

複数の交絡変数がある場合のプロキシバイアスに関する考察

樊怡舟（広島大学）・中尾走（広島市立大学）

1 序論

本研究は、直接、測定不可能な交絡変数が複数に存在する場合に、それらを一つの代理指標にまとめて統制することの是非を、因果推論のバイアスという視点から検討するものである。

社会科学の研究では、構成概念を取り扱う場合が多く、その際に概念そのものがそもそも完璧に測定できないものとなっており、何らかの操作的定義に従い、代理指標を用いてあくまで間接的に分析を行なっている。つまり、代理指標と構成概念の間に隔たりが存在しており、何らかの測定誤差が入っていると考えられる。従来、誤差を抑制し、より精度の高い代理指標を使用することが精緻な推論につながるとされており、代理指標を評価する際にも主に「測りたいものが適切に測れているか」という妥当性の議論が心理学を中心に展開されてきた（Streiner et al., 2015=2016）。

一方、構成概念と代理指標の不一致を受け入れながら、測定誤差が発生している代理指標の使用がどのようなバイアスを伴っているかという感度分析（Sensitivity Analysis）の視点から、分析結果の信頼性について改めて検討する必要があると考えられる。

加えて、中尾ほか（2022）は、交絡変数の代理指標（proxy）が従属変数の誤差項と従属非差異な関係（Dependence nondifferential）（VanderWeele & Hernán 2012）を持つ場合に、交絡変数の部分的統制（partial control）のほか、合流点バイアスも同時に引き起こすことを指摘した。そのうえで、交絡変数が1つというモデルに限定し、代理指標の測定誤差の抑制が必ずしも正確な因果効果の推定につながっていないことを明らかにした。

しかしながら、現実の観察データを用いた因果推論では、交絡変数が複数あることが一般的であろう。以下では具体的に、交絡変数が複数であり、かつプロキシの合流点バイアス問題も生じるケースを考えていく。ここでは一例として、通塾と大学進学との関係を取り上げてみよう。この例の場合、以下の図1のような有向非巡回グラフ（Directed Acyclic Graph: DAG）が想定される。このように、通塾が大学進学にもたらす効果を推定する際に、これまでの社会学や経済学の研究成果一般を踏まえると、常識的には交絡変数として少なくとも経済資本と文化資本など、社会経済的地位（Social economics status : SES）に相当する変数が複数想定できる。しかも二つの交絡変数間もまた独立しておらず、互いに関係性が生じていると考えられる（図1ではとりあえず経済資本から文化資本に矢印を引いている）。

この場合、通塾の因果効果を求める際に、文化資本と経済資本を同時に統制する必要がある。しかし、文化資本も経済資本も完全には測定できない構成概念であるため、「本の数」や「暮らしむき」などをアンケート調査で収集し、代理指標として交絡要因の部分的統制

(partial control) 図ることが一般的である（そのほかに、代理指標として、「親の職業」、「親の年収」「親の学歴」なども考えられる）。但し、親の年収・学歴など SES に関連する何らかの代理指標を調査する場合、本来なら直接的に測ることが必要なのだが、日本の調査では、親の情報を本人に間接的に尋ねるといった代理回答が多かった。そのため、系統誤差 u が混入している可能性が指摘されており（白川 2021）、その要因として虚偽・誤回答、記憶の曖昧さ、社会的望ましきバイアスなどが考えられる。さらにそれら系統誤差 u を引き起こす要因は、同時に従属変数の誤差項にも影響を及ぼす可能性もあり、測定誤差間の従属非差異な関係（Dependence nondifferential）が想定される（VanderWeele & Hernán 2012）。

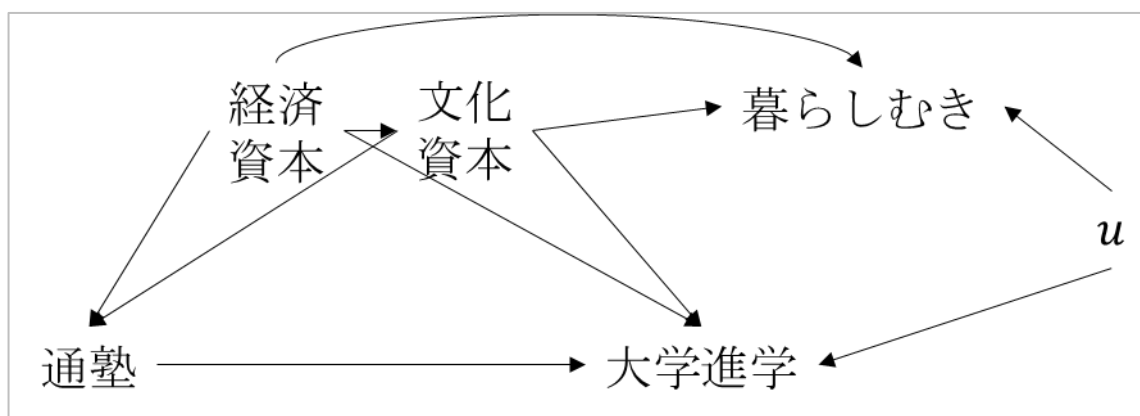


図1 想定される DAG の事例

もし u が存在していなければ、代理指標の「暮らしむき」を統制することで、本来なら測定不可能だった交絡要因に対して部分的統制（partial control）が果たされ、不完全とはいえ処置効果の推定値が真の因果効果に近づいていくと期待される。しかしながら、図1の場合、 u の存在によって、暮らしむきを統制することで、 $u \rightarrow$ 暮らしむき \leftarrow 文化資本というバックドアパスが開き、即ち合流点バイアス（collider bias）が発生する（林・黒木 2016）。つまり、「暮らしむき」の統制をめぐるジレンマが生じている。この場合、一概に代理指標を統制したほうが良いとは限らない。中尾ほか（2022）が指摘しているように、極端な条件下では、交絡変数の未統制よりも、代理指標による合流点バイアスの方が問題になる場合もある。

