x_1} = y_0), & u_2 &= (Y_{x_0} = y_0, Y_{x_1} = y_1), \\
 u_3 &= (Y_{x_0} = y_1, Y_{x_1} = y_0), & u_4 &= (Y_{x_0} = y_1, Y_{x_1} = y_1)
 \end{aligned}$$

によって表され、対応する評価したい潜在結果タイプの確率は  $p(u_1)$ ,  $p(u_2)$ ,  $p(u_3)$ ,  $p(u_4)$  で与えられる。これらの潜在結果タイプの確率は次の意味で、因果推論において基礎的な役割を果たす:

- (i) 潜在結果タイプの確率が識別可能であれば、因果リスクも識別可能であるが、逆は成り立たない [Pearl, 2009, Tian and Pearl, 2000],
- (ii) 潜在結果タイプの確率は「必要な原因」「十分な原因」「必要十分な原因」の確率的な側面を評価することを可能にする。

しかしながら、 $Y_{x_0}$  と  $Y_{x_1}$  は各被験者に対してランダム化実験の下でさえも同時に観測できないので、通常これらの確率は識別できない。なお、ランダム化実験において、単調性 (monotonicity), 即ち、 $p(u_3) = 0$  の下で、 $p(u_1)$ ,  $p(u_2)$ ,  $p(u_4)$  は識別可能であり、 $p(u_1) = p(y_0|x_1)$ ,  $p(u_2) = p(y_1|x_1) - p(y_1|x_0)$ ,  $p(u_3) = 0$  and  $p(u_4) = p(y_1|x_0)$  で与えられることに注意する [Pearl, 2009, Tian and Pearl, 2000]。以下では、 $p(u_3) \neq 0$  を仮定し、 $p(u_3) \neq 0$  の状況で潜在結果タイプの確率の識別可能性について考える。

## 4.2 識別

前節の問題設定について、図 4.2a で示される有向グラフで表せる、ランダム化実験に基づいた潜在結果タイプの確率を評価する問題を定式化する。ここでは、共変量  $Z$  と  $W$  は共変量の集合  $U$  の代理変数として測定されている。この図で示されるデータ生成過程は

$$Y = g_y(X, \mathbf{U}, \epsilon_y), \quad X = g_x(\epsilon_x), \quad Z = g_z(\mathbf{U}, \epsilon_z), \quad W = g_w(\mathbf{U}, \epsilon_w) \quad (4.1)$$

となる。ここで、 $\epsilon_x, \epsilon_y, \epsilon_z, \epsilon_w$  は独立な攪乱項 (random disturbances) であり、 $\mathbf{U}$  とも独立である。式 (4.1) のような構造方程式モデルがデータ生成過程を表すのに用いられるとき、図 4.2a で示される対応するグラフは因果ダイアグラムと呼ばれる。図 4.2a で示されるような状況では、 $\mathbf{U}$  は被験者の処置への反応の仕方に影響する

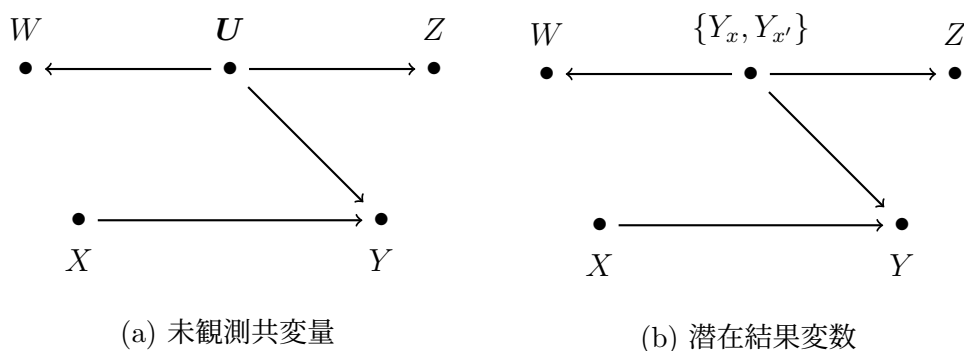


図 4.2: ランダム化実験における因果的な従属性のグラフ表現.

すべての共変量を含むことができるため、 $U$  が与えられた下で  $\{X, Y\}$  と独立である共変量  $Z$  と  $W$  の存在を ( $U$  の中から取り出すことで) 仮定するのは不自然なことではない. 以降,  $U \cup \{\epsilon_y, \epsilon_z, \epsilon_w\}$  を改めて  $U$  と表し, Pearl [2009, 8.2.2 節] に従って,  $U$  の複雑さに関係なく  $U$  の  $Y$  への影響を  $X$  と  $Y$  の間の関数関係の修正に制限する.

図 4.2a で示されるような因果モデルにおける二つの二値変数  $Y \in \{x_0, x_1\}$  と  $X \in \{y_0, y_1\}$  を結びつける構造方程式

$$y = f(x, \mathbf{u})$$

について考える. このとき, 任意に与えられた  $\mathbf{u}$  に対して,  $f_{\mathbf{u}} = f(\cdot, \mathbf{u}): X \rightarrow Y$  のグラフ

$$\Gamma^{\mathbf{u}} = \{(x, f(x, \mathbf{u})) \in X \times Y \mid x \in X\}$$

は次の四つの集合の何れかとなる:

$$\Gamma_1 = \{(x_0, y_0), (x_1, y_0)\}, \quad \Gamma_2 = \{(x_0, y_0), (x_1, y_1)\},$$

$$\Gamma_3 = \{(x_0, y_1), (x_1, y_0)\}, \quad \Gamma_4 = \{(x_0, y_1), (x_1, y_1)\}.$$

$\mathbf{u}$  がその定義域に沿って変化するとき, その変化がどれほど複雑であっても, モデルに与える影響はこれら四つのグラフに対応する関数の中で  $X$  と  $Y$  の関係が入れ

替わるだけである。したがって、個体  $i$  と  $j$  に対して式 (4.1) からサンプリングされる  $(X, Y, U)$  をそれぞれ、 $(X(i), Y(i), U(i))$  と  $(X(j), Y(j), U(j))$  と置き、 $U(i)$  と  $U(j)$  の間に定まる

$$\begin{aligned} U(i) \sim U(j) &\iff Y_x(i) = Y_x(j) \text{ for } x \in \{x_0, x_1\} \\ &\iff f(\text{do}(X(i) = x), U(i)) = f(\text{do}(X(j) = x), U(j)) \text{ for } x \in \{x_0, x_1\} \\ &\iff \Gamma^{U(i)} = \Gamma^{U(j)} \end{aligned}$$

のような同値関係  $\sim$  に基づいて、図 4.3 で示されるように  $U$  の定義域  $\mathcal{U}$  を、潜在結果変数の組合せ  $(Y_{x_0}, Y_{x_1})$  を表す四つの同値類に分割することができる: このことから、 $U$  が取る値は四つの関数の一つを選択することに対応することがわかる [Pearl, 2009].

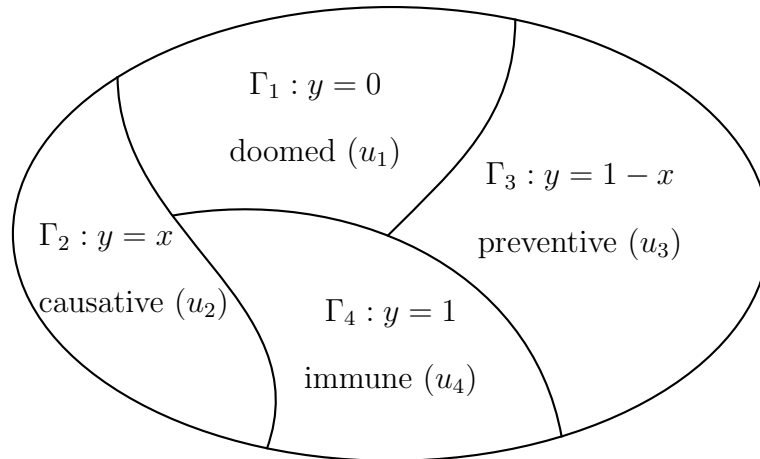


図 4.3:  $U$  の四つの同値類への分割.

