

各章内に BOX などの形で収録されるはずだった文章 (ver.20240521)

目次

各章内に BOX などの形で収録されるはずだった文章 (VER. 20240521)

- X2-1 バックドア基準のよりフォーマルな解説
- X2-2 偶然によるバイアス (平均への回帰)
- X3-1 「交絡」「外生性」「交換可能性」「無視可能性」を数式でおさらいする
- X4-1 処置 T 以外の共変量の偏回帰係数はどう捉えるべきか
- X5-1 傾向スコアを用いた IPW の計算
- X7-1 操作変数法と不遵守の繋がり
- X7-2 フロントドア基準のフォーマルな説明
- X8-1 二重盲検法を構造方程式モデルでみる
- X8-2 因果ダイアグラムは効果の異質性を捉えられない？
- X8-3 介入できないものを「処置」とみなすのはアリかナシか？
- X9-1 *BIAS* 内的のバイアスは本当に最も基底的なのか？
- X9-2 構造的因果モデルから見る「固有性」と「誤差」

X2-1 バックドア基準のよりフォーマルな解説

フォーマルに記述すると、バックドア基準は以下の2つの条件から構成されます¹。

因果ダイアグラム G において、 T から Y へと有向道があるとする。このとき、次の2つの条件を満たす頂点集合 S は、 (T, Y) についてバックドア基準を満たすという。

- (1) T から S の任意の要素に有向道がない
- (2) 因果ダイアグラム G より T から出る矢線を除いたグラフにおいて、 S が T と Y を有向分離する

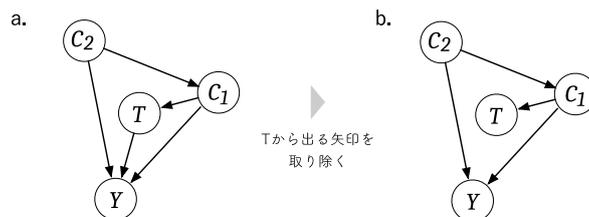
色々で見慣れない用語がありますが、今の私たちにとっては解読可能な内容です。怖気づかずに1つ1つの文の内容を見ていきましょう。

まず、ここで『 T から Y へと有向道があるとする』というのは、「池 T から水路を下流側にたどると池 Y に至る経路がある」という意味になります (T から下流側に矢印をたどったときに Y に繋がる経路を「 T から Y への有向道」と呼びます)。その次の『頂点集合 S は、 (T, Y) についてバックドア基準を満たす』というのは、重回帰分析の枠組みを念頭においた場合には、「モデルに追加した説明変数の組は、 $T \rightarrow Y$ の介入効果の推定においてバックドア基準を満たす」という意味に相当します。

では条件(1)を見ていきましょう。『 T から S の任意の要素に有向道がない』というのは、「モデルに追加した説明変数は T の下流側がない」という意味になります。この条件は、本編 2.2 節でみた「中間変数を説明変数として加えてはいけない」という一般則に対応するものです。また、「 T の下流側にある合流点を加えてはいけない」ことも含意しています。

次に、条件(2)を見ていきます。まず前半の『因果ダイアグラム G より T から出る矢線を除いたグラフにおいて』というのは、「 T から出る矢印を除いたときの因果構造において」という意味になります。たとえば、ももとの因果構造が図 X2.1a のときには、図 X2.1b が「 T から出る矢印を除いたときの因果構造」になります。後半の『 S が T と Y を有向分離する』というのは、「モデルに追加した説明変数の組により、 $(T$ と $Y)$ の上流側の共通要因からの T と Y の両方に影響を与える流れが全て遮断されている」という意味になります。ここでの「遮断されている」ということの意味内容をより詳細に補足すると、「 $(T$ と $Y)$ の上流側の共通要因からの T への有向道もしくは Y への有向道の道上において、合流点ではない変数が少なくとも1つ、説明変数としてモデルに追加されている」という意味になります。この条件(2)は、交絡を消失させるための一般則に対応するものです。条件(2)は、上流側から「 T への入力となる矢印 (T の「裏口のドア=バックドア」から入る矢印)」を通して交絡をもたらす全ての経路が、追加された説明変数の組により遮断されていることを意味しています。

図 X2.1 T から出る矢印を除いたときの因果構造



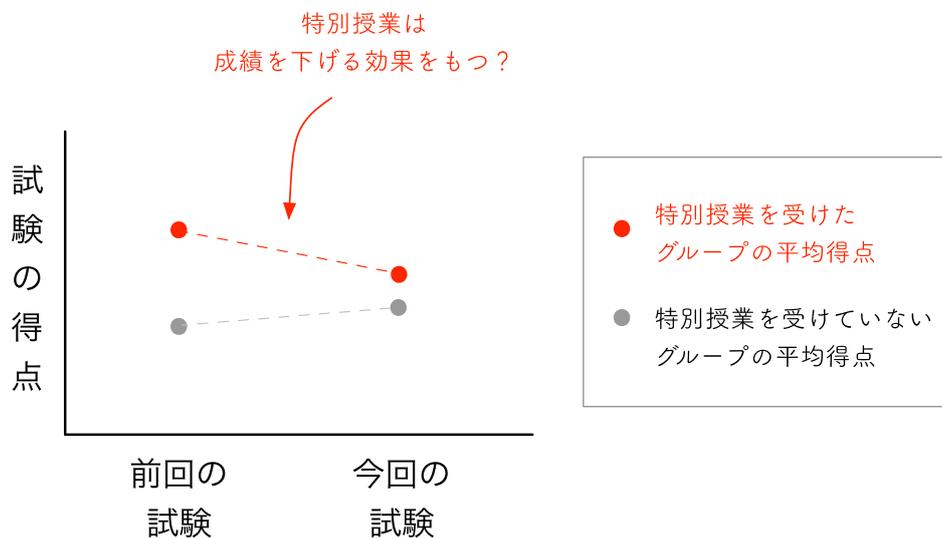
¹ 宮川雅巳(2004)『統計的因果推論—重回帰分析の新しい枠組み』朝倉書店, p82 より引用