これまで、中尾ほか（2022）を含め、交絡変数の部分的統制の問題は、Ogburn and VanderWeele（2012）、Miao et al.（2017）、Elwert and Pfeffer（2019）などが取り上げている。しかしながら、いずれも交絡変数がある場合を想定しており、上記の事例に対応しきれない部分もある。本稿は、中尾ほか（2022）のモデルを拡張し、複数の交絡変数をもつ代理指標にまとめて統制する場合のバイアス（以下、 p -bias）の定式化とともに、代理指標の精度による p -bias の挙動について検討していく。

2 未観測交絡の局所独立性（Local Independence）と脱落変数バイアス

複数の交絡変数が想定されるモデルの難点の一つは、交絡変数間の因果関係も想定される点である。このことによって、交絡変数と処置変数との相関の計算はかなり複雑になる。そこで、本稿は局所独立性、即ち、未観測交絡間は互いに独立しているという仮定を置くこととする。

実は局所独立性の仮定は、複雑に影響し合う交絡変数の主成分を操作的に真の交絡要因として認識すること（即ち交絡変数の分散共分散行列に対する固有値分解を行うこと）に相当しており、ゆえにこの仮定の下でも本稿のモデルの一般性が依然として保たれる。図2で示しているように、本稿では、局所独立性を仮定したモデルにおける交絡変数の行列を Q （ $N \times R$ の行列： N はサンプルサイズで、 R は交絡変数の数）で表記することとする。また、第 j の交絡変数 Q_j がそれぞれ処置変数 A と従属変数 Y へのパスの係数を b_{Aj} と b_{Yj} で表記する。それらの係数を並べると、 Q から A への係数ベクトル b_A と Q から Y へのパスベクトル b_Y が出来上がる。

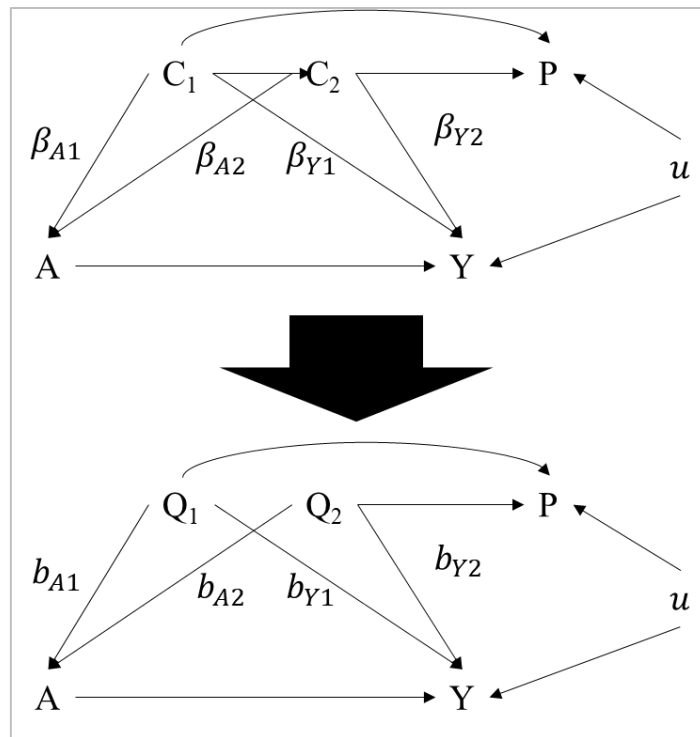


図2 固有値分解による局所独立性

図2の下部のDAGから読めるように、直接YとAの単回帰を行うと、Aの回帰係数は真の因果効果（ β とする）にならずに、いわゆる脱落変数バイアスが生じる。その値を a -bias と表記すると、次のように計算できる。

$$a - bias = \frac{cov(A, Y)}{var(A)} - \beta = \frac{cov(A, Qb_Y + A\beta + U)}{var(A)} - \beta$$

$$= \frac{\text{cov}(Qb_A, Qb_Y)}{\text{var}(A)} = \frac{\sum_j b_{Aj}b_{Yj}}{\text{var}(A)}$$

なお、モデルの解釈を簡単化するため、次節から、**a-bias** が正であると仮定したうえで議論を展開していく。**a-bias** がマイナスなら、本稿の解釈は正負逆転すればそのまま対応できるので、本稿の一般性が保たれる。

3 代理指標（プロキシ）の生成とプロキシバイアスの定式化

DAG に従い、プロキシ **P** を **Q** と **U** から生成する。**T** は、第 **j** の交絡変数 Q_j から代理指標 **P** へのパス係数 T_j を並べた、 $R \times 1$ の生成係数ベクトルである。

$$P = QT + U$$

線形回帰モデルで **P** を統制した時の推定結果を β^* とすると、その値は式①から求められる。

$$\beta^* = \frac{\text{corr}(A, Y) - \text{corr}(A, P) \cdot \text{corr}(P, Y)}{1 - \text{corr}(A, P)^2} \frac{\text{std}(Y)}{\text{std}(A)} \quad \textcircled{1}$$

また、式①における $\text{corr}(A, P)$ 、 $\text{corr}(Y, P)$ は次になる。

$$\begin{aligned} \text{corr}(A, P) &= \sum_j \frac{b_{Aj}T_j}{\text{std}(P)\text{std}(A)} = \sum_j \frac{b_{Aj}\text{corr}(p, Q_j)}{\text{std}(A)} \\ \text{corr}(Y, P) &= \sum_j \frac{(b_{Yj} + b_{Aj}\beta)T_j}{\text{std}(P)\text{std}(Y)} + \frac{\text{var}(U)}{\text{std}(P)\text{std}(Y)} \\ &= \sum_j \frac{b_{Yj}\text{corr}(P, Q_j)}{\text{std}(Y)} + \beta \cdot \text{corr}(A, P) \frac{\text{std}(A)}{\text{std}(Y)} + \frac{\text{var}(U)}{\text{std}(P)\text{std}(Y)} \end{aligned}$$