これらの考察に基づき、Lash et al. [2021, p. 59] に従って、 $U$  の状態を次の四つの潜在結果タイプ “doomed” ( $u_1$ ), “causative” ( $u_2$ ), “preventive” ( $u_3$ ), “immune” ( $u_4$ ) の何れかを取る変数  $U$  として再定義し、任意の  $x, y, z, w$  に対して、図 4.2a を図 4.2b のように再表現すれば、 $X$  が与えられた下での  $Y, Z, W$  の同時確率  $p(y, z, w | x)$  の逐次分解は

$$p(y, z, w | x) = \sum_{i=1}^4 p(y | x, u_i) p(z | u_i) p(w | u_i) p(u_i)$$

のように与えられる. つまり,  $Z$  は二つ以上の値  $z \in \{z_1, z_2, \dots, z_k\}$  ( $k \geq 2$ ) を取り,  $W$  は三つ以上の値  $w \in \{w_1, w_2, \dots, w_l\}$  ( $l \geq 3$ ) を取り得るとき,  $x \in \{x_0, x_1\}$ ,  $y \in \{y_0, y_1\}$ ,  $w \in \{w_1, w_2, \dots, w_l\}$  に対して,

$$Q_{xyw} = \begin{pmatrix} p(y | x) & p(y, w | x) \\ p(y, z_1 | x) & p(y, z_1, w | x) \end{pmatrix},$$

$$R_{x_0y_0} = \begin{pmatrix} 1 & p(z_1 | u_1) \\ 1 & p(z_1 | u_2) \end{pmatrix}, \quad R_{x_0y_1} = \begin{pmatrix} 1 & p(z_1 | u_3) \\ 1 & p(z_1 | u_4) \end{pmatrix},$$

$$R_{x_1y_0} = \begin{pmatrix} 1 & p(z_1 | u_1) \\ 1 & p(z_1 | u_3) \end{pmatrix}, \quad R_{x_1y_1} = \begin{pmatrix} 1 & p(z_1 | u_2) \\ 1 & p(z_1 | u_4) \end{pmatrix},$$

$$S_{x_0y_0w} = \begin{pmatrix} 1 & p(w | u_1) \\ 1 & p(w | u_2) \end{pmatrix}, \quad S_{x_0y_1w} = \begin{pmatrix} 1 & p(w | u_3) \\ 1 & p(w | u_4) \end{pmatrix},$$

$$S_{x_1y_0w} = \begin{pmatrix} 1 & p(w | u_1) \\ 1 & p(w | u_3) \end{pmatrix}, \quad S_{x_1y_1w} = \begin{pmatrix} 1 & p(w | u_2) \\ 1 & p(w | u_4) \end{pmatrix},$$

$$\Delta_{x_0y_0} = \begin{pmatrix} p(u_1) & 0 \\ 0 & p(u_2) \end{pmatrix}, \quad \Delta_{x_0y_1} = \begin{pmatrix} p(u_3) & 0 \\ 0 & p(u_4) \end{pmatrix},$$

$$\Delta_{x_1y_0} = \begin{pmatrix} p(u_1) & 0 \\ 0 & p(u_3) \end{pmatrix}, \quad \Delta_{x_1y_1} = \begin{pmatrix} p(u_2) & 0 \\ 0 & p(u_4) \end{pmatrix}$$

と置けば,  $x \in \{x_0, x_1\}$ ,  $w_{l_1}, w_{l_2} \in \{w_1, w_2, \dots, w_l\}$  ( $w_{l_1} \neq w_{l_2}$ ) に対して

$$Q_{xy_0w_{l_1}} = R_{xy_0}^\top \Delta_{xy_0} S_{xy_0w_{l_1}}, \quad Q_{xy_0w_{l_2}} = R_{xy_0}^\top \Delta_{xy_0} S_{xy_0w_{l_2}},$$

$$Q_{xy_1w_{l_1}} = R_{xy_1}^\top \Delta_{xy_1} S_{xy_1w_{l_1}}, \quad Q_{xy_1w_{l_2}} = R_{xy_1}^\top \Delta_{xy_1} S_{xy_1w_{l_2}}$$

と表せる. ここで, “ $\top$ ” は転置を表す. また,  $x \in \{x_0, x_1\}$ ,  $y \in \{y_0, y_1\}$ ,  $w \in \{w_1, w_2, \dots, w_l\}$  に対して  $Q_{xyw}$  が可逆であるとき,

$$S_{xy_0w_{l_1}} Q_{xy_0w_{l_1}}^{-1} = S_{xy_0w_{l_2}} Q_{xy_0w_{l_2}}^{-1}, \quad S_{xy_1w_{l_1}} Q_{xy_1w_{l_1}}^{-1} = S_{xy_1w_{l_2}} Q_{xy_1w_{l_2}}^{-1}$$

となる. これらの行列方程式を成分ごとに  $p(w | u)$  について解くことで, 次の定理を得ることができる:

**定理 4.1.**  $Z, W$  はそれぞれ  $k (\geq 2), l (\geq 3)$  個の値  $z \in \{z_1, z_2, \dots, z_k\}, w \in \{w_1, w_2, \dots, w_l\}$  を取る変数であるとする. このとき, 次の条件が満たされるならば, 潜在結果タイプの確率  $p(u_1), p(u_2), p(u_3), p(u_4)$  は識別可能である:

(1)  $x \in \{x_0, x_1\}, y \in \{y_0, y_1\}, z \in \{z_1, z_2, \dots, z_k\}, w \in \{w_1, w_2, \dots, w_l\}$  に対して, 正の確率  $p(y, z, w|x)$  が得られる.

(2)  $W \perp\!\!\!\perp Z | U$  かつ  $X \perp\!\!\!\perp Z | U$  が成り立つ.

(3)  $x \in \{x_0, x_1\}, y \in \{y_0, y_1\}, w \in \{w_1, w_2, \dots, w_l\}$  に対して,  $Q_{xyw}$  は可逆であり,  $w_{l_1}, w_{l_2} \in \{w_1, w_2, \dots, w_l\}$  ( $w_{l_1} \neq w_{l_2}$ ) に対して

$$\begin{aligned} \frac{|Q_{x_0 y_0 w_{l_1}}|}{|Q_{x_0 y_0 w_{l_2}}|} &\neq \frac{|Q_{x_1 y_0 w_{l_1}}|}{|Q_{x_1 y_0 w_{l_2}}|}, & \frac{|Q_{x_0 y_0 w_{l_1}}|}{|Q_{x_0 y_0 w_{l_2}}|} &\neq \frac{|Q_{x_1 y_1 w_{l_1}}|}{|Q_{x_1 y_1 w_{l_2}}|}, \\ \frac{|Q_{x_0 y_1 w_{l_1}}|}{|Q_{x_0 y_1 w_{l_2}}|} &\neq \frac{|Q_{x_1 y_0 w_{l_1}}|}{|Q_{x_1 y_0 w_{l_2}}|}, & \frac{|Q_{x_0 y_1 w_{l_1}}|}{|Q_{x_0 y_1 w_{l_2}}|} &\neq \frac{|Q_{x_1 y_1 w_{l_1}}|}{|Q_{x_1 y_1 w_{l_2}}|} \end{aligned}$$

が成り立つ.

ここで,  $|\cdot|$  は行列式を表す.

**証明.** 条件 (2) から, 一致性により,  $u \in \{u_1, u_2, u_3, u_4\}, z \in \{z_1, z_2, \dots, z_k\}, w \in \{w_1, w_2, \dots, w_l\}, x \in \{x_0, x_1\}, y \in \{y_0, y_1\}$  に対して,

$$p(u|x) = p(u), \quad p(y, z, w|x) = p(y_x, z, w|x) = p(y_x, z, w),$$

$$p(z, w|x) = p(z, w), \quad p(w|x, u) = p(w|u)$$

となる. 一致性から,

$$p(y_0, z, w, u_i | x_0) = p(z, w, u_i | x_0), \quad i = 1, 2;$$

$$p(y_1, z, w, u_i | x_0) = p(z, w, u_i | x_0), \quad i = 3, 4,$$

$$p(y_0, z, w, u_i | x_1) = p(z, w, u_i | x_1), \quad i = 1, 3;$$

$$p(y_1, z, w, u_i | x_1) = p(z, w, u_i | x_1), \quad i = 2, 4,$$

$$p(y_0, z, w, u_i | x_0) = 0, \quad i = 3, 4;$$

$$p(y_1, z, w, u_i | x_0) = 0, \quad i = 1, 2,$$

$$p(y_0, z, w, u_i | x_1) = 0, \quad i = 2, 4;$$

$$p(y_1, z, w, u_i | x_1) = 0, \quad i = 1, 3$$

であることに注意して,  $z \in \{z_1, z_2, \dots, z_k\}$ ,  $w \in \{w_1, w_2, \dots, w_l\}$  に対して,

$$p(y_0, z, w | x_0) = \sum_{i=1,2} p(w | u_i)p(z | u_i)p(u_i),$$

$$p(y_0, z, w | x_1) = \sum_{i=1,3} p(w | u_i)p(z | u_i)p(u_i),$$

$$p(y_1, z, w | x_0) = \sum_{i=3,4} p(w | u_i)p(z | u_i)p(u_i),$$

$$p(y_1, z, w | x_1) = \sum_{i=2,4} p(w | u_i)p(z | u_i)p(u_i),$$

$$p(z, w) = \sum_{i=1}^4 p(w | u_i)p(z | u_i)p(u_i),$$

$$p(y_0 | x_0) = \sum_{i=1,2} p(u_i), \quad p(y_0 | x_1) = \sum_{i=1,3} p(u_i)$$

を得る. このとき,  $x \in \{x_0, x_1\}$ ,  $y \in \{y_0, y_1\}$ ,  $w \in \{w_1, w_2\}$  に対して,

$$Q_{xyw} = \begin{pmatrix} p(y | x) & p(y, w | x) \\ p(y, z_1 | x) & p(y, z_1, w | x) \end{pmatrix} = p(y | x) \begin{pmatrix} 1 & p(w | x, y) \\ p(z_1 | x, y) & p(z_1, w | x, y) \end{pmatrix}$$