X2-2 偶然によるバイアス（平均への回帰）

因果ダイアグラムや潜在結果モデルで表現しにくい種類のバイアスとして、偶然に起因するバイアスについても触れておきたいと思います。

あなたは高校のカリキュラム作成担当者で、試験の成績が良かった生徒に対して行う追加の特別授業を導入するかどうかを決めるために、特別授業の因果効果を知りたいと考えています。そこで、1000人の生徒を対象に行った前回の試験での上位50人に追加の特別授業を受講させ、その特別授業を受けた生徒たちの次回の試験の得点を調べてみました。その結果をまとめてみると、意外なことに、特別授業を受けた生徒たちの平均点は前回よりも下がっていることが分かりました（図X2.2）。さて、これは、特別授業には負の因果効果があることを意味するのでしょうか？

図 X2.2 平均への回帰の例



一般にこうした状況では、偶然に起因するバイアスにより、特別授業の有無に関わらず平均点の低下が起こりやすいことが知られています。なぜなら、試験の点数は多かれ少なかれ（得意な設問が出題されるか否かなど）「偶然」や「運」によって左右される部分があり、ある試験での上位の成績を取った生徒の中には、「運が良かった」ことによって、実力よりも高い点数をたまたまとれた生徒が含まれています。そうした生徒たちは、次の試験ではそこまでの幸運には恵まれない（より平均的な幸運に留まる）ことが多いため、次回の試験では「運による高得点」の分だけ、前回よりも平均的に得点が低くなりやすいわけです。

こうした「偶然によって極端な値を取ったあと、その値のような極端な値になる確率は少ない」という現象は「平均への回帰」として知られています。原理としては単純なものですが、こうしたパターンで因果効果の推定におけるバイアスが生じているケースは現実においても散見されるため、注意が必要です。

X3-1 「交絡」「外生性」「交換可能性」「無視可能性」を数式でおさらいする

本章では、無作為化について潜在結果モデルの枠組みを用いて見てきました。ここでは、潜在結果モデルの枠組みを用いて「交絡」「無視可能性 (ignorability)」「交換可能性 (exchangeability)」といった統計的因果推論における重要概念を数式でも定義をしておきたいと思います。逆に言うと、潜在結果モデルを導入することによって、これらの統計的因果推論における中核的な概念の話を明確にできるようになった、ということでもあります。

まず、「交換可能性が成り立っている」とは、処置 T が0/1の二値のみの場合、

$$\begin{aligned}P(Y^{if(T=0)}|T=1) &= P(Y^{if(T=0)}|T=0) = P(Y^{if(T=0)}), \\P(Y^{if(T=1)}|T=1) &= P(Y^{if(T=1)}|T=0) = P(Y^{if(T=1)})\end{aligned}$$

が満たされることを指します。本章で見てきたとおり、この式が成り立つときには、反事実データを観測データで代替できることが期待できるため、観測データからバイアスのない因果効果を推定できます。また、上式は、外生性（もしくは無交絡性）の数学的表現である

$$Y^{if(T=0)} \perp\!\!\!\perp T, \quad Y^{if(T=1)} \perp\!\!\!\perp T$$

を

満たすことも意味しています²。この外生性の条件（潜在結果と処置の独立性）がなりたっていない場合を「内生性がある」と表現します。また、これら交換可能性や外生性を表す式は、以下の「条件付き無視可能性 (conditional ignorability)」の条件

$$Y^{if(T=0)} \perp\!\!\!\perp T|C, \quad Y^{if(T=1)} \perp\!\!\!\perp T|C$$

における³、共変量で条件付けなくても独立が成立している（ C が空集合である⁴）という特別な場合に相当しています。バックドア基準を満たす共変量での条件付けは、この条件付き無視可能性をみたすことが知られており、バックドア基準に基づく条件付けにより観察データから因果効果を推定することが理論的に可能である（というか、因果ダイアグラムに基づく因果推論と、潜在結果モデルに基づく因果推論は、結局のところ目指している数学的条件は同じである）ことが分かります。

こうした因果ダイアグラムと潜在結果モデルの関係や、内生性と回帰分析における誤差との関係性については、オンライン補遺XA, XB章でより深掘りして解説していますので、そちらも適宜ご参照ください。

² ここで $\perp\!\!\!\perp$ は独立を現す記号である。また、本BOXの内容は、岩崎学『統計的因果推論』第3章、星野崇宏『調査観察データの統計科学』第2章、Hernan and Robins『What if: causal inference』第2章、黒木学『構造的因果モデルの基礎』第6章を参考とした。

³ バックドア基準や条件付き無視可能性の成立には、条件付き正值性 (positivity) と呼ばれる、任意の条件付けに対して任意の割付の確率がゼロにならない ($P(T|C) > 0$) という条件も必要です。

⁴ ここでの C は共変量ベクトルとしての表現として見なしてください。

X4-1 処置 T 以外の共変量の偏回帰係数はどう捉えるべきか

重回帰分析では通常、興味がある処置 T 以外の共変量の偏回帰係数も推定値として得られます。このとき、これらの偏回帰係数の値をどう捉えるべきでしょうか？

もし純粋に興味の対象が処置 T の因果効果だけである場合には、話はシンプルであり、単純に共変量の偏回帰係数は無視しましょう。ここで悩ましいのは、分析対象自体に広く興味がある場合に、処置 T 以外の共変量の”影響”についてもついつい語りたくなってしまう場合です。ここで、処置 T への興味とはまた別に、ある共変量(たとえば変数 C_A)に興味がある場合に、もし $C_A \rightarrow Y$ についてもバックドア基準が満たされているときには、 C_A の偏回帰係数を $C_A \rightarrow Y$ の因果効果のバイアスのない推定量として解釈できます。一般に、ある因果ダイアグラムにおいて「何が処置であるか」は、因果ダイアグラムの構造に依存する問題ではなく、あくまで解釈者の問題設定に従って決まる問題です。そのため、着目した共変量についてバックドア基準が満たされている場合には、その共変量を”処置”とみなして、その偏回帰係数を因果効果の推定値として解釈して構いません。