上記の計算を踏まえると、**p-bias** は次のように整理できる。

$$\begin{aligned} p\text{-bias} &= \beta^* - \beta \\ &= \frac{\frac{\sum_j b_{Aj}b_{Yj}}{\text{std}(A)\text{std}(Y)} - \left(\sum_j \frac{b_{Aj}\text{corr}(P, Q_j)}{\text{std}(A)}\right) \cdot \left(\sum_j \frac{b_{Yj}\text{corr}(P, Q_j)}{\text{std}(Y)} + \frac{\text{var}(U)}{\text{std}(P)\text{std}(Y)}\right)}{1 - \left(\sum_j \frac{b_{Aj}\text{corr}(P, Q_j)}{\text{std}(A)}\right)^2} \frac{\text{std}(Y)}{\text{std}(A)} \end{aligned}$$

ここで、表記を簡単化するために、部分決定係数の考え方を用いて、 ρ という記号を導入する。具体的には、以下の通りである。

$$\frac{b_{Aj}\text{std}(Q_j)}{\text{std}(A)} = \sqrt{\rho_{Aj}}$$

$$\frac{\text{std}(U_A)}{\text{std}(A)} = \sqrt{\rho_{AU}}$$

$$\frac{b_{Yj}\text{std}(Q_j)}{\text{std}(Y)} = \sqrt{\rho_{Yj}}$$

$$\frac{\text{std}(U_Y)}{\text{std}(Y)} = \sqrt{\rho_{YU}}$$

$$\beta \frac{std(A)}{std(Y)} = \sqrt{\rho_{YA}}$$

$$\frac{T_j std(Q_j)}{std(P)} = \sqrt{\rho_{Pj}}$$

$$\frac{std(U_Y)}{std(P)} = \sqrt{\rho_{PU}}$$

以上を踏まえると、 $p - bias$ は、 ρ という記号を用いて式②のように表記できる。

$$p - bias = \frac{\sum_j \sqrt{\rho_{Aj\rho_{Yj}} - \sum_j \sqrt{\rho_{Aj\rho_{Pj}} \cdot \sum_j \sqrt{\rho_{Yj\rho_{Pj}} - \sum_j \sqrt{\rho_{Aj\rho_{Pj}\rho_{YU}\rho_{PU}}}} \cdot \frac{std(Y)}{std(A)}}}{1 - \left(\sum_j \sqrt{\rho_{Aj\rho_{Pj}}}\right)^2} \quad ②$$

4 プロキシバイアスの構成要素

ここで、式②は、以下に示すとおり4つの部分に分解することが可能である。

$$p - bias = \frac{(p - bias_1) - (p - bias_2) - (p - bias_3)}{(p - bias_4)} \quad ③$$

$$\left\{ \begin{array}{l} p - bias_1 = \sum_j \sqrt{\rho_{Aj\rho_{Yj}} \cdot \frac{std(Y)}{std(A)}} \\ p - bias_2 = \sum_j \sqrt{\rho_{Aj\rho_{Pj}} \cdot \sum_j \sqrt{\rho_{Yj\rho_{Pj}} \cdot \frac{std(Y)}{std(A)}}} \\ p - bias_3 = \sum_j \sqrt{\rho_{Aj\rho_{Pj}\rho_{YU}\rho_{PU}} \cdot \frac{std(Y)}{std(A)}} \\ p - bias_4 = 1 - \left(\sum_j \sqrt{\rho_{Aj\rho_{Pj}}}\right)^2 \end{array} \right.$$

そして、それぞれの構成要素は、次のような意味合いを有する。

まず、 $p - bias_1$ は、統制変数を何も投入していない場合の、結果変数Yに対するAの回帰係数に含まれる脱落変数バイアス (a-bias) である。

次に、 $p - bias_2$ は、交絡変数行列QがPによって集約した際に、各 $A \leftarrow Q_j \rightarrow Y$ のバックドアパスが疑似的に $A \leftarrow P \rightarrow Y$ に集約して統制された部分を表している。 $p - bias_1$ と $p - bias_2$ の差分は代理指標Pが統制されてもなお残っている a-bias の一部と解釈できる。

三つ目の $p - bias_3$ は、Pを統制する際の $Q \rightarrow P \leftarrow U$ のバックドアパスによって生じる合流点バイアスである。 $p - bias_3$ は上記の部分的統制をされてもなお残っている部分、即ち $p - bias_1$ と $p - bias_2$ の差分を相殺していき、場合によってはその差分より $p - bias_3$ が大きい場合 $p - bias$ の符号をマイナスの方向へもっていく。

最後に、 $p - bias_4$ は、代理指標PによってAの分散の内、説明されていない部分の割合を意味している。そして、実は $\frac{1}{p - bias_4}$ はPを統制した時のAの分散拡大係数 (variance inflation

factor : VIF) である。

5 プロキシの精度

ここまで複数の交絡変数を一つの代理指標にまとめて統制する場合のプロキシバイアスを定式化した。それでは、ある代理指標に含まれる誤差を抑えることによって因果効果の推定が必ずうまくいくのか、即ち、プロキシの精度の向上が必ず *p-bias* の軽減につながるだろうか。

本節から、 T (プロキシの生成係数ベクトル) の方向が一定という前提のもとで、誤差項 U が P に対する部分決定係数 ρ_{PU} の増減に対し、*p-bias* がいかに挙動するかについて検討していく。

ρ_{PU} の範囲は $[0,1]$ なので、変動をより鮮明に把握するために、本稿は、 $q = \sqrt{\frac{1-\rho_{PU}}{\rho_{PU}}}$ を使用して、 ρ_{PU} の増減を検討していく。 q が無限に近づいていくにつれて、 ρ_{PU} が 0 となり、 q が 0 の時に、 ρ_{PU} が 1 となる。 q が大きいほど、プロキシとしての精度が高いといえる。

p-bias (式②) の内、 q と関連しているのは、 ρ_{PU} と ρ_{Pj} の部分である。

ρ_{PU} を q の数式で書き換えると、式④のようになる。

$$\rho_{PU} = \frac{1}{q^2 + 1} \quad \text{④}$$

また、 ρ_{PU} と ρ_{Pj} の部分決定係数の性質から次の関係式が分かる。

$$\left\{ \begin{array}{l} \sum_j \rho_{Pj} = 1 - \rho_{PU} = \frac{q^2}{q^2 + 1} \\ \rho_{Pj} = \frac{T_j^2 \text{var}(Q_j)}{\text{var}(P)} = \frac{T_j^2 \text{var}(Q_j)}{\sum_j T_j^2 \text{var}(Q_j) + \text{var}(U)} \end{array} \right.$$

$$\Rightarrow \frac{q^2}{q^2 + 1} = \frac{\sum_j T_j^2 \text{var}(Q_j)}{\sum_j T_j^2 \text{var}(Q_j) + \text{var}(U)} \quad \text{⑤}$$