かつ

$$\begin{aligned} R_{x_0y_0} &= \begin{pmatrix} 1 & p(z_1 | u_1) \\ 1 & p(z_1 | u_2) \end{pmatrix}, & R_{x_0y_1} &= \begin{pmatrix} 1 & p(z_1 | u_3) \\ 1 & p(z_1 | u_4) \end{pmatrix}, \\ R_{x_1y_0} &= \begin{pmatrix} 1 & p(z_1 | u_1) \\ 1 & p(z_1 | u_3) \end{pmatrix}, & R_{x_1y_1} &= \begin{pmatrix} 1 & p(z_1 | u_2) \\ 1 & p(z_1 | u_4) \end{pmatrix}, \\ S_{x_0y_0w_1} &= \begin{pmatrix} 1 & p(w_1 | u_1) \\ 1 & p(w_1 | u_2) \end{pmatrix}, & S_{x_0y_1w_1} &= \begin{pmatrix} 1 & p(w_1 | u_3) \\ 1 & p(w_1 | u_4) \end{pmatrix}, \end{aligned}$$



$$\begin{aligned}
S_{x_1 y_0 w_1} &= \begin{pmatrix} 1 & p(w_1 | u_1) \\ 1 & p(w_1 | u_3) \end{pmatrix}, & S_{x_1 y_1 w_1} &= \begin{pmatrix} 1 & p(w_1 | u_2) \\ 1 & p(w_1 | u_4) \end{pmatrix}, \\
S_{x_0 y_0 w_2} &= \begin{pmatrix} 1 & p(w_2 | u_1) \\ 1 & p(w_2 | u_2) \end{pmatrix}, & S_{x_0 y_1 w_2} &= \begin{pmatrix} 1 & p(w_2 | u_3) \\ 1 & p(w_2 | u_4) \end{pmatrix}, \\
S_{x_1 y_0 w_2} &= \begin{pmatrix} 1 & p(w_2 | u_1) \\ 1 & p(w_2 | u_3) \end{pmatrix}, & S_{x_1 y_1 w_2} &= \begin{pmatrix} 1 & p(w_2 | u_2) \\ 1 & p(w_2 | u_4) \end{pmatrix}, \\
\Delta_{x_0 y_0} &= \begin{pmatrix} p(u_1) & 0 \\ 0 & p(u_2) \end{pmatrix}, & \Delta_{x_0 y_1} &= \begin{pmatrix} p(u_3) & 0 \\ 0 & p(u_4) \end{pmatrix}, \\
\Delta_{x_1 y_0} &= \begin{pmatrix} p(u_1) & 0 \\ 0 & p(u_3) \end{pmatrix}, & \Delta_{x_1 y_1} &= \begin{pmatrix} p(u_2) & 0 \\ 0 & p(u_4) \end{pmatrix}
\end{aligned}$$

と置くと,  $x \in \{x_0, x_1\}$  に対して,

$$\begin{aligned}
Q_{x y_0 w_1} &= R_{x y_0}^\top \Delta_{x y_0} S_{x y_0 w_1}, & Q_{x y_0 w_2} &= R_{x y_0}^\top \Delta_{x y_0} S_{x y_0 w_2}, \\
Q_{x y_1 w_1} &= R_{x y_1}^\top \Delta_{x y_1} S_{x y_1 w_1}, & Q_{x y_1 w_2} &= R_{x y_1}^\top \Delta_{x y_1} S_{x y_1 w_2}
\end{aligned} \tag{4.2}$$

を得る.  $x \in \{x_0, x_1\}$ ,  $y \in \{y_0, y_1\}$ ,  $w \in \{w_1, w_2\}$  に対して,  $Q_{x y w}$  は可逆なので,  $x \in \{x_0, x_1\}$  に対して,

$$S_{x y_0 w_1} Q_{x y_0 w_1}^{-1} = S_{x y_0 w_2} Q_{x y_0 w_2}^{-1}, \quad S_{x y_1 w_1} Q_{x y_1 w_1}^{-1} = S_{x y_1 w_2} Q_{x y_1 w_2}^{-1} \tag{4.3}$$

となる. よって,

$$\begin{aligned}
& \frac{1}{|Q_{x_0 y_0 w_1}|} \begin{pmatrix} 1 & p(w_1 | u_1) \\ 1 & p(w_1 | u_2) \end{pmatrix} \begin{pmatrix} p(y_0, z_1, w_1 | x_0) & -p(y_0, w_1 | x_0) \\ -p(y_0, z_1 | x_0) & p(y_0 | x_0) \end{pmatrix} \\
&= \frac{1}{|Q_{x_0 y_0 w_2}|} \begin{pmatrix} 1 & p(w_2 | u_1) \\ 1 & p(w_2 | u_2) \end{pmatrix} \begin{pmatrix} p(y_0, z_1, w_2 | x_0) & -p(y_0, w_2 | x_0) \\ -p(y_0, z_1 | x_0) & p(y_0 | x_0) \end{pmatrix}, \tag{4.4} \\
& \frac{1}{|Q_{x_0 y_1 w_1}|} \begin{pmatrix} 1 & p(w_1 | u_3) \\ 1 & p(w_1 | u_4) \end{pmatrix} \begin{pmatrix} p(y_1, z_1, w_1 | x_0) & -p(y_1, w_1 | x_0) \\ -p(y_1, z_1 | x_0) & p(y_1 | x_0) \end{pmatrix}
\end{aligned}$$

$$= \frac{1}{|Q_{x_0 y_1 w_2}|} \begin{pmatrix} 1 & p(w_2 | u_3) \\ 1 & p(w_2 | u_4) \end{pmatrix} \begin{pmatrix} p(y_1, z_1, w_2 | x_0) & -p(y_1, w_2 | x_0) \\ -p(y_1, z_1 | x_0) & p(y_1 | x_0) \end{pmatrix}, \quad (4.5)$$

$$\begin{aligned} & \frac{1}{|Q_{x_1 y_0 w_1}|} \begin{pmatrix} 1 & p(w_1 | u_1) \\ 1 & p(w_1 | u_3) \end{pmatrix} \begin{pmatrix} p(y_0, z_1, w_1 | x_1) & -p(y_0, w_1 | x_1) \\ -p(y_0, z_1 | x_1) & p(y_0 | x_0) \end{pmatrix} \\ &= \frac{1}{|Q_{x_1 y_0 w_2}|} \begin{pmatrix} 1 & p(w_2 | u_1) \\ 1 & p(w_2 | u_3) \end{pmatrix} \begin{pmatrix} p(y_0, z_1, w_2 | x_1) & -p(y_0, w_2 | x_1) \\ -p(y_0, z_1 | x_1) & p(y_0 | x_1) \end{pmatrix}, \quad (4.6) \end{aligned}$$

$$\begin{aligned} & \frac{1}{|Q_{x_1 y_1 w_1}|} \begin{pmatrix} 1 & p(w_1 | u_2) \\ 1 & p(w_1 | u_4) \end{pmatrix} \begin{pmatrix} p(y_1, z_1, w_1 | x_1) & -p(y_1, w_1 | x_1) \\ -p(y_1, z_1 | x_1) & p(y_1 | x_1) \end{pmatrix} \\ &= \frac{1}{|Q_{x_1 y_1 w_2}|} \begin{pmatrix} 1 & p(w_2 | u_2) \\ 1 & p(w_2 | u_4) \end{pmatrix} \begin{pmatrix} p(y_1, z_1, w_2 | x_1) & -p(y_1, w_2 | x_1) \\ -p(y_1, z_1 | x_1) & p(y_1 | x_1) \end{pmatrix} \quad (4.7) \end{aligned}$$

であるので、(4.4)の(1, 2)成分から

$$p(w_1 | u_1) = \frac{|Q_{x_0 y_0 w_1}|}{|Q_{x_0 y_0 w_2}|} (p(w_2 | u_1) - p(w_2 | x_0, y_0)) + p(w_1 | x_0, y_0),$$

(4.4)の(2, 2)成分から

$$p(w_1 | u_2) = \frac{|Q_{x_0 y_0 w_1}|}{|Q_{x_0 y_0 w_2}|} (p(w_2 | u_2) - p(w_2 | x_0, y_0)) + p(w_1 | x_0, y_0),$$

(4.5)の(1, 2)成分から

$$p(w_1 | u_3) = \frac{|Q_{x_0 y_1 w_1}|}{|Q_{x_0 y_1 w_2}|} (p(w_2 | u_3) - p(w_2 | x_0, y_1)) + p(w_1 | x_0, y_1),$$

(4.5)の(2, 2)成分から

$$p(w_1 | u_4) = \frac{|Q_{x_0 y_1 w_1}|}{|Q_{x_0 y_1 w_2}|} (p(w_2 | u_4) - p(w_2 | x_0, y_1)) + p(w_1 | x_0, y_1),$$

(4.6)の(1, 2)成分から

$$p(w_1 | u_1) = \frac{|Q_{x_1 y_0 w_1}|}{|Q_{x_1 y_0 w_2}|} (p(w_2 | u_1) - p(w_2 | x_1, y_0)) + p(w_1 | x_1, y_0),$$