さて、ここからはおそらく大きく意見の分かれる領域に入っていきますが、もし興味のある共変量についてバックドア基準が満たされていない場合には、その共変量の偏回帰係数についてどう考えるべきでしょうか？もちろん、その偏回帰係数を、その共変量に関する因果効果の推定量として解釈するのは間違いです。とはいえ、その偏回帰係数が対象の系について何の情報も持っていないかという、そういうわけでもありません。ある共変量の(標準化)偏回帰係数が大きいことは、少なくとも、その共変量と目的変数との間に相関があることを意味しています。このことは、「その共変量もしくはその共変量と遠からず関連する何らかの要因が、目的変数のあり方と実質的な形で関連している」可能性を示唆すると考えられます。かなり漠然としていますが、もし解析者が処置 T の因果効果以外の対象の系全体のあり方やサンプル集団の特性にも興味がある場合には、こうした関連の確認からも、何らかの洞察が得られることがあるかもしれません。

この話に関連して、”Table 2 fallacy (表 2 の錯誤)”という言葉があります。この言葉は、統計的因果推論による分析がされている論文の査読において、「(ある処置 T の因果効果を推定することが解析の目的であるにもかかわらず) 本来解析の目的外である一連の共変量の偏回帰係数も論文の表 2 にまとめろ」という査読コメントがつくことが多いことを嘆く言葉です。確かに統計的因果推論の立場からは、共変量の偏回帰係数を載せたり議論したりすることは、本来の議論の幹からは外れた行為であり、その意味で、上記のような査読コメントは的外れであり、査読者の因果推論リテラシーを疑わせるものと言えます・・・とは理解しつつ、実は私は(表 2 とは言わないまでも補遺の表くらいには)一連の共変量の偏回帰係数を載せて欲しいと思うタイプの分析者だったりします。その理由は、処置以外の共変量の偏回帰係数の値には解析対象の集団の特徴に関する情報が多分に含まれており、そのサンプル集団から得られた処置の因果効果の推定を他の集団に移設することができるかどうかなどを考える際には、その情報が貴重なヒントになりうるからです。

私の意見としては、”Table 2 fallacy”の嘆きは確かに正当なものですが、その一方で、その論文の内的妥当性しか視野に入っていないある種の狭量さも感じます。分析対象の集団の情報は、解析結果の色々な部分に反映されています。統計的因果推論は確かに処置効果を推定するためのものですが、もしその結果から何らかの一般則を導き出すつもりがあるのであれば、移設可能性や外的妥当性(本編第 9 章)のヒントとなりうる数値情報をみすみすお蔵入りさせないための一定の考慮はあってもよいのではと私は思います。

X5-1 傾向スコアを用いた IPW の計算

ここでは本編の表 5.1 のデータに対して、傾向スコアを重み付けの値として用いた IPW 法を適用した計算法の例を見ていきます。本編の表 5.1 のデータに、傾向スコアと IPW（傾向スコアの逆数）の値の列を追加したものが以下になります。

表 X5-1 表 5.1 に傾向スコアを追加した表 + IPW を用いた計算の概要

個体 i	体重 C_W	年齢 C_A	投薬 T_i	駆除までの日数 Y_i	処置T=0の個体		処置T=1の個体		(共変量のインバランスがまとめて補正される) IPWでYを重み付ける	
					傾向スコア $1-e_j$	IPW $1/(1-e_j)$	傾向スコア e_j	IPW $1/e_j$	$Y^{IPW} T=0$	$Y^{IPW} T=1$
びかそ	3.7	4	0	7	0.74	1.35	(0.26)		9.42	1.35 x 7
だり	4.1	4	0	9	0.70	1.43	(0.30)		12.86	3.09 x 6
まちす	4.3	4	1	6			0.32	3.09		18.57
まぐりと	3.5	6	1	9			0.61	1.64		14.77
しゃがる	4.4	4	1	6			0.34	2.98		17.91
みろ	4.8	4	0	14	0.62	1.62	(0.38)		22.74	
あんり	5.1	6	0	23	0.21	4.68	(0.79)		107.73	
くりむと	3.6	5	0	10	0.58	1.73	(0.42)		17.35	
ごっほ	4.2	6	1	13			0.69	1.44		18.72
むんく	4.5	4	0	12	0.65	1.53	(0.35)		18.38	
ぶらつく	5.6	5	1	17			0.68	1.47		24.91
きたへふ	6.0	6	1	24			0.86	1.17		28.02
集計	$E[C_W]$ =4.5	$E[C_A]$ =4.8	$P(T)$ =6/12	$E[Y]$ =12.5		$\sum[1/(1-e_i)]$ T=0 =12.34	$E[e_i]$ =0.5	$\sum[1/e_i]$ T=1 =11.8	$\sum[Y^{IPW} T=0]$ =188.48	$\sum[Y^{IPW} T=1]$ =122.89



この表をもとに、本編表 A2.1 での割付確率の逆数を用いた重み付け (IPW) によるバランスングと全く同じ考え方で、傾向スコア e の逆数を用いて重み付けの計算していきましょう。まずは、投薬あり $T=1$ の例として「まちす」の重み付けを考えてみます。このとき傾向スコアは $e_{\text{まちす}} = 0.32$ 、その逆数は $1/e_{\text{まちす}} = 3.09$ です。この $1/e_{\text{まちす}}$ を IPW として Y を重みづけた値は $Y_{\text{まちす}}^{IPW} = 3.09 \times 6 = 18.57$ となります。一方、投薬なし $T=0$ の例