式⑤において、 $\sum_j T_j^2 \text{var}(Q_j)$ は代理指標 P における誤差以外の部分の分散を意味しており、しかも代理指標の精度 q と共変し、 $\sum_j T_j^2 \text{var}(Q_j)$ が大きいほど、 ρ_{PU} が抑制される。

T (プロキシの生成係数ベクトル) の方向が一定という本節の設定を踏まえ、そこで、 $T_j^2 \text{var}(Q_j)$ を式⑥のように分解する。

$$T_j^2 \text{var}(Q_j) = t_j^2 \cdot f(q)^2 \quad \text{⑥}$$

t_j は T (プロキシの生成係数ベクトル) と同方向の 1 ベクトル t の各次元である。即ち、

$$\sum_j t_j^2 = 1$$

本節は、 t が T の方向によって一義的に決まるので、 t_j は ρ_{PU} の増減によらずに一定である。そして、 ρ_{PU} の増減 (即ち精度 q の変化) に反応するのは、 $f(q)$ の部分である。以上のことを踏まえ、式⑤が次のように整理できる。

$$\frac{q^2}{q^2 + 1} = \frac{f(q)^2 \sum_j t_j^2}{f(q)^2 \sum_j t_j^2 + \text{var}(U)} = \frac{f(q)^2}{f(q)^2 + \text{var}(U)}$$

そうすると、 ρ_{Pj} を式⑦のように、 q の数式に書き換えられるようになる。即ち、

$$\Rightarrow \rho_{Pj} = \frac{T_j^2 \text{var}(Q_j)}{\sum_j T_j^2 \text{var}(Q_j) + \text{var}(U)} = \frac{f(q)^2}{f(q)^2 + \text{var}(U)} t_j^2 = \frac{q^2}{q^2 + 1} t_j^2 \quad \textcircled{7}$$

式④と式⑦を p -bias の各構成要素（式③）に代入すると、次のようになる。

$$\left\{ \begin{array}{l} p - bias_1 = \sum_j \sqrt{\rho_{Aj} \rho_{Yj}} \cdot \frac{\text{std}(Y)}{\text{std}(A)} \\ p - bias_2 = \frac{q^2}{q^2 + 1} \cdot \sum_j \sqrt{\rho_{Aj} t_j^2} \cdot \sum_j \sqrt{\rho_{Yj} t_j^2} \cdot \frac{\text{std}(Y)}{\text{std}(A)} \\ p - bias_3 = \frac{q}{q^2 + 1} \cdot \sqrt{\rho_{YU}} \cdot \sum_j \sqrt{\rho_{Aj} t_j^2} \cdot \frac{\text{std}(Y)}{\text{std}(A)} \\ p - bias_4 = 1 - \frac{q^2}{q^2 + 1} \cdot \left(\sum_j \sqrt{\rho_{Aj} t_j^2} \right)^2 \end{array} \right. \quad \textcircled{8}$$

式⑧において、プロキシの精度 q （即ち ρ_{PU} の増減）に影響されるのは四角に囲われる部分のみである。

6 p -bias の各構成部分の変動

q に対する p -bias の各部分の微分をもとめて、それぞれの変動について確認する。なお $p - bias_1$ は即ち a -bias のことで、代理指標の精度 q と関係ない。

1), $p - bias_2$ の変動

$$\begin{aligned} \frac{\partial p - bias_2}{\partial q} &= \frac{\partial \frac{q^2}{q^2 + 1} \cdot \sum_j \sqrt{\rho_{Aj} t_j^2} \cdot \sum_j \sqrt{\rho_{Yj} t_j^2} \cdot \frac{\text{std}(Y)}{\text{std}(A)}}{\partial q} \\ &= \sum_j \sqrt{\rho_{Aj} t_j^2} \cdot \sum_j \sqrt{\rho_{Yj} t_j^2} \cdot \frac{\text{std}(Y)}{\text{std}(A)} \cdot \frac{\partial \frac{q^2}{q^2 + 1}}{\partial q} \\ &= \sum_j \sqrt{\rho_{Aj} t_j^2} \cdot \sum_j \sqrt{\rho_{Yj} t_j^2} \cdot \frac{\text{std}(Y)}{\text{std}(A)} \cdot \frac{2q}{(q^2 + 1)^2} > 0 \end{aligned}$$

上記の計算から、 $p - bias_2$ 、即ち $A \leftarrow P \rightarrow Y$ という仮想的な交絡統制という部分は、 q の増加、即ち P に対する U の影響力の減少により大きくなっていくことが分かる。

2), $p - bias_3$ の変動

$$\begin{aligned}
\frac{\partial p - bias_3}{\partial q} &= \frac{\partial \frac{q}{q^2 + 1} \cdot \sqrt{\rho_{YU}} \cdot \sum_j \sqrt{\rho_{Aj} t_j^2} \cdot \frac{std(Y)}{std(A)}}{\partial q} \\
&= \sqrt{\rho_{YU}} \cdot \sum_j \sqrt{\rho_{Aj} t_j^2} \cdot \frac{std(Y)}{std(A)} \cdot \frac{\partial \frac{q}{q^2 + 1}}{\partial q} \\
&= \sqrt{\rho_{YU}} \cdot \sum_j \sqrt{\rho_{Aj} t_j^2} \cdot \frac{std(Y)}{std(A)} \cdot \frac{1 - q^2}{(1 + q^2)^2}
\end{aligned}$$

上記の式の符号は $1 - q^2$ によって決まる。

$p - bias_3$, 即ち P を統制することによって引き起こされる合流点バイアスは q の増加, 即ち P に対する U の影響力の増加により, まずは小さくなって, その後また大きくなっていく。 $q=1$, 即ち U がちょうど P の分散の 50%を決めるときに, $p - bias_3$ が最大になる。