(4.6)の(2, 2)成分から

$$p(w_1 | u_3) = \frac{|Q_{x_1 y_0 w_1}|}{|Q_{x_1 y_0 w_2}|} (p(w_2 | u_3) - p(w_2 | x_1, y_0)) + p(w_1 | x_1, y_0),$$

(4.7) の (1, 2) 成分から

$$p(w_1 | u_2) = \frac{|Q_{x_1 y_1 w_1}|}{|Q_{x_1 y_1 w_2}|} (p(w_2 | u_2) - p(w_2 | x_1, y_1)) + p(w_1 | x_1, y_1),$$

(4.7) の (2, 2) 成分から

$$p(w_1 | u_4) = \frac{|Q_{x_1 y_1 w_1}|}{|Q_{x_1 y_1 w_2}|} (p(w_2 | u_4) - p(w_2 | x_1, y_1)) + p(w_1 | x_1, y_1).$$

となる。よって,

$$\begin{pmatrix} 1 & -\frac{|Q_{x_0 y_0 w_1}|}{|Q_{x_0 y_0 w_2}|} \\ 1 & -\frac{|Q_{x_1 y_0 w_1}|}{|Q_{x_1 y_0 w_2}|} \end{pmatrix} \begin{pmatrix} p(w_1 | u_1) \\ p(w_2 | u_1) \end{pmatrix} = \begin{pmatrix} p(w_1 | x_0, y_0) - \frac{|Q_{x_0 y_0 w_1}|}{|Q_{x_0 y_0 w_2}|} p(w_2 | x_0, y_0) \\ p(w_1 | x_1, y_0) - \frac{|Q_{x_1 y_0 w_1}|}{|Q_{x_1 y_0 w_2}|} p(w_2 | x_1, y_0) \end{pmatrix}, \quad (4.8)$$

$$\begin{pmatrix} 1 & -\frac{|Q_{x_0 y_0 w_1}|}{|Q_{x_0 y_0 w_2}|} \\ 1 & -\frac{|Q_{x_1 y_1 w_1}|}{|Q_{x_1 y_1 w_2}|} \end{pmatrix} \begin{pmatrix} p(w_1 | u_2) \\ p(w_2 | u_2) \end{pmatrix} = \begin{pmatrix} p(w_1 | x_0, y_0) - \frac{|Q_{x_0 y_0 w_1}|}{|Q_{x_0 y_0 w_2}|} p(w_2 | x_0, y_0) \\ p(w_1 | x_1, y_1) - \frac{|Q_{x_1 y_1 w_1}|}{|Q_{x_1 y_1 w_2}|} p(w_2 | x_1, y_1) \end{pmatrix}, \quad (4.9)$$

$$\begin{pmatrix} 1 & -\frac{|Q_{x_0 y_1 w_1}|}{|Q_{x_0 y_1 w_2}|} \\ 1 & -\frac{|Q_{x_1 y_0 w_1}|}{|Q_{x_1 y_0 w_2}|} \end{pmatrix} \begin{pmatrix} p(w_1 | u_3) \\ p(w_2 | u_3) \end{pmatrix} = \begin{pmatrix} p(w_1 | x_0, y_1) - \frac{|Q_{x_0 y_1 w_1}|}{|Q_{x_0 y_1 w_2}|} p(w_2 | x_0, y_1) \\ p(w_1 | x_1, y_0) - \frac{|Q_{x_1 y_0 w_1}|}{|Q_{x_1 y_0 w_2}|} p(w_2 | x_1, y_0) \end{pmatrix}, \quad (4.10)$$

$$\begin{pmatrix} 1 & -\frac{|Q_{x_0 y_1 w_1}|}{|Q_{x_0 y_1 w_2}|} \\ 1 & -\frac{|Q_{x_1 y_1 w_1}|}{|Q_{x_1 y_1 w_2}|} \end{pmatrix} \begin{pmatrix} p(w_1 | u_4) \\ p(w_2 | u_4) \end{pmatrix} = \begin{pmatrix} p(w_1 | x_0, y_1) - \frac{|Q_{x_0 y_1 w_1}|}{|Q_{x_0 y_1 w_2}|} p(w_2 | x_0, y_1) \\ p(w_1 | x_1, y_1) - \frac{|Q_{x_1 y_1 w_1}|}{|Q_{x_1 y_1 w_2}|} p(w_2 | x_1, y_1) \end{pmatrix}. \quad (4.11)$$

条件 (3) から

$$\begin{aligned} \frac{|Q_{x_0 y_0 w_1}|}{|Q_{x_0 y_0 w_2}|} &\neq \frac{|Q_{x_1 y_0 w_1}|}{|Q_{x_1 y_0 w_2}|}, & \frac{|Q_{x_0 y_0 w_1}|}{|Q_{x_0 y_0 w_2}|} &\neq \frac{|Q_{x_1 y_1 w_1}|}{|Q_{x_1 y_1 w_2}|}, \\ \frac{|Q_{x_0 y_1 w_1}|}{|Q_{x_0 y_1 w_2}|} &\neq \frac{|Q_{x_1 y_0 w_1}|}{|Q_{x_1 y_0 w_2}|}, & \frac{|Q_{x_0 y_1 w_1}|}{|Q_{x_0 y_1 w_2}|} &\neq \frac{|Q_{x_1 y_1 w_1}|}{|Q_{x_1 y_1 w_2}|} \end{aligned}$$

が成り立つので,  $u \in \{u_1, u_2, u_3, u_4\}$ ,  $w \in \{w_1, w_2\}$  に対して,  $p(w|u)$  についてこれらの方程式を解くことで,  $p(w|u)$  は識別可能となる.

次に,  $x \in \{x_0, x_1\}$  に対して,  $S_{xy_0w_1}$  は識別可能かつ可逆であり,  $Q_{xy_0w_1}$  は観測できることに注意すると,

$$Q_{x_0y_0w_1} S_{x_0y_0w_1}^{-1} = R_{x_0y_0}^\top \Delta_{x_0y_0} = \begin{pmatrix} p(u_1) & p(u_2) \\ p(u_1, z_1) & p(u_2, z_1) \end{pmatrix}, \quad (4.12)$$

$$Q_{x_1y_0w_1} S_{x_1y_0w_1}^{-1} = R_{x_1y_0}^\top \Delta_{x_1y_0} = \begin{pmatrix} p(u_1) & p(u_3) \\ p(u_1, z_1) & p(u_3, z_1) \end{pmatrix}, \quad (4.13)$$

$$Q_{x_0y_1w_1} S_{x_0y_1w_1}^{-1} = R_{x_0y_1}^\top \Delta_{x_0y_1} = \begin{pmatrix} p(u_3) & p(u_4) \\ p(u_3, z_1) & p(u_4, z_1) \end{pmatrix}, \quad (4.14)$$

$$Q_{x_1y_1w_1} S_{x_1y_1w_1}^{-1} = R_{x_1y_1}^\top \Delta_{x_1y_1} = \begin{pmatrix} p(u_2) & p(u_4) \\ p(u_2, z_1) & p(u_4, z_1) \end{pmatrix} \quad (4.15)$$

となる. よって,  $p(u_1), p(u_2), p(u_3), p(u_4)$  は識別可能である.  $\square$

定理 4.1 についていくつか補足する. 条件 (1) と (3) は観測データから検証可能である. 条件 (2) については,  $X \perp\!\!\!\perp \{U, Z, W\}$  はランダム化実験では自ずと満たされる. また, Kuroki and Pearl [2014] では, 未観測変数を含む確率は完全に識別されないが, 定理 4.1 は潜在結果タイプの確率が識別可能であることを示している. さらに, 第 3 部の定理 3.2 では, 二つの代理共変量が少なくとも四つの値を取ることがを要請していたが, 定理 4.1 の二つの代理共変量は片方が二つ, もう片方が三つ以上の値を取れば十分である.