である「びかそ」の傾向スコアは $e_{\text{びかそ}} = 0.26$ です。ここで割付けられているのは「投薬なし $T = 0$ 」の方なので、重み付けでは $1 - e_{\text{びかそ}}$ の方を参照します。このとき $1 - e_{\text{びかそ}} = 0.74$ 、その逆数は $1/(1 - e_{\text{びかそ}}) = 1.35$ です。この値を IPW として Y を重みづけた値は $Y_{\text{びかそ}}^{\text{IPW}} = 1.35 \times 7 = 9.42$ となります⁵。

ここでそれぞれの個体もつ IPW の大小を見ると、「傾向スコア e_i が大きいけど $T = 0$ が割り当てられている（例：あんり）」もしくは「傾向スコア e_i が小さいけど $T = 1$ が割り当てられている（例：まちす）」場合に IPW の値が大きくなっていることが分かります。こうして実現されづらいデータに（その”実現されにくさ”に応じて）大きな重みをつけることにより、実現しなかった組み合わせのデータの”補完代入”的な計算が行われ、処置群間の共変量の分布がバランスングされます。原理としては、本編の補遺 A2 でみた Y 変量の IPW によるバランスング（図 A2.1）と同じです。

さて、算出した各個体 i の Y_i^{IPW} をそれぞれの処置ごとに足し合わせると $\sum[Y_i^{\text{IPW}}|T = 0] = 188.48$ 、 $\sum[Y_i^{\text{IPW}}|T = 1] = 122.89$ になります。ただ、これらはあくまで IPW での重み付けにより嵩増しされた量なので、単純にそれぞれの匹数（6 匹）で割ることによって平均を求めることはできません。IPW によって実質的に何匹分の量になっているかは、IPW の総和の値でみる事ができます。 $T = 0$ の群での IPW の総和は $\sum[1/(1 - e_i)|T = 0] = 12.34$ であり、 $T = 0$ 群での重み付け後の総和は「12.34 匹分」の量に相当します。同様に、 $T = 1$ の群での IPW の総和は $\sum[1/e_i|T = 1] = 11.8$ であり、 $T = 1$ 群での重み付け後の総和は「11.8 匹分」の量に相当します。これらの量で Y_{IPW} の総和を割ることにより、IPW で重み付けされた平均値を計算できます。また、ここでは IPW での重み付けにより共変量がバランスングされているため、それらの平均値は元々のサンプル集団の潜在結果の平均値 $E[Y^{\text{if}(T)}]$ として解釈できます。式でまとめると：

$$E[Y^{\text{if}(T=0)}] = \sum[Y_i \times (1/(1 - e_i))] / \sum[1/(1 - e_i)] = 15.26 \quad \leftarrow \text{「T=0」の処置グループにおける、IPW で重み付けた Y の総和 / IPW の総和}$$

$$E[Y^{\text{if}(T=1)}] = \sum[Y_i \times (1/e_i)] / \sum[1/e_i] = 10.41 \quad \leftarrow \text{「T=1」の処置グループにおける、IPW で重み付けた Y の総和 / IPW の総和}$$

となります。これらの潜在結果における平均値の差が因果効果の推定値となるので

$$E[Y^{\text{if}(T=1)}] - E[Y^{\text{if}(T=0)}] = 10.42 - 15.27 = -4.846$$

となり、投薬により駆除までに要する日数が 4.86 日ぶん短くなっていると推定されました（真の値は 5.0 日）。単純に観察された処置群間の差を見ると $E[Y|T = 0] - E[Y|T = 1] = 0$ 日であり、傾向スコアを用いた IPW により、バランスングにより推定のバイアスがかなり改善したことが分かります。

さて、今回は「真の因果効果の値があらかじめ分かっている」という仮想的な状況で説明をしてきましたが、

⁵ この辺りの計算の数字、小数点以下 2 桁の表示と、実際の数字の端数の関係で計算が微妙に合っていないですが、単なる端数計算による差異なのであまり気にしないいただければと思います。すみません。

実際の統計解析では「真の因果効果」があらかじめ分かっていることは稀です。そのため、推定のバイアスが改善したかを「真の因果効果」と比較して確かめることはできません。しかし、次善の策として、傾向スコア法では共変量のバランスが改善されているかを実際にチェックすることができます。もしかしらささやかなことに感じるかもしれませんが、これはとても素晴らしいことです！上記の解析の中で、傾向スコアを用いた IPW による調整によって、どのくらい共変量のバランスが改善したかを見てみましょう。

表 X5-2 IPW による重み付けによって処置群間で共変量のバランスが改善しているのを計算で確認

個体 i	投薬 T_i	もともとの C_W		もともとの C_A		傾向スコアに基づく IPW		IPW調整後の C_W		IPW調整後の C_A	
		体重 $C_W T=0$	体重 $C_W T=1$	年齢 $C_A T=0$	年齢 $C_A T=1$	$1/(1-e_i)$	$1/e_i$	$C_W^{IPW} T=0$	$C_W^{IPW} T=1$	$C_A^{IPW} T=0$	$C_A^{IPW} T=1$
ぴかそ	0	3.7		4		1.35		5.00		5.40	
だり	0	4.1		4		1.43		5.86		5.72	
まちす	1		4.3		4		3.09		13.29		12.36
まぐりと	1		3.5		6		1.64		5.74		9.84
しゃがる	1		4.4		4		2.98		13.11		11.92
みろ	0	4.8		4		1.62		7.78		6.48	
あんり	0	5.1		6		4.68		23.87		28.08	
くりむと	0	3.6		5		1.73		6.23		8.65	
ごっほ	1		4.2		6		1.44		6.05		8.64
むんく	0	4.5		4		1.53		6.89		6.12	
ぶらつく	1		5.6		5		1.47		8.23		7.35
きたへふ	1		6		6		1.17		7.02		7.02
集計	$P(T) = 6/12$	$E[C_W T=0] = 4.3$	$E[C_W T=1] = 4.67$	$E[C_A T=0] = 4.5$	$E[C_A T=1] = 5.12$	$\sum[1/(1-e_i) T=0] = 12.34$	$\sum[1/e_i T=1] = 11.8$	$\sum[C_W^{IPW} T=0] = 55.61$	$\sum[C_W^{IPW} T=1] = 53.44$	$\sum[C_A^{IPW} T=0] = 60.45$	$\sum[C_A^{IPW} T=1] = 57.13$