3), $p - bias_4$ の変動

$$\begin{aligned}
\frac{\partial p - bias_4}{\partial q_U} &= \frac{\partial \left[1 - \frac{q^2}{q^2 + 1} \cdot \left(\sum_j \sqrt{\rho_{Aj} t_j^2} \right)^2 \right]}{\partial q_U} \\
&= - \left(\sum_j \sqrt{\rho_{Aj} t_j^2} \right)^2 \cdot \frac{2q}{(q^2 + 1)^2} < 0
\end{aligned}$$

$p - bias_4$, 即ち A の分散の内, 代理指標 P によって説明されていない部分の割合である。上記の計算から, 代理指標の精度 q の増加とともに, $p - bias_4$ は小さくなっていくことが分る。 $p - bias_4$ はプロキシバイアスの分母に位置している。 $p - bias_4$ の減少はプロキシバイアスの拡大を意味する。このことについて, 次のような解釈ができる。

すでに説明したように, $p - bias_4$ は, A の分散拡大係数の逆数である。つまり, $p - bias_4$ が極端に小さい場合に起きる現象は, 即ち A と P の多重共線性である。それによって, A の係数の推定が不安定になることが推察される。多重共線性は一般的に, 推定値の標準誤差を拡大させる要因として知られており, 推定の有効性が疑われるが, 推定値 b_1 の不偏性に支障は出ないものとされている。しかし, 本稿のモデル設定においては, プロキシバイアスは分散拡大係数に比例して拡大していることが確認できる。

7 p-bias の挙動

代理指標の精度 q が大きくなっていくにつれ, p-bias がいかに変化するだろうか。式⑧を p-bias の計算式 (式③) に代入して整理すると, 次のようになる。

$$p - bias = \frac{\left(\sum_j \sqrt{\rho_{Aj}\rho_{Yj}} - \sum_j \sqrt{\rho_{Aj}t_j^2} \sum_j \sqrt{\rho_{Yj}t_j^2}\right) \cdot q^2 - \sqrt{\rho_{YU}} \sum_j \sqrt{\rho_{Aj}t_j^2} \cdot q + \sum_j \sqrt{\rho_{Aj}\rho_{Yj}}}{q^2 \left[1 - \left(\sum_j \sqrt{\rho_{Aj}t_j^2}\right)^2\right] + 1}$$

$$\cdot \frac{std(Y)}{std(A)} \quad \textcircled{9}$$

$\frac{\partial(p-bias)}{\partial q}$ の符号は次のものによって決まる。

$$\begin{aligned} & \left(2 \left(\sum_j \sqrt{\rho_{Aj}\rho_{Yj}} - \sum_j \sqrt{\rho_{Aj}t_j^2} \sum_j \sqrt{\rho_{Yj}t_j^2}\right) q - \sqrt{\rho_{YU}} \sum_j \sqrt{\rho_{Aj}t_j^2}\right) \left(q^2 \left[1 - \left(\sum_j \sqrt{\rho_{Aj}t_j^2}\right)^2\right] + 1\right) \\ & - 2 \left(\left(\sum_j \sqrt{\rho_{Aj}\rho_{Yj}} - \sum_j \sqrt{\rho_{Aj}t_j^2} \sum_j \sqrt{\rho_{Yj}t_j^2}\right) \cdot q^2 - \sqrt{\rho_{YU}} \sum_j \sqrt{\rho_{Aj}t_j^2} \cdot q \right. \\ & \left. + \sum_j \sqrt{\rho_{Aj}\rho_{Yj}}\right) \left[1 - \left(\sum_j \sqrt{\rho_{Aj}t_j^2}\right)^2\right] q \\ = & \left(2 \left(\sum_j \sqrt{\rho_{Aj}\rho_{Yj}} - \sum_j \sqrt{\rho_{Aj}t_j^2} \sum_j \sqrt{\rho_{Yj}t_j^2}\right) q - \sqrt{\rho_{YU}} \sum_j \sqrt{\rho_{Aj}t_j^2}\right) \\ & + \left(2 \left(\sum_j \sqrt{\rho_{Aj}\rho_{Yj}} - \sum_j \sqrt{\rho_{Aj}t_j^2} \sum_j \sqrt{\rho_{Yj}t_j^2}\right) q^3 - \sqrt{\rho_{YU}} \sum_j \sqrt{\rho_{Aj}t_j^2} q^2\right) \left[1 \right. \\ & \left. - \left(\sum_j \sqrt{\rho_{Aj}t_j^2}\right)^2\right] \\ & - \left(2 \left(\sum_j \sqrt{\rho_{Aj}\rho_{Yj}} - \sum_j \sqrt{\rho_{Aj}t_j^2} \sum_j \sqrt{\rho_{Yj}t_j^2}\right) q^3 - 2\sqrt{\rho_{YU}} \sum_j \sqrt{\rho_{Aj}t_j^2} q^2 \right. \\ & \left. + 2 \sum_j \sqrt{\rho_{Aj}\rho_{Yj}} q\right) \left[1 - \left(\sum_j \sqrt{\rho_{Aj}t_j^2}\right)^2\right] \\ = & \left(2 \left(\sum_j \sqrt{\rho_{Aj}\rho_{Yj}} - \sum_j \sqrt{\rho_{Aj}t_j^2} \sum_j \sqrt{\rho_{Yj}t_j^2}\right) q - \sqrt{\rho_{YU}} \sum_j \sqrt{\rho_{Aj}t_j^2}\right) \\ & + \left(\sqrt{\rho_{YU}} \sum_j \sqrt{\rho_{Aj}t_j^2} q^2 - 2 \sum_j \sqrt{\rho_{Aj}\rho_{Yj}} q\right) \left[1 - \left(\sum_j \sqrt{\rho_{Aj}t_j^2}\right)^2\right] \end{aligned}$$

$$\begin{aligned}
&= -\sqrt{\rho_{YU}} \sum_j \sqrt{\rho_{Aj} t_j^2} + \sqrt{\rho_{YU}} \sum_j \sqrt{\rho_{Aj} t_j^2} \left[1 - \left(\sum_j \sqrt{\rho_{Aj} t_j^2} \right)^2 \right] q^2 \\
&\quad + 2 \left(- \sum_j \sqrt{\rho_{Aj} t_j^2} \sum_j \sqrt{\rho_{Yj} t_j^2} + \sum_j \sqrt{\rho_{Aj} \rho_{Yj}} \left(\sum_j \sqrt{\rho_{Aj} t_j^2} \right)^2 \right) q
\end{aligned}$$