### 4.3 推定

定理 4.1 の証明からわかるように, 前節の条件の下で識別できる潜在結果タイプの確率に対して, 次のような推定量を考えることができる.  $Q_{xyw}$  の各成分  $p(y|x)$ ,  $p(w|x, y)$ ,  $p(z, w|x, y)$ ,  $p(z|x, y)$  を最尤推定量  $\hat{p}(y|x)$ ,  $\hat{p}(w|x, y)$ ,  $\hat{p}(z, w|x, y)$ ,  $\hat{p}(z|x, y)$  で置き換えて得られる行列  $\hat{Q}_{xyw}$  を考える. このとき, 式 (4.8)–(4.11) か

ら,  $w \in \{w_1, w_2\}$ ,  $u \in \{u_1, u_2, u_3, u_4\}$  に対して,  $p(w|u)$  の一致推定量は

$$\begin{aligned} \begin{pmatrix} \hat{p}(w_1 | u_1) \\ \hat{p}(w_2 | u_1) \end{pmatrix} &= \begin{pmatrix} 1 & -\frac{|\hat{Q}_{x_0y_0w_1}|}{|\hat{Q}_{x_0y_0w_2}|} \\ 1 & -\frac{|\hat{Q}_{x_1y_0w_1}|}{|\hat{Q}_{x_1y_0w_2}|} \end{pmatrix}^{-1} \begin{pmatrix} \hat{p}(w_1 | x_0, y_0) - \frac{|\hat{Q}_{x_0y_0w_1}|}{|\hat{Q}_{x_0y_0w_2}|} \hat{p}(w_2 | x_0, y_0) \\ \hat{p}(w_1 | x_1, y_0) - \frac{|\hat{Q}_{x_1y_0w_1}|}{|\hat{Q}_{x_1y_0w_2}|} \hat{p}(w_2 | x_1, y_0) \end{pmatrix}, \\ \begin{pmatrix} \hat{p}(w_1 | u_2) \\ \hat{p}(w_2 | u_2) \end{pmatrix} &= \begin{pmatrix} 1 & -\frac{|\hat{Q}_{x_0y_0w_1}|}{|\hat{Q}_{x_0y_0w_2}|} \\ 1 & -\frac{|\hat{Q}_{x_1y_1w_1}|}{|\hat{Q}_{x_1y_1w_2}|} \end{pmatrix}^{-1} \begin{pmatrix} \hat{p}(w_1 | x_0, y_0) - \frac{|\hat{Q}_{x_0y_0w_1}|}{|\hat{Q}_{x_0y_0w_2}|} \hat{p}(w_2 | x_0, y_0) \\ \hat{p}(w_1 | x_1, y_1) - \frac{|\hat{Q}_{x_1y_1w_1}|}{|\hat{Q}_{x_1y_1w_2}|} \hat{p}(w_2 | x_1, y_1) \end{pmatrix}, \\ \begin{pmatrix} \hat{p}(w_1 | u_3) \\ \hat{p}(w_2 | u_3) \end{pmatrix} &= \begin{pmatrix} 1 & -\frac{|\hat{Q}_{x_0y_1w_1}|}{|\hat{Q}_{x_0y_1w_2}|} \\ 1 & -\frac{|\hat{Q}_{x_1y_0w_1}|}{|\hat{Q}_{x_1y_0w_2}|} \end{pmatrix}^{-1} \begin{pmatrix} \hat{p}(w_1 | x_0, y_1) - \frac{|\hat{Q}_{x_0y_1w_1}|}{|\hat{Q}_{x_0y_1w_2}|} \hat{p}(w_2 | x_0, y_1) \\ \hat{p}(w_1 | x_1, y_0) - \frac{|\hat{Q}_{x_1y_0w_1}|}{|\hat{Q}_{x_1y_0w_2}|} \hat{p}(w_2 | x_1, y_0) \end{pmatrix}, \\ \begin{pmatrix} \hat{p}(w_1 | u_4) \\ \hat{p}(w_2 | u_4) \end{pmatrix} &= \begin{pmatrix} 1 & -\frac{|\hat{Q}_{x_0y_1w_1}|}{|\hat{Q}_{x_0y_1w_2}|} \\ 1 & -\frac{|\hat{Q}_{x_1y_1w_1}|}{|\hat{Q}_{x_1y_1w_2}|} \end{pmatrix}^{-1} \begin{pmatrix} \hat{p}(w_1 | x_0, y_1) - \frac{|\hat{Q}_{x_0y_1w_1}|}{|\hat{Q}_{x_0y_1w_2}|} \hat{p}(w_2 | x_0, y_1) \\ \hat{p}(w_1 | x_1, y_1) - \frac{|\hat{Q}_{x_1y_1w_1}|}{|\hat{Q}_{x_1y_1w_2}|} \hat{p}(w_2 | x_1, y_1) \end{pmatrix} \end{aligned}$$

で与えられる. このとき,

$$\hat{S}_{x_0y_0w_1} = \begin{pmatrix} 1 & \hat{p}(w_1 | u_1) \\ 1 & \hat{p}(w_1 | u_2) \end{pmatrix}, \quad \hat{S}_{x_0y_1w_1} = \begin{pmatrix} 1 & \hat{p}(w_1 | u_3) \\ 1 & \hat{p}(w_1 | u_4) \end{pmatrix}.$$

と置けば,  $\hat{Q}_{x_0y_0w_1} \hat{S}_{x_0y_0w_1}^{-1}$  と  $\hat{Q}_{x_0y_1w_1} \hat{S}_{x_0y_1w_1}^{-1}$  の第1行はそれぞれ  $(p(u_1), p(u_2))$  と  $(p(u_3), p(u_4))$  の一致推定量を与える.

さらに, これらの推定量は次の漸近正規性を持つ:

**定理 4.2.**  $i, j \in \{0, 1\}$  に対して,  $\boldsymbol{\theta}_{ij} = (\theta_1^{ij}, \theta_2^{ij}, \theta_3^{ij}, \theta_4^{ij}, \theta_5^{ij})^\top = (p(w_1 | x_i, y_j), p(w_2 | x_i, y_j), p(z_1, w_1 | x_i, y_j), p(z_1, w_2 | x_i, y_j), p(z_1 | x_i, y_j))^\top$  かつ  $T_{ij}^{(n)} = (\hat{p}(w_1 | x_i, y_j), \hat{p}(w_2 | x_i, y_j), \hat{p}(z_1, w_1 | x_i, y_j), \hat{p}(z_1, w_2 | x_i, y_j), \hat{p}(z_1 | x_i, y_j))^\top$  と置く.  $\sqrt{n} (T_{ij}^{(n)} - \boldsymbol{\theta}_{ij})$  が漸近的に平均ゼロベクトル, 共分散行列  $\Sigma_{ij}$  の同時正規分布に従うとき, 定理 4.1 の仮定の下で,  $\sqrt{n} (\hat{S}_{x_0y_0w_1}^{-\top} \hat{Q}_{x_0y_0w_1}^\top \mathbf{e} - (p(u_1), p(u_2))^\top)$  は漸近的に平均ゼロベクトル, 共分散

散行列

$$\begin{pmatrix} \frac{\partial S_{x_0 y_0 w_1}^{-\top} Q_{x_0 y_0 w_1}^{\top} \mathbf{e}}{\partial \boldsymbol{\theta}_{00}^{\top}} \\ \frac{\partial S_{x_0 y_0 w_1}^{-\top} Q_{x_0 y_0 w_1}^{\top} \mathbf{e}}{\partial \boldsymbol{\theta}_{10}^{\top}} \\ \frac{\partial S_{x_0 y_0 w_1}^{-\top} Q_{x_0 y_0 w_1}^{\top} \mathbf{e}}{\partial \boldsymbol{\theta}_{11}} \end{pmatrix}^{\top} \begin{pmatrix} \Sigma_{00} & O & O \\ O & \Sigma_{10} & O \\ O & O & \Sigma_{11} \end{pmatrix} \begin{pmatrix} \frac{\partial S_{x_0 y_0 w_1}^{-\top} Q_{x_0 y_0 w_1}^{\top} \mathbf{e}}{\partial \boldsymbol{\theta}_{00}^{\top}} \\ \frac{\partial S_{x_0 y_0 w_1}^{-\top} Q_{x_0 y_0 w_1}^{\top} \mathbf{e}}{\partial \boldsymbol{\theta}_{10}^{\top}} \\ \frac{\partial S_{x_0 y_0 w_1}^{-\top} Q_{x_0 y_0 w_1}^{\top} \mathbf{e}}{\partial \boldsymbol{\theta}_{11}} \end{pmatrix},$$

の同時正規分布に従い、 $\sqrt{n} \left( \widehat{S}_{x_0 y_1 w_1}^{-\top} \widehat{Q}_{x_0 y_1 w_1}^{\top} \mathbf{e} - (p(u_3), p(u_4))^{\top} \right)$  は漸近的に平均ゼロベクトル、共分散行列

$$\begin{pmatrix} \frac{\partial S_{x_0 y_1 w_1}^{-\top} Q_{x_0 y_1 w_1}^{\top} \mathbf{e}}{\partial \boldsymbol{\theta}_{01}^{\top}} \\ \frac{\partial S_{x_0 y_1 w_1}^{-\top} Q_{x_0 y_1 w_1}^{\top} \mathbf{e}}{\partial \boldsymbol{\theta}_{10}^{\top}} \\ \frac{\partial S_{x_0 y_1 w_1}^{-\top} Q_{x_0 y_1 w_1}^{\top} \mathbf{e}}{\partial \boldsymbol{\theta}_{11}} \end{pmatrix}^{\top} \begin{pmatrix} \Sigma_{01} & O & O \\ O & \Sigma_{10} & O \\ O & O & \Sigma_{11} \end{pmatrix} \begin{pmatrix} \frac{\partial S_{x_0 y_1 w_1}^{-\top} Q_{x_0 y_1 w_1}^{\top} \mathbf{e}}{\partial \boldsymbol{\theta}_{01}^{\top}} \\ \frac{\partial S_{x_0 y_1 w_1}^{-\top} Q_{x_0 y_1 w_1}^{\top} \mathbf{e}}{\partial \boldsymbol{\theta}_{10}^{\top}} \\ \frac{\partial S_{x_0 y_1 w_1}^{-\top} Q_{x_0 y_1 w_1}^{\top} \mathbf{e}}{\partial \boldsymbol{\theta}_{11}} \end{pmatrix}$$