C_W の処理群間の差は 0.37
 C_A の処理群間の差は 0.62
 共変量のバランスの崩れあり

重み付けを考慮して平均をとる
 $55.61/12.34$ $53.44/11.8$ $60.45/12.34$ $57.13/11.8$
 4.50 4.53 4.89 4.84

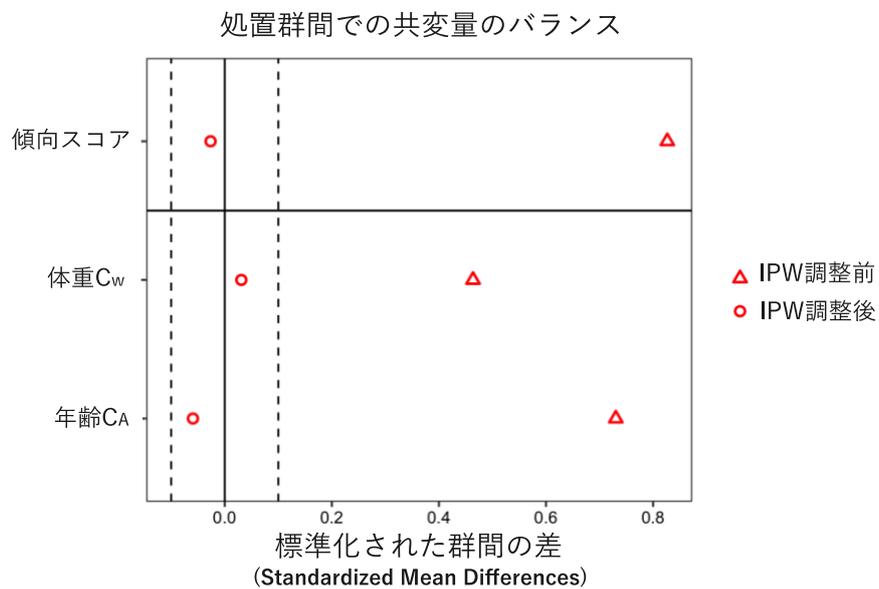
傾向スコアに基づく IPW による調整
 C_W の処理群間の差は 0.03
 C_A の処理群間の差は -0.05
 共変量のバランスが大幅に改善！

この表から、処置群間での各変量の平均値の差が IPW による調整前と調整後でどう変化したかを表しています。体重 C_W についてはおおよそ $-0.37 \text{ kg} \rightarrow 0.03 \text{ kg}$ 、年齢 C_A についてはおおよそ $-0.62 \rightarrow -0.05$ くらいまで処置グループ間での平均値の差が縮小しており、バランスがかなり改善されたことが分かります。この傾向スコア法による共変量のバランスの改善については、以下のような図で表すこともできます⁶。

⁶ この図は R の cobalt パッケージの love.plot 関数を使って作図しました。尚、図内の横軸の群間差には標準偏差で標準化した standardized mean difference が使われており、本文中での（標準化していない）数値とは少しずれています。

図 X5-2 ラブプロットによる IPW 調整前-後の共変量のバランスチェック

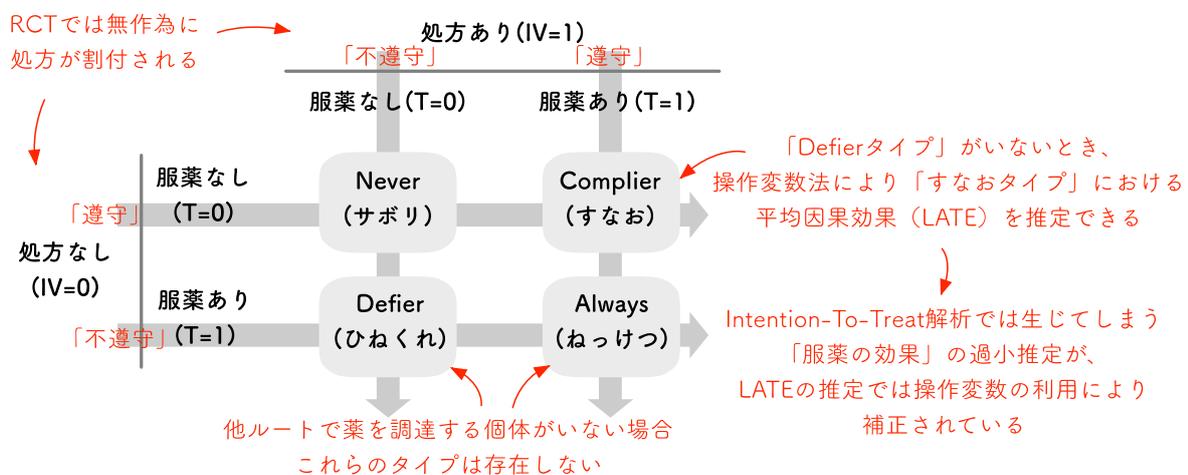
この図では、△は IPW 調整前の処置群間の差を、○は IPW 調整後の群間の差を表しています。傾向スコアに基づく IPW 重み付けによる調整を行うことで、群間での共変量のバランスが実際に大幅に改善されていることが視覚的にも確認できます。このように、共変量のバランスが実際に改善したかをデータを用いてチェックできることは、統計的因果推論の他の手法にはあまり無い、傾向スコア法ならではの美点です。



X7-1 操作変数法と不遵守の繋がり

さてちょっと意外かもしれませんが、操作変数法とRCTでの「不遵守」の問題は同じ枠組みで捉えられることが知られています。例として、猫に対して無作為にある薬の処方の有無を割り付けるタイプのRCTを考えます。本来の興味があるのは、「服薬の割り付けが遵守された状況での因果効果」です。しかし困ったことに、「処方あり」を割り付けた猫の中には、どうにも服薬してくれない「不遵守」の猫が少なからず生じているとします。「実際に服薬したかどうか」は、きちんとデータとして記録されているとします。ここで、最初の「処方の割り付け」を操作変数、服薬の有無を T 、健康アウトカムを Y として捉えることにより、操作変数法の枠組みを用いて、不遵守の問題を考えることができます (図 X7.1)。