上記の式展開からわかるように、 q の三次項が消去され、実質 q に関する二次関数である。

そして、 q の二次項の係数 $\sqrt{\rho_{YU}} \sum_j \sqrt{\rho_{Aj} t_j^2} \left[1 - \left(\sum_j \sqrt{\rho_{Aj} t_j^2} \right)^2 \right]$ が正で、切片 $-\sqrt{\rho_{YU}} \sum_j \sqrt{\rho_{Aj} t_j^2}$ が負であることから以下のような考察ができる。

q が0の場合、代理指標 P は完全に誤差項 U によって構成されており、その時の p -bias は脱落変数バイアス a -bias そのものである。 q が0から大きくなるにつれ、 $\frac{\partial(p-bias)}{\partial q}$ の符号が最初は負だったので、最初の部分は p -bias が小さくなっていく（場合によってマイナスになる）。その後、 $\frac{\partial(p-bias)}{\partial q}$ の符号が「0」を経て、徐々に正になっていき、 p -bias の値は、一旦最小値をとった後徐々に上昇していくという挙動をする。このプロセスにおいて、 q が無量大に向かっていくにつれ、 p -bias は式⑩に漸近していく。

$$\lim_{q \rightarrow \infty} (p - bias) = \frac{\sum_j \sqrt{\rho_{Aj} \rho_{Yj}} - \sum_j \sqrt{\rho_{Aj} t_j^2} \sum_j \sqrt{\rho_{Yj} t_j^2}}{1 - \left(\sum_j \sqrt{\rho_{Aj} t_j^2} \right)^2} \cdot \frac{std(Y)}{std(A)} \quad \text{⑩}$$

8 p -bias の最大値・最小値

前節では代理指標の精度 q の改善による p -bias の挙動の形を明らかにした。次に、 p -bias の範囲を求める。

さらに表記を簡単化するために、本節より、 $\sum_j \sqrt{\rho_{Aj} t_j^2}$ を X 、 $\sum_j \sqrt{\rho_{Yj} t_j^2}$ を Y 、 $\sqrt{\rho_{YU}}$ を \sqrt{U} 、そして、 $\sum_j \sqrt{\rho_{Aj} \rho_{Yj}}$ を Z と表記する。すると、 p -bias の定式は式⑩のようになる。

$$p - bias = \frac{(Z - XY) \cdot q^2 - \sqrt{U} X \cdot q + Z}{q^2 \cdot (1 - X^2) + 1} \cdot \frac{std(Y)}{std(A)} \quad \text{⑪}$$

p -bias の挙動の範囲は式⑪である（計算の詳細は、appendix を参照）。

$$\begin{aligned}
&\frac{(2Z - XY - X^2 Z) - X \sqrt{U(1 - X^2)} + (XZ - Y)^2}{2(1 - X^2)} \frac{std(Y)}{std(A)} \leq p - bias \\
&\leq \max \left\{ \frac{Z - XY}{1 - X^2} \frac{std(Y)}{std(A)}, Z \frac{std(Y)}{std(A)} \right\} \quad \text{⑫}
\end{aligned}$$

式⑫からわかるように、p-bias の上限値 $\max\left\{\frac{Z-XY}{1-X^2} \frac{std(Y)}{std(A)}, Z \frac{std(Y)}{std(A)}\right\}$ の2項目は実はそれぞれ、 q が無限大に向かっていくときの p-bias の漸近値と、 q が 0 の場合の p-bias (即ち脱落変数バイアス a-bias) である。このことは前節で検討した p-bias の挙動の形と整合的であって、つまり、 q の増大に伴い、p-bias はいったん小さくなってから大きくなっていくものであり、その最大値は初期値 ($q = 0$) と漸近値 ($q \rightarrow \infty$) によって決まる。

9 考察：p-bias と代理指標の是非

ここまでで、p-bias の挙動の範囲が分かった。それでは、p-bias は常に a-bias より抑えられているのか、即ち、代理指標 P はとりあえず統制すればよいものなのだろうか。それらを考察する上で、p-bias が a-bias を超えるのは以下の 2 つの場合に限定されると考えられる。①p-bias の漸近値が a-bias よりも大きい場合、そして、②p-bias の下限値が $-(a-bias)$ よりも小さい場合である。この二つの場合に限って、代理指標の統制が逆に大きなバイアスを導入していることを意味する。ここで、双方の場合の p-bias の挙動を図 3 に示す。

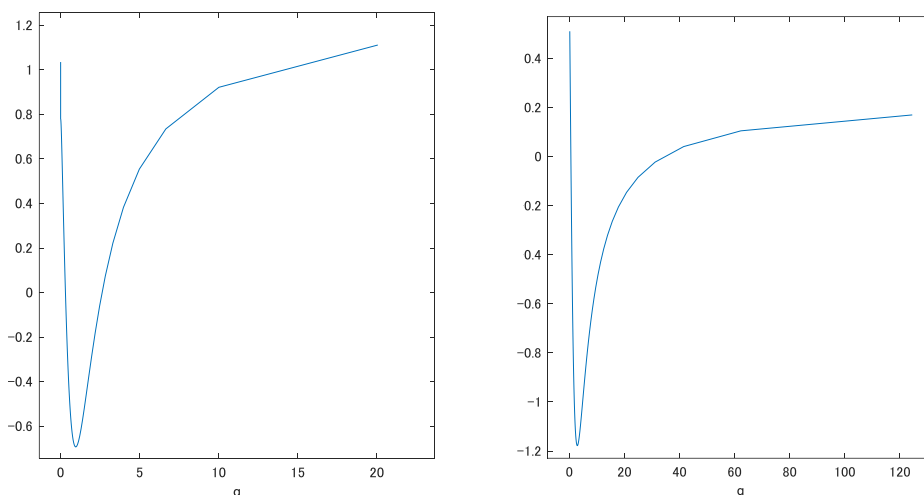


図 3 p-bias が a-bias を超える場合

(左：漸近値が a-bias よりも大きい場合、右：p-bias の下限値が $-(a-bias)$ 以下)