の同時正規分布に従う。

定理 4.2 はデルタ法 [e.g., van der Vaart, 1998] により簡単に証明できる。

#### 4.4 数値実験

ここでは、前節で提案した推定法の性質を調べるために、原因の確率の文脈で議論されてきた必要十分性の確率 (probability of necessity and sufficiency, PNS) と因果リスク差  $p(Y_{x_1} = y_1) - p(Y_{x_0} = y_1) = p(u_2) - p(u_3)$  を通して行った数値実験の結果について述べる。図 4.2 で示された  $(X, Y, Z, W, U)$  の同時確率が表 4.1 に従って与えられるような因果ダイアグラムについて考える。 $(X, Y, Z, W)$  は観測できるが、 $U$  は観測できない状況の下で、 $p(u_2)$  の推定量  $\widehat{p}(u_2)$  と  $p(u_2) - p(u_3)$  の推定量  $\widehat{p}(u_2) - \widehat{p}(u_3)$  の性質をサンプルサイズ  $n = 500, 1000, 5000, 10000$  の設定の数値実験で検証する。この状況では、 $p(u_2) = 0.25$  かつ  $p(u_2) - p(u_3) = 0.00$  なので、 $\widehat{p}(u_2)$  と  $\widehat{p}(u_2) - \widehat{p}(u_3)$  の標本平均はそれぞれ 0.25 と 0.00 に近い値になることが期待される。表 4.2 と図 4.4 は与えられたサンプルサイズ  $n$  で 1000 回繰り返し実験したときの、 $\widehat{p}(u_2)$  および  $\widehat{p}(u_2) - \widehat{p}(u_3)$  の基本統計量と箱型図を表す。 $\widehat{Q}_{xyw}$  の行列式がゼロに近

いとき、定理 4.1 の条件 (3) は満たしていないことが懸念されるため、 $\hat{Q}_{xyw}$  の行列式の絶対値が  $10^{-5}$  未満になった場合は除いて表 4.2 と図 4.4 を作成した。図 4.4 の水平線は  $p(u_2)$  と  $p(u_2) - p(u_3)$  の真の値を示している。ランダム化実験を考えているため、提案手法を用いることなく  $p(u_2) - p(u_3)$  を精度よく推定することができるが、ここでは、提案手法の性質を調べるために、 $p(u_2) - p(u_3)$  を提案手法により推定する。

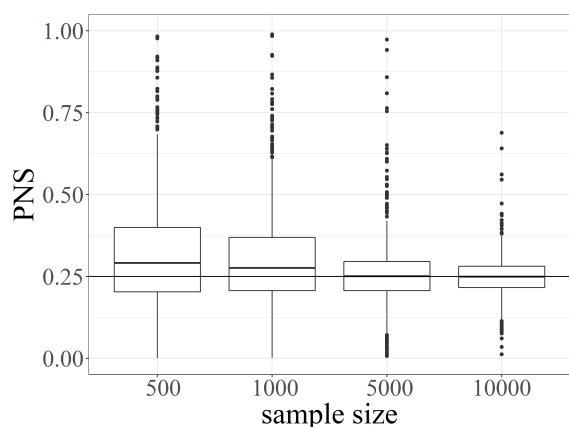
表 4.1: 数値実験のデータ生成に用いた条件付き確率。

	$p(Z U)$		$p(W U)$			$p(Y=0 X,U)$		$p(U)$	$p(X)$	
	$Z=1$	$Z=2$	$W=1$	$W=2$	$W=3$	$X=1$	$X=0$			
$U=1$	9/10	1/10	8/10	3/20	1/20	1	1	1/4	$X=0$	1/2
$U=2$	8/10	2/10	3/20	1/20	8/10	0	1	1/4		
$U=3$	2/10	8/10	7/10	2/10	1/10	1	0	1/4	$X=1$	1/2
$U=4$	1/10	9/10	2/10	1/10	7/10	0	0	1/4		

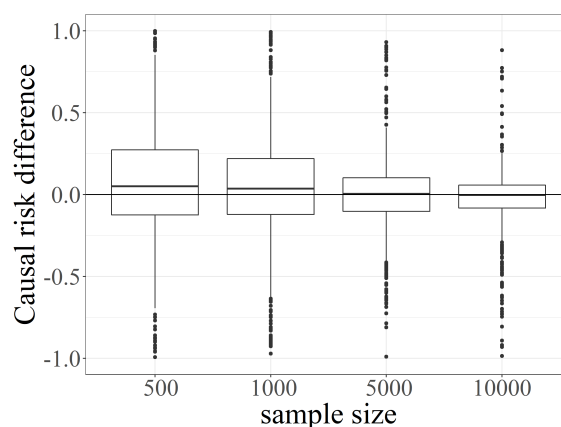
表 4.2 から、サンプルサイズが大きくなるにつれて、 $\hat{p}(u_2)$  と  $\hat{p}(u_2) - \hat{p}(u_3)$  の標本平均は真の値に近づき、標本標準誤差 (standard errors, s.e.) は小さくなることが読み取れ、提案した推定量が  $p(u_2)$  と  $p(u_2) - p(u_3)$  に対する一致推定量になっている様子が窺える。さらに、図 4.4 から、サンプルサイズが大きくなるにつれて、 $\hat{p}(u_2)$  と  $\hat{p}(u_2) - \hat{p}(u_3)$  に対する第一四分位点と第三四分位点の範囲が、真の値を含みつつ小さくなっていることがわかる。ここで、各サンプルサイズにおいて多くの外れ値が生じていることに注意する。 $\hat{Q}_{xyw}$  の行列式の絶対値が  $10^{-5}$  に近い値になるなど、定理 4.1 の条件が成り立つことを観察データから判断することが難しいときに外れ値が生じやすいためである。

表 4.2: 提案手法に基づいた推定値の基本統計量

	(a) $\hat{p}(u_2)$				(b) $\hat{p}(u_2) - \hat{p}(u_3)$			
	$n = 500$	$n = 1000$	$n = 5000$	$n = 10000$	$n = 500$	$n = 1000$	$n = 5000$	$n = 10000$
Minimum	-9.609	-15.024	-2.530	-0.480	-19.961	-14.944	-4.966	-12.621
1st Quantile	0.164	0.177	0.194	0.213	-0.157	-0.135	-0.110	-0.084
Median	0.275	0.267	0.247	0.247	0.051	0.044	0.003	-0.003
Mean	0.314	0.291	0.284	0.243	0.067	0.084	0.050	-0.026
3rd	0.406	0.369	0.295	0.280	0.325	0.255	0.110	0.059
Maximum	17.233	8.404	15.692	2.053	17.236	8.196	15.494	3.194
s.e.	1.047	0.921	0.764	0.116	1.603	1.285	0.932	0.494



(a) PNS の箱型図



(b) 因果リスク差の箱型図

図 4.4: 提案手法に基づいた推定値の箱型図

## 4.5 適用例

ここでは, Kevin Hillstrom のブログ “MineThatData” から得られる実際のマーケティングデータセットを通して第 4 部の結果を説明する [Hillstrom, 2008]. 64,000 人の (匿名化された) 顧客の情報を含むこのデータセットは, Hillstorm のウェブページ<sup>1</sup> から入手できる.

<sup>1</sup><https://blog.minethatdata.com/2008/03/minethatdata-e-mail-analytics-and-data.html>



このデータセットには、あるインターネットベースの小売業者のEメールキャンペーンの結果が含まれており、前年度の消費金額や最後に購入した時期などの基本的な特徴が含まれており、合計10個の変数が利用できる。顧客はランダムに三つのグループに分けられ、一つ目のグループは男性用商品の広告Eメールキャンペーンを受け、二つ目のグループは女性用商品の広告Eメールキャンペーンを受け、三つ目のグループはコントロールとして使用される。

以下、無作為に選ばれた女性向け商品のEメールキャンペーンを受け取った顧客と、そうでない顧客との結果の違いについて分析する。ウェブサイトに掲載されているサンプルサイズは42,613であり、関心のある変数は以下の通りである：

X: 処置 ( $x_1$ : Eメールキャンペーンを受け取る;  $x_0$ : Eメールキャンペーンを受け取らない)

Y: コンバージョン ( $y_1$ : 次の二週間に顧客が商品を購入する;  $y_0$ : 次の二週間に顧客が商品を購入しない)

Z: 新規顧客かどうか ( $z_1$ : 過去12箇月の新規顧客ではない;  $z_2$ : 過去12箇月の新規顧客である)

W: チャンネル ( $w_1$ : マルチチャンネル;  $w_2$ : ウェブ;  $w_3$ : 電話)

この状況の下で、図4.2で示されるように本調査のデータ生成過程が符号化されると仮定する。

入手できるデータセットから、 $Q_{x_0y_0w_1}$ ,  $Q_{x_0y_1w_1}$ ,  $Q_{x_1y_0w_1}$ ,  $Q_{x_1y_1w_1}$ ,  $Q_{x_0y_0w_2}$ ,  $Q_{x_0y_1w_2}$ ,