図 X7.1 操作変数法の枠組みからみた「RCTにおける不遵守」



ここでは、処方は無作為に割り付けられているため、Intention-to-Treat 解析 (本編 3.4.3 節参照) による因果効果 ($IV \rightarrow Y$ の因果効果。ここでは ITT と表記します) は、各層で観察されたアウトカム Y の差分、すなわち:

$$ITT = E(Y|IV = 1) - E(Y|IV = 0)$$

と計算できます。しかし、この ITT はあくまで処方意図に基づく $IV \rightarrow Y$ の因果効果であり、薬の服薬 $T \rightarrow Y$ の因果効果ではありません。この例では、処方あり ($IV = 1$) の中にも処方されても実際には服薬をしていない猫の個体が含まれるので、上記の ITT は実際の服薬の因果効果と比較して一般に過小評価になります。このとき、実際に服薬したかどうかのデータがある場合には、以下の服薬 T の有無による層別の差を見ることができます。

$$\text{服薬の有無による健康アウトカムの差} = E(Y|T = 1) - E(Y|T = 0)$$

しかしこの量は、 $T \rightarrow Y$ に対して閉じられていないバックドアパスがありうるため⁷、 $T \rightarrow Y$ の因果効果の推定量としてはバイアスを含みうるものとなります（本編 3.4.3 節参照）。

ここで次善の策を考えましょう。処方 IV は操作変数の条件を満たしているため、服薬の因果効果の ATE は計算できなくとも、服薬の因果効果の LATE、つまり「すなおタイプ」における因果効果に限れば、操作変数法の考え方で計算できます。具体的には、処方の割り付けを操作変数として用いることで、 $T \rightarrow Y$ の因果効果を以下の式で計算できます。

$$\text{LATE} = \frac{E[Y|IV = 1] - E[Y|IV = 0]}{E[T|IV = 1] - E[T|IV = 0]} = \frac{\text{ITT}}{E[T|IV = 1] - E[T|IV = 0]}$$

この LATE の式は、操作変数法における LATE の推定式である本編の式 7.3 と一見違う形に見えるかも知れませんが、この分子は「 IV が 1 変化したときの Y の平均変化量」を意味しており、回帰直線の傾き（ IV が 1 単位変化したときの Y の変化量の推定値）である $\hat{\beta}_{IVonY}$ に対応しています。同様に、分母は「 IV が 1 変化したときの T の平均変化量」であり、回帰直線の傾き $\hat{\beta}_{IVonT}$ にきちんと対応していることが分かるかと思います。この薬の効果の例では、もともとの興味の対象が「服薬の割り付けが遵守された状況での RCT による因果効果」なので、「すなおタイプ」に対する効果である LATE は、もともとの興味の対象と近く、概念解釈上も筋の良い推定となっていると言えます。

⁷ たとえば、食欲がないと服薬しないケースが生じる場合には、「服薬 $T \leftarrow$ 食欲 \rightarrow 健康アウトカム Y 」となり食欲による交絡が生じる

X7-2 フロントドア基準のフォーマルな説明

本編の 7.2.2 節で扱ったフロントドア基準について、黒木(2017) ⁸p110 からのフォーマルな定義を引用します (各変数の記号の表記は本書の文脈に対応させるために変更しています)。

非巡回的有向グラフ G においては T は Y の非子孫であるとする。このとき、次の 3 条件を満たす頂点集合 M は (T, Y) についてフロントドア基準を満たすという。

- (1) T から Y への任意の有向道に M の要素が存在する
- (2) G より T から出る矢線 ($T \rightarrow$) をすべて除いたグラフにおいて、空集合は T と M を有向分離する
- (3) G より M の任意の要素から出る矢線 ($M \rightarrow$) をすべて除いたグラフにおいて、 T は M と Y を有向分離する

これを意識すると、中間変数 M が一つの場合には、背景にある因果構造において、 $T \rightarrow M$ のバックドア基準が満たされており (上記(2)に対応)、また、 T を条件づけたときには $M \rightarrow Y$ のバックドア基準が満たされている (上記(3)に対応)、という 2 段階のバックドア基準が満たされていることが条件として示されています。

⁸ 黒木学(2017)『構造的因果モデルの基礎』共立出版

X8-1 二重盲検法を構造方程式モデルでみる

本編の8.1節の二重盲検法の例を、数式でもおさらいしてみましょう。まず、本編の図8.3の素朴な因果ダイアグラムに対応した構造方程式は

$Y = \beta_{\text{処方}T} T + \epsilon$ ←Yは、処方Tを受けたとき ($T = 1$)、処方による効果と、その他要因 ϵ により決定される

となります。ここで「処方あり」が $T = 1$ 、「処方なし」が $T = 0$ であり、 $\beta_{\text{処方}T}$ は「処方 $T \rightarrow$ 健康指標 Y 」の因果効果になります。ここで ϵ は、 T と独立な未観測要因からの影響を表します。一方、図8.1.3右の効果の媒介経路の内訳を考慮して書くと：

$$Y = (\beta_{\text{薬Aの生理的効果}} + \beta_{\text{プラセボ効果}} + \beta_{\text{観測者効果}}) T + \epsilon$$