この二つの場合はそれぞれどのような条件の下で起きるのだろうか。

p-bias の漸近値が a-bias よりも大きい場合、式⑬が満たされる。

$$\frac{Z - XY}{1 - X^2} > Z$$

$$\Rightarrow \frac{Y}{X} < Z \quad \text{⑬}$$

p-bias の下限値が $-(a-bias)$ よりも小さい場合、式⑭が満たされる。

$$\frac{(2Z - XY - X^2Z) - X\sqrt{U(1 - X^2) + (XZ - Y)^2}}{2(1 - X^2)} < -Z$$

$$4Z - XY - 3X^2Z < X\sqrt{U(1 - X^2) + (XZ - Y)^2}$$

$$(4Z - XY - 3X^2Z)^2 < X^2[U(1 - X^2) + (XZ - Y)^2]$$

$$16Z^2 + X^2Y^2 + 9X^4Z^2 - 24X^2Z^2 - 8XYZ + 6X^3YZ < X^2[U(1 - X^2) + X^2Z^2 - 2XYZ + Y^2]$$

$$16Z^2 + 8X^4Z^2 - 24X^2Z^2 - 8XYZ + 8X^3YZ < U(1 - X^2)X^2$$

$$8(1 - X^2)(2 - X^2)Z^2 - 8XYZ(1 - X^2) < U(1 - X^2)X^2$$

$$8(2 - X^2)Z^2 - 8XYZ - UX^2 < 0$$

$$\frac{8XY - \sqrt{64X^2Y^2 + 32(2 - X^2)UX^2}}{16(2 - X^2)} < 0 < Z < \frac{8XY + \sqrt{64X^2Y^2 + 32(2 - X^2)UX^2}}{16(2 - X^2)}$$

$$\Rightarrow Z < \frac{2Y + \sqrt{4Y^2 + 2(2 - X^2)U}}{4(2 - X^2)} X \quad \text{⑭}$$

式⑬及び式⑭のいずれかを成立させるという XYZ の組み合わせは、 q の値によって **p-bias** が **a-bias** を超えることもあり、代理指標を統制することによってより正確な推定結果が必ずしも期待できるものではない。

逆に、式⑬及び式⑭の両方を同時に成立させない XYZ の組み合わせであれば、代理指標の精度に依らず、統制することが望ましいといえる。その際に XYZ は以下の関係にある。

$$\frac{2Y + \sqrt{4Y^2 + 2(2 - X^2)U}}{4(2 - X^2)} X < Z < \frac{Y}{X}$$

10 補論：漸近値の符号と X, Y, Z の意味合い——球面三角法からのヒント

本稿で提起した、複数の交絡変数が存在するモデルにおける **p-bias** の挙動は、中尾ほか (2022) で計算した交絡変数がある場合と似たような形である。ただ、中尾ほか (2022) は、交絡変数があるモデルにおいて、代理指標の精度が無限に高くなれば **p-bias** の漸近値は必ず 0 になることが確認できた。一方、すでに論じていたように、複数の交絡変数が想定される場合、一つの代理指標のみ統制する際の **p-bias** は精度を無限に高くしたとしても漸近値が必ずしも 0 になるわけではない。

これまで、代理指標の改善としてまず誤差を除去することが考えられてきたが、精度が無限に高いときの **p-bias** の漸近値は代理指標の性質を理解するのに最も重要な議論ともいえよう。前節で挙げた例からもわかるように、複数の交絡が存在する際に、そもそも漸近値そのものが **a-bias** を超えている可能性もあり、いくら誤差項を除去したところで、推定値が真値から遠ざかっていくことも考えられる。

p-bias の漸近値の符号は式⑩からわかるように、 $Z - XY$ によって決まる。本節では、幾何学からのヒントをもとに XYZ の意味合いとともに、漸近値の符号を決める条件について解釈してみることにする。

本稿では、 $\sum_j \sqrt{\rho_{Aj}t_j^2}$ を X、 $\sum_j \sqrt{\rho_{Yj}t_j^2}$ を Y、 $\sum_j \sqrt{\rho_{Aj}\rho_{Yj}}$ を Z としており、この定義から、X は即ち $\sqrt{\rho_{Aj}}$ を並べたベクトルと t_j を並べたベクトルの内積であり、Y は即ち $\sqrt{\rho_{Yj}}$ を並べたベクトルと t_j を並べたベクトルの内積であり、Z は即ち $\sqrt{\rho_{Aj}}$ を並べたベクトルと $\sqrt{\rho_{Yj}}$ を並べたベクトルの内積であることが分かる。

また、 $\sqrt{\rho_{Aj}}$ 、 t_j と $\sqrt{\rho_{Yj}}$ の意味合いはそれぞれ次のようになる。 ρ_{Aj} とは A に対する Q_j の部分決定係数であり、即ち第 j の交絡変数が処置変数に与える影響の相対的強さである。 t_j は T (プロキシの生成係数ベクトル) と同方向の 1 ベクトル t の各次元で、即ち第 j の交絡変数が代理指標に対する影響を表している。 ρ_{Yj} とは Y に対する Q_j の部分決定係数であり、即ち第 j の交絡変数が従属変数に与える影響の相対的強さである。つまり、三本のベクトルはそれぞれ、処置変数 A、代理指標 P、従属変数 Y を生成する際の各交絡変数の影響力の配分を示している。

このことを踏まえると、三本のベクトル間の内積とは、構成要件の「近さ」ととらえられる。内積の値が大きければ大きいほど、両ベクトル・両変数の構成が相似しているといえる。例えば、X が大きければ、A と P の生成式における各次元の交絡変数の影響の大きさが互いに似ており、結果として A と P の相関も高くなる。Z - XY の符号は、本質的に上記三本のベクトルの方向関係によって決められる。言い換えると、いかに P を経由して $Q \rightarrow A$ と $Q \rightarrow Y$ の関係が記述されるかによって決まる。