$Q_{x_1y_0w_2}, Q_{x_1y_1w_2}$  の標本推定値は

$$\begin{aligned}\widehat{Q}_{x_0y_0w_1} &= \begin{pmatrix} 0.994 & 0.121 \\ 0.494 & 0.049 \end{pmatrix}, & \widehat{Q}_{x_0y_1w_1} &= \begin{pmatrix} 0.006 & 0.001 \\ 0.004 & 0.000 \end{pmatrix}, \\ \widehat{Q}_{x_1y_0w_1} &= \begin{pmatrix} 0.991 & 0.119 \\ 0.492 & 0.047 \end{pmatrix}, & \widehat{Q}_{x_1y_1w_1} &= \begin{pmatrix} 0.009 & 0.002 \\ 0.004 & 0.001 \end{pmatrix}, \\ \widehat{Q}_{x_0y_0w_2} &= \begin{pmatrix} 0.994 & 0.437 \\ 0.494 & 0.226 \end{pmatrix}, & \widehat{Q}_{x_0y_1w_2} &= \begin{pmatrix} 0.006 & 0.003 \\ 0.004 & 0.002 \end{pmatrix}, \\ \widehat{Q}_{x_1y_0w_2} &= \begin{pmatrix} 0.991 & 0.433 \\ 0.492 & 0.222 \end{pmatrix}, & \widehat{Q}_{x_1y_1w_2} &= \begin{pmatrix} 0.009 & 0.004 \\ 0.004 & 0.002 \end{pmatrix},\end{aligned}$$

で与えられる。これらの式は定理 4.1 の条件 (1) と (3) が成り立つことを示す。このとき、定理 4.1 の条件 (2) が成り立つという仮定の下で、潜在結果変数タイプの確率  $p(u_1), p(u_2), p(u_3), p(u_4)$  を、提案した推定方法により

$$\widehat{p}(u_1) = 0.952, \quad \widehat{p}(u_2) = 0.043, \quad \widehat{p}(u_3) = 0.003, \quad \widehat{p}(u_4) = 0.002,$$

のように推定することができる。これらの確率から、PNS  $p(u_2)$  と因果リスク差  $p(u_2) - p(u_3)$  はそれぞれ  $\widehat{p}(u_2) = 0.043$  と  $\widehat{p}(u_2) - \widehat{p}(u_3) = 0.039$  のように評価される。図 4.3 は推定値  $\widehat{\mathbf{u}} = (\widehat{p}(u_1), \widehat{p}(u_2), \widehat{p}(u_3), \widehat{p}(u_4))^\top$  の箱型図を示し、表 4.3 は 95% 信頼区間を持つ PNS と因果リスク差の推定値を示す。ここで、推定値の 5,000 個のブートストラップ複製の 2.5% 分位点と 97.5% 分位点が 95% 信頼区間の導出に用いられる<sup>2</sup>。

## 4.6 考察

第 4 部では、実験研究の枠組みにおいて、共変量情報を利用することにより潜在結果タイプの確率を識別するための新たな条件と推定方法について述べた。ランダム

<sup>2</sup>正確にいうと、 $\widehat{p}(u_1), \widehat{p}(u_2), \widehat{p}(u_3), \widehat{p}(u_4)$  が  $[0, 1]$  に入り、定理 4.1 の条件 (1) と (3) が成り立つような 1,073 個の複製のみを用いた。

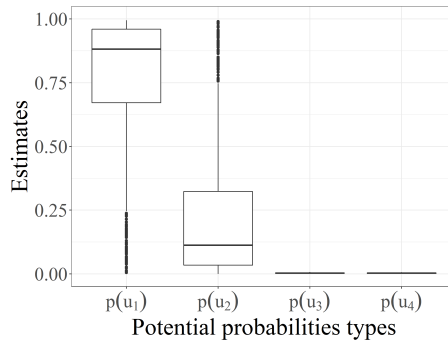


図 4.5: “MineThatData E-Mail Analytics And Data Mining Challenge” データセットにおける提案手法に基づいた推定値  $\hat{\mathbf{u}} = (\hat{p}(u_1), \hat{p}(u_2), \hat{p}(u_3), \hat{p}(u_4))^T$  の箱型図.

表 4.3: データセットにおける PNS と因果リスク差の推定値.

	Estimate (95%CI)
PNS	0.043 (0.002, 0.899)
causal risk difference	0.039 (-0.002, 0.895)

化実験のような実験研究においても，潜在結果タイプの確率を識別することは難しい問題であるが，第 3 部と同様に二つの代理変数を用いることで識別を可能にする条件を見出した．さらに，第 3 部で用いた二つの代理変数はどちらも四つの異なる値を取ることを前提としていたが，ここで提案した方法によれば，代理変数が取り得る値は二つ乃至三つ以上であればよい．また，潜在結果タイプの確率に対する一致推定量として閉じた式を提示することにより，第 3 部で述べたような最適化問題を解く手間を省くことができ，定理 4.2 のような推定量の漸近正規性を容易に導出することができる．

一方で，図 4.4 からわかるように，提案手法による推定は逆行列の計算により実用的には安定しないことが起こり得る．経験的には，(1)  $U$  と代理共変量の間に関連が弱いとき，(2) サンプルサイズが小さいとき，に不安定な推定値を引き起こしやすい傾向にある．このような推定の安定化は今後の課題である．また，第 3 部で述べ

た方法では、推定対象である確率を表す変数が  $[0, 1]$  に収まるような制約条件の下で最適化問題を解くことにより、潜在結果タイプの確率を推定していきたいが、ここで述べた提案方法で得られる推定値は  $[0, 1]$  の範囲に入るとは限らない。これについては、推定対象である確率  $p(w | u)$  のパラメタリゼーションを  $p(w | u) = 1 / \{1 + \exp(-\varphi)\}$  のように変更し、 $\varphi$  について解くことが有効であると考えられるが、このようなパラメタリゼーションによる推定精度や安定性への影響の調査についても今後の課題である。

## 第 5 部

### まとめ

本論文では共変量情報に基づいた新たな潜在結果変数の同時確率の識別および推定方法を提案した。これらの内容は、[Shingaki and Kuroki, 2021] および [Shingaki and Kuroki, 2023] において公表されている。

第 3 部では、観察研究において、共変量の代理変数を利用することで、既存研究で調べられている単調性のような強い仮定なしに、潜在結果変数の同時確率が識別可能であることを示した。特に、3.2.1 節では、共変量の一つの代理変数に加えて因果リスクが既知であれば、潜在結果変数の同時確率を識別できることを示した。観察研究において因果リスクを識別および推定するための条件や方法は既存研究でよく知られており、因果リスクが識別可能であるような状況を想定することは難しくないが、このような状況では本節で述べた一つの共変量を活用することにより潜在結果変数の同時確率についても識別および推定できることとなる。また、3.2.2 節では、因果リスクが既知でない場合でも、共変量の一つの代理変数を観測できていれば、潜在結果変数の同時確率を識別できることを示した。既存研究ではこのような場合、潜在結果変数の同時確率の取り得る値の範囲を調べることしかできなかったため、ここで述べた識別および推定法は潜在結果変数の同時確率を調べるのに非常に有用であると考えられる。

第 4 部では、実験研究においても、共変量の一つの代理変数が観測できていれば、潜在結果変数の同時確率を識別できることを示した。実験研究においても潜在

結果変数の同時確率を識別することは難しく、観察研究のときと同様、既存研究では単調性のような強い仮定がないと識別することはできなかった。ここでは実験研究においても、3.2.2節と同様の条件の下で潜在結果変数の同時確率を識別できることを示しているが、3.2.2節の状況よりも各代理変数の取り得る値の種類数を少なく設定できるほか、3.2.2節のように最適化問題を解いて推定量を得るのではなく、潜在結果変数の同時確率の推定量を具体的に構成することに成功している。

第4部で述べた方法は第3部のような観察研究の状況にも拡張できると考えられる。但し、具体的に推定量を構成できたとしても、第3部のような最適化問題を解くことで得られる推定量を用いる場合と、第4部のような具体的な推定量を構成する場合とで、推定性能に差があるかどうかは明らかではない。異なる推定方法を用いたときの推定性能の違いを明らかにすることについては今後の課題である。

なお、第3部および第4部で述べた方法はどちらも単調性が成り立つ場合（即ち  $p(u_3) = 0$  の場合）には適用することができない。これは、証明で用いた (3.5) の  $P = R^\top \Delta S$ , (3.10) の  $P_x = R_x^\top \Delta_x S$  (4.2) の  $Q_{xy} = R_{xy}^\top \Delta_{xy} S_{xyw}$  が、 $p(u_3) = 0$  のときには非可逆となるためである。単調性が成り立つ場合には、Tian and Pearl [2000] により識別可能であることが既に調べられており、そちらの結果を用いればよい。