←Yは、薬Aの処方を受けたとき ($T=1$)、薬の生理的効果、プラセボ効果、観測者効果、その他要因 ϵ により決定される

と書けます（説明の便宜のため、ここでは服薬は処方 T の意図通りに行われている状況を考えます）。ここで、処置が単に「処方あり／なし」のままだと、「処方の有無」によって生じる差 ($\beta_{\text{薬Aの生理的効果}} + \beta_{\text{プラセボ効果}} + \beta_{\text{観測者効果}}$) は「薬の生理的効果の有無」だけでなく「プラセボ効果」と「観測者効果」も含んでしまうため、そもそもの興味の対象である「薬の生理的効果 ($\beta_{\text{薬Aの生理的効果}}$)」のみを分離した推定にはなりません。そのため、処置を「本物の薬」と「プラセボ」に変えてみます。薬Aの場合の処置を T_{DrugA} として

$$Y = (\beta_{\text{薬Aの生理的効果}} + \beta_{\text{プラセボ効果}} + \beta_{\text{観測者効果}}) T_{\text{DrugA}} + \epsilon$$

←Yは、薬Aの処方を受けたとき ($T_{\text{DrugA}}=1$)、薬の生理的効果、プラセボ効果、観測者効果、その他要因 ϵ により決定される

と表すと、薬Aが処方された場合 ($T_{\text{DrugA}} = 1$) の結果 Y は

$$Y = (\beta_{\text{薬Aの生理的効果}} + \beta_{\text{プラセボ効果}} + \beta_{\text{観測者効果}}) + \epsilon$$

←Yは、薬の生理的効果、プラセボ効果、観測者効果、その他要因 ϵ により決定される

となります。一方、プラセボの場合の処置を T_{Placebo} とすると、薬の生理的効果はゼロと考えられるので、

$$Y = (\beta_{\text{プラセボ効果}} + \beta_{\text{観測者効果}}) T_{\text{Placebo}} + \epsilon$$

←Yは、プラセボの処方を受けたとき ($T_{\text{Placebo}} = 1$)、プラセボ効果、観測者効果に影響を受ける

と表され、プラセボが処方された場合 ($T_{\text{Placebo}} = 1$) の結果 Y は、

$$Y = (\beta_{\text{プラセボ効果}} + \beta_{\text{観測者効果}}) + \epsilon$$

←Yは、薬の生理的効果、プラセボ効果、観測者効果、その他要因 ϵ により決定される

となります。ここで、患者にも観測者にも見分けがつかないように薬Aとプラセボが処方されているときには、それぞれの異なる処方を受けたサブグループの間で、 $\beta_{\text{プラセボ効果}}$ および $\beta_{\text{観測者効果}}$ の項に差は生じないものと期待できます。そのため、各サブグループ間での平均の差分を取ると、

$$E(Y|T_{\text{DrugA}} = 1) - E(Y|T_{\text{Pracebo}} = 1) = \beta_{\text{薬Aの生理的効果}}$$

←「薬Aを処方した集団のYの平均」と「プラセボを処方した集団のYの平均」の差は「薬Aの生理的効果」に等しい

となり、薬Aの生理作用を介した平均因果効果を推定できると期待できます。

X8-2 因果ダイアグラムは効果の異質性を捉えられない？

一般に、効果の異質性について因果ダイアグラムで描くことは困難です。なぜなら、因果ダイアグラムにおける効果はシンプルに一本の矢線で表現されるため、効果の凹凸まではうまく表すことができないからです。しかし、別の見方をすると、効果の異質性がある場合は「一枚の因果ダイアグラムでは描けない」という言い方もできるかもしれません。「効果に異質性がある」ことは、サンプル集団内に実質的に「(定性的・定量的に)異なる生成メカニズムをもつサブグループが混在している」ことを含意しているからです。複数の異なる生成メカニズムが存在しているとき、その記述には複数の因果ダイアグラムが必要となり、効果に異質性がある状況は、それらの複数の生成メカニズムが併存しながらデータが生成されている状況として(原理的には)描きうると考えられます。こうした場合に、異質性の背後にある複数の生成メカニズムの構造を複数の因果ダイアグラムを描いて検討することで、なぜ、どのように効果の異質性が生じているのかについての手がかりが得られることもあります。

X8-3 介入できないものを「処置」とみなすのはアリかナシか？

「介入なくして因果なし (no causation without manipulation)」の原則に基づき、統計的因果推論においては、介入できないものを処置としてみなすことは否定的に語られる傾向があります⁹。例えば、ある業界での賃金の男女差について、「性別」を処置としてその「性別→賃金」の因果効果を推定することを考えてみましょう。一般に、「性別」は介入の対象とはなりえません。例えば頭痛薬を投与したり投与しなかったりするには、個体に対して男性にしたり女性にしたりする介入を行うことは困難です¹⁰。ひとまずこの問題を「介入可能性」の問題と呼びます。もう一つの問題として、実際に介入することが考えにくい場合に、処置としての「性別」が分析概念として一体何を表しているのかが不明瞭なままになることが挙げられます。これは SUTVA 条件に関する問題であり、「性別」という分析概念が含意している versions of treatment --- 社会的役割期待の話なのか、性染色体の話なのか、性的指向性の話なのか、性自認の話なのか、もっと他の側面のものなのか --- が不明瞭となるという問題です。この「介入可能性がないこと」と「実際の介入を伴わないために version of treatment が不明瞭なままとなること」により、「この世界の反事的世界として、どの世界を対置したらよいのか」が不明瞭になってしまうため、因果効果を明確に定義することができない状態になります(本編 BOX 8.2)。これらのことを考えると、介入できないものを処置とみなすのを禁ずることには相応の理由があります。

しかし --- これは論者によって見解が分かれると思いますが --- 介入できないものを、あたかも外形上は「処置」であるかのようにみなして解析することに、学術的に一定の意味がある場合もあると私は考えています。たとえば、性別以外の全ての想定しうる共変量を揃えた上で、異なる性別群間での結果を比較したとき、もしその結果に少なくない群間差がみられた場合には、そこには何らかの「語られるべき issue の存在」が示されているのではないのでしょうか。統計的因果推論的には、ここでの「性別」は処置として ill-defined かもしれませんが、そのことと学術的分析としての有意義さは本質的に別の問題です。ここで生じている問題はむしろ、上記の群間差を「因果効果」と呼ぶことによる単なる概念的混乱であるように思います。統計的因果推論(特に潜在結果モデル)の枠組みの観点からは、「介入なくして因果なし」の立場に立つことには相応の理由がありますが、潜在結果モデルの方法論的な枠組みに整合しているかどうかという話と、「介入できないものを処置として見なして解析する」ことが学術的あるいは社会的な問いとして重要でありうるかは別の話であり、それらの異なる水準の話を混同するべきはない、と私は考えています。