最後に、潜在結果変数の同時確率を識別するための条件について補足する。3.2.2節や4.2節で述べた提案手法では、潜在結果タイプの確率を推定するために定理3.2や定理4.1の条件を満たす二つの代理変数を入力とする必要があるが、そのような代理変数を見つけることは難しいことではないことに注意する。 $U$  は被験者が処置に反応する仕方に影響を与えるあらゆる共変量を含むことができるため、多くの状況で  $U$  が与えられた下で条件付き独立であるような共変量  $Z$  と  $W$  の存在を仮定することができると思われる。例えば、GWAS (Genome-Wide Association Study) のように、膨大な数の共変量が観測できるような研究では定理4.1の条件を満たす代理変数を複数選ぶことが可能である。但し、共変量に対して複数の代理変数が観測できる場合に本研究で述べた方法を実際に適用する際には、どの代理変数を選択す

るかが問題となる。複数の代理変数を縮約することで推定精度を向上させるようなより良い方法が考えられるかもしれない。そのような代理共変量の選択方法についても今後の課題である。

## 謝辞

指導教官の黒木先生には、2018年4月から2024年3月の6年間にわたり、因果推論の基礎から研究活動の基本までご指導いただきました。深く感謝いたします。特に、博士論文のテーマ以外にも公平性や代替エンドポイントの評価、操作変数の一般化などの幅広い話題についても共同研究させていただき、大変貴重な経験をさせていただきました。これらの共同研究の共著者となっただいた当時黒木研究室の神田博さん、川上裕大さん、および Eli Lilly and Company の Yongming Qu 氏にも感謝いたします。

また、学位論文の博論審査等にあたり、本田先生、梶原先生、竹居先生、植木先生には大変お世話になりました。発表および博士論文に関して丁寧にご指導をいただき深く感謝申し上げます。



## 参考文献

- Joshua D. Angrist, Guido W. Imbens, and Donald B. Rubin. Identification of causal effects using instrumental variables. *J. Amer. Statist. Assoc.*, 91(434):444–455, 1996.
- Alexander Balke and Judea Pearl. Bounds on treatment effects from studies with imperfect compliance. *J. Amer. Statist. Assoc.*, 92(439):1171–1176, 1997.
- Elias Bareinboim, Juan D. Correa, Duligur Ibeling, and Thomas Icard. *On Pearl’s Hierarchy and the Foundations of Causal Inference*, page 507–556. Association for Computing Machinery, New York, NY, USA, 1 edition, 2022. ISBN 9781450395861. URL <https://doi.org/10.1145/3501714.3501743>.
- Dimitri P. Bertsekas. *Constrained Optimization and Lagrange Multiplier Methods*. Athena Scientific, 1996.
- A. P. Dawid. Causal inference without counterfactuals. *Journal of the American Statistical Association*, 95(450):407–424, 2000. doi: 10.1080/01621459.2000.10474210. URL <https://www.tandfonline.com/doi/abs/10.1080/01621459.2000.10474210>.
- Rajeev H. Dehejia and Sadek Wahba. Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *J. Amer. Statist. Assoc.*, 94(448):1053–1062, 1999.

- Constantine E. Frangakis and Donald B. Rubin. Principal stratification in causal inference. *Biometrics*, 58(1):21–29, 2002.
- Sainyam Galhotra, Romila Pradhan, and Babak Salimi. Explaining black-box algorithms using probabilistic contrastive counterfactuals. In *Proceedings of the 2021 International Conference on Management of Data*, pages 577–590, 2021.
- David Galles and Judea Pearl. An axiomatic characterization of causal counterfactuals. *Foundations of Science*, 3:151–182, 1998.
- Jesse L Goodman, John D Grabenstein, and M Miles Braun. Answering key questions about covid-19 vaccines. *Journal of the American Medical Association*, 324(20):2027–2028, 2020.
- James J. Heckman, Jeffrey Smith, and Nancy Clements. Making the most out of programme evaluations and social experiments: Accounting for heterogeneity in programme impacts. *The Review of Economic Studies*, 64(4):487–535, 1997.
- Kevin Hillstrom. The minethatdata e-mail analytics and data mining challenge. <https://blog.minethatdata.com/2008/03/minethatdata-e-mail-analytics-and-data.html>, 2008. [Online; accessed 12-Jan-2024].
- G. W. Imbens and D. B. Rubin. *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press, 2015.
- Ramaravind Kommiya Mothilal, Divyat Mahajan, Chenhao Tan, and Amit Sharma. Towards unifying feature attribution and counterfactual explanations: Different means to the same end. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, AIES '21, page 652–663, New York, NY,

- USA, 2021. Association for Computing Machinery. ISBN 9781450384735. doi: 10.1145/3461702.3462597. URL <https://doi.org/10.1145/3461702.3462597>.
- Manabu Kuroki and Zhihong Cai. Statistical analysis of ‘probabilities of causation’ using co-variate information. *Scand. J. Statist.*, 38(3):564–577, 2011.
- Manabu Kuroki and Judea Pearl. Measurement bias and effect restoration in causal inference. *Biometrika*, 101(2):423–437, 03 2014.
- Robert LaLonde. Evaluating the econometric evaluations of training programs with experimental data. *American Economic Review*, 76(4):604–20, 1986.
- T.L. Lash, T.J. VanderWeele, S. Haneuse, and K.J. Rothman. *Modern Epidemiology*. Lippincott Williams & Wilkins, 4th edition, 2021.
- Ang Li and Judea Pearl. Unit selection based on counterfactual logic. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence*, pages 1793–1799. International Joint Conferences on Artificial Intelligence Organization, 2019.
- Kung-Jong Lui. *Binary Data Analysis of Randomized Clinical Trials with Noncompliance*. John Wiley & Sons, 2011.
- Wang Miao, Zhi Geng, and Eric J Tchetgen Tchetgen. Identifying causal effects with proxy variables of an unmeasured confounder. *Biometrika*, 105(4):987–993, 08 2018. ISSN 0006-3444. doi: 10.1093/biomet/asy038. URL <https://doi.org/10.1093/biomet/asy038>.
- Judea Pearl. Direct and indirect effects. In *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence*, pages 411–420, San Francisco, CA, USA, 2001. Morgan Kaufmann Publishers Inc.

Judea Pearl. *Causality: Models, Reasoning and Inference*. Cambridge University Press, 2nd edition, 2009.

Judea Pearl. On measurement bias in causal inference. In *Proceedings of the Twenty-Sixth Conference on Uncertainty in Artificial Intelligence*, pages 425–432, Arlington, Virginia, USA, 2010. AUAI Press.

Drago Plečko and Elias Bareinboim. Causal fairness analysis. *Foundations and Trends in Machine Learning*, forthcoming.

James Robins. The analysis of randomized and non-randomized aids treatment trials using a new approach to causal inference in longitudinal studies. In L. Sechrest, H. Freeman, and A. Mulley, editors, *Health Service Research Methodology: A Focus on AIDS*, pages 113–159. U.S. Public Health Service, Washington, DC., 1989.

James Robins and Thomas S. Richardson. Alternative graphical causal models and the identification of direct effects. In Patrick E Shrout, Katherine M Keyes, and Katherine Ornstein, editors, *Causality and psychopathology: Finding the determinants of disorders and their cures*, pages 103–158. Oxford University Press, 2011.

K.J. Rothman, S. Greenland, and T.L. Lash. *Modern Epidemiology*. Lippincott Williams & Wilkins, 3rd edition, 2008.

Donald B Rubin. Estimating causal effects of treatments in randomized and non-randomized studies. *Journal of Educational Psychology*, 66(5):688–701, 1974. doi: 10.1037/h0037350.

Ryusei Shingaki and Manabu Kuroki. Identification and estimation of joint probabilities of potential outcomes in observational studies with covariate

information. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P.S. Liang, and J. Wortman Vaughan, editors, *Advances in Neural Information Processing Systems*, volume 34, pages 26475–26486. Curran Associates, Inc., 2021. URL [https://proceedings.neurips.cc/paper\\_files/paper/2021/file/dea9ddb25cbf2352cf4dec30222a02a5-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2021/file/dea9ddb25cbf2352cf4dec30222a02a5-Paper.pdf).

Ryusei Shingaki and Manabu Kuroki. Probabilities of potential outcome types in experimental studies: Identification and estimation based on proxy covariate information. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37(10):12287–12294, Jun. 2023. doi: 10.1609/aaai.v37i10.26448. URL <https://ojs.aaai.org/index.php/AAAI/article/view/26448>.

Eric Joel Tchetgen Tchetgen, Andrew Ying, Yifan Cui, Xu Shi, and Wang Miao. An introduction to proximal causal inference. *Statistical Science*, forthcoming.

Jin Tian and Judea Pearl. Probabilities of causation: Bounds and identification. *Annals of Mathematics and Artificial Intelligence*, 28(1):287–313, 2000.

A. W. van der Vaart. *Asymptotic Statistics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1998.

David S. Watson, Limor Gultchin, Ankur Taly, and Luciano Floridi. Local explanations via necessity and sufficiency: unifying theory and practice. In Cassio de Campos and Marloes H. Maathuis, editors, *Proceedings of the Thirty-Seventh Conference on Uncertainty in Artificial Intelligence*, volume 161 of *Proceedings of Machine Learning Research*, pages 1382–1392. PMLR, 27–30 Jul 2021.