⁹ Holland (1986) Statistical and causal inference. Journal of the American Statistical Association. Vol. 81, pp945-960. Imbens and Rubin (2015) など参照。

¹⁰ 介入可能性も versions of treatment も明確なケースとしては、例えば書面上だけで行われる審査に対して、同一のプロファイルで「性別」の記載だけを変えた書類を作成し、その記載の差によって結果の差をみるような実験デザインの場合には、「性別」は明確に(操作的に)定義可能、かつ介入可能であると言えます。

X9-1 $BIAS_{\text{内的}}$ のバイアスは本当に最も基底的なのか？

本編の9.1.4節では、手持ちのデータからターゲット集団における因果効果の推定を行う際のバイアスの内訳を以下のように整理しました。

$$BIAS_{\text{移設}} = BIAS_{\text{移設元と移設先}} + BIAS_{\text{サンプリング}} + BIAS_{\text{内的}}$$

しばしば、これらのバイアスの中で $BIAS_{\text{内的}}$ が最も基底的であり、 $BIAS_{\text{内的}}$ が小さいこと (= 内的妥当性) を満たした推定結果のみが、その次の「一般化可能性」や「移設可能性」の検討に値するものである、と主張されることがあります。しかし、たとえば、サンプル集団とターゲット集団の特性がかけ離れており、 $BIAS_{\text{移設元と移設先}} \gg BIAS_{\text{内的}}$ となっている場合には、内的妥当性の高さは推論全体における妥当性の担保には実質的にはほとんど貢献しません。ターゲット妥当性や Evidence-Based Practices (エビデンスに基づく実践) の観点からは、上記のバイアスの各項はそれぞれ平行な重要性をもつものであり、 $BIAS_{\text{内的}}$ をもっとも基底的な項であると見なす主張には論理的な根拠が乏しいと筆者は考えています。

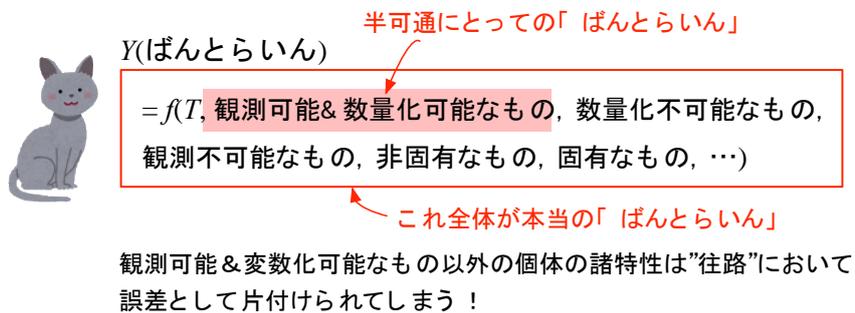
尚、研究者共同体での論文の査読基準において内的妥当性が重要視されることには (個々の論文は必ずしもターゲット妥当性の観点から評価されるわけではないので)、一定の妥当性があると思われます。しかし、個々の論文の査読評価において求められる妥当性境界と、EBPs において求められる妥当性境界は当然異なるので、前者の妥当性境界に基づき EBPs の文脈での妥当性を主張することは短絡的な論理であると言えます。

X9-2 構造的因果モデルから見る「固有性」と「誤差」

本編の9.2節で議論した「法則性」と「固有性」と「誤差」の関係について、構造的因果モデルの枠組み（詳しくはオンライン補遺XA章参照）をもとに、少し別の角度から考察してみたいと思います。

ここでは、「ばんとらいん」という名前の一匹の猫の個体を考えます。着目する結果としては、皮膚病の治癒までの日数 Y を考えます。理念的な話として、ここで観察される結果である「 Y (ばんとらいん)」が生成される過程には、一般に非常に多くの要因 --- 原理的には、定量的なもの、非定量的なもの、観測可能なもの、非観測可能なもの、他の個体とは概念的に交換されようがない固有の文脈に属するようなもの、あるいはいかなる文脈にも回収されえない断片的なもの --- などの諸々が関わりうると考えられます（図X9.1）。

図X9.1 「ばんとらいんの Y 」を構成する諸特性のイメージ



これらの諸要因のうち、測定可能な要因はごくわずかです。ここで測定可能な共変量を集めたところで、測定された共変量群 C_1, \dots, C_J だけで表現されうるものは --- もしかしたらデータ分析者は全てを知ったつもりになっているかもしれませんが --- それは「半可通にとっての『ばんとらいん』」でしかなく、「固有性の世界」にいる「真の『ばんとらいん』」の全てを反映した表現では決してありません。

ここで、こうしたそれぞれの個体 i の $Y(i)$ をもとに、多くの個体からの測定値が集計されたとき、「測定された共変量群 C_1, \dots, C_J で表現されうるもの」以外のものに起因する Y の変動は、「誤差項」と呼ばれるものの中に包摂されます。この意味で、「誤差項」とは、“測定”という切り取り行為では捉えることが叶わなかった、「固有性の世界」に属する「測定されえなかったものの断片」が、うず高く積もったものであると言えます。

往々にして、「誤差」は取り上げるに足らない「ノイズやガレキの山」のように扱われることがあります。しかし、この観点から言えば、その一見「ガレキの山」のように見えるものは、「固有性の世界」では意味をもっていた（し、そこでは「かけがえのない意味」をもっていたかもしれない）ものの集積としての「山」と捉えるべきものなのかもしれません（図X9.2）。

図 X9.2 構造的因果モデルの枠組みから見た「法則性」と「固有性」と「誤差」に関する諸特性のイメージ