では、p-bias の漸近値がちょうどゼロになるのはどういう状況だろうか。この場合、 $Z - XY = 0$ となる。このことは幾何学の観点から以下のように解釈される (図 4)。

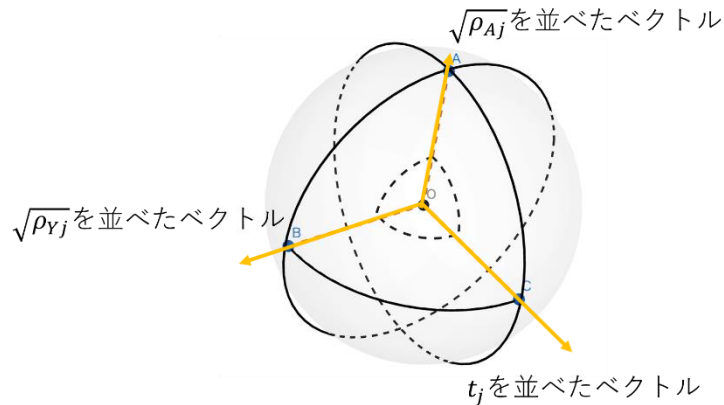


図 4 Z - XY がちょうど 0 となる際の三本のベクトルの位置関係

球面三角法より、AOC が決まる平面と BOC が決まる平面が垂直であれば、以下の余弦値間の関係を満たすことが知られている。

$$\cos(AOB) = \cos(AOC)\cos(BOC)$$

以上の関係式に基づいて、OA を $\sqrt{\rho_{Aj}}$ を並べたベクトル、OC を t_j を並べたベクトル、そして OB を $\sqrt{\rho_{Yj}}$ を並べたベクトルに見立てるとする。そして、 $Z - XY = 0$ という条件は即

ち、 $\sqrt{\rho_{Aj}}$ を並べたベクトルと t_j を並べたベクトルがなす平面と、 $\sqrt{\rho_{Yj}}$ を並べたベクトルと t_j を並べたベクトルがなす平面とがちょうど垂直することを意味している。

三本のベクトルの意味合いをさらに鑑みて、このことは、**p-bias**の漸近値がちょうど0になるような代理指標を作り上げるためには、一方的に交絡変数からAへの影響のみ、もしくはYへの影響のみを考慮する（即ちXまたはYを一方的に高める）だけでは不十分で、両者を同時に「いいバランス」のもとで組み入れて t_j を決定する必要がある。また、ここでいう「良いバランス」とは、 $\sqrt{\rho_{Aj}}$ を並べたベクトルと t_j を並べたベクトルがなす平面と、 $\sqrt{\rho_{Yj}}$ を並べたベクトルと t_j を並べたベクトルがなす平面とがちょうど垂直というような位置関係である。

以上、幾何学の視点から**p-bias**の漸近値について解釈を試みたが、現段階ではまだ初歩的な考察にとどまる。しかし、本稿の検討は以下の示唆と含意を有する。本稿の取り組みは幾何学の観点から交絡変数の代理指標の性質を理解するアプローチの可能性を示している。また、本稿の知見に基づき、代理指標の測定誤差の観点を鑑みれば、傾向スコアなどQからAの影響に基づいた交絡統制が、QからYの影響を考慮できていない点からすると、**p-bias**が0に漸近していない場合も考えられる。

代理指標を通して間接的に分析を行うという社会科学の実証分析の基本的特性を念頭に、代理指標の性質および因果推論において使用することの是非についてさらなる深い理解を目指し、以下の取り組みを今後の課題とする。

一つ目に、本稿は複数の交絡変数を有するモデルに一つのプロキシを投入する際の是非について検討したが、実際の研究では複数の代理指標を投入することが普通である。これからFWL定理などに基づき、複数のプロキシを投入する場合の**p-bias**について検討していく。次に、傾向スコアなど確率的や非線形的に生成した代理指標に対応するように本稿のモデルをさらに拡張していく。三つ目に、本稿の知見に基づいて、代理指標の使用の是非に関する感度分析の枠組みを開発し、実際の分析で用いることができるように、また代理指標の実査に活用できるような取り組みに繋げていくことも重要であろう。またそのうえで、実際の分析に使用されている代理指標の性質を経験的に理解することを進めていく。

参考文献

- Elwert, Felix and Pfeffer, Fabian. T., 2019, “The Future Strikes Back: Using Future Treatments to Detect and Reduce Hidden Bias,” *Sociological Method & Research*, 1-38.
- 樊怡舟・中尾走, 2022, 「従属変数の測定誤差が統計的因果推論へ与える影響」『*Advancement of Higher Education Research: RIHE Monograph Series*』10 : 1-19.
- Miao Wang, Geng Zhi, and Tchetgen, E. T., 2018, “Identifying Causal Effects With Proxy Variables of an Unmeasured Confounder,” *Biometrika*, 105(4): 987-993.
- Ogburn, E. L. and VanderWeele, T. J., 2012, “On the Nondifferential Misclassification of a Binary Confounder,” *Epidemiology*, 23(3): 433-439.

- Streiner, D. L., Norman, G. R., Cairney, J., 2015, *Health Measurement Scales*, Oxford University Press, (=2016, 木原雅子・加治正行・木原正博訳, 『医学的測定尺度の理論と応用』メディカル・サイエンス・インターナショナル) .
- VanderWeele TJ, Hernán MA., 2012, “Results on differential and dependent measurement error of the exposure and the outcome using signed Directed Acyclic Graphs. *American Journal of Epidemiology*, 175(12): 1303–1310.
- 中尾走・樊怡舟・村澤昌崇, 2022, 「交絡変数の測定誤差が合流点バイアスをもたらす影響について」『理論と方法』71 : 34-52.
- 中尾走・樊怡舟・康凱翔, 2023, 「因果効果の推定における交絡変数の欠損と測定誤差のトレードオフの検討：いくつかの想定から」第74回数理社会学会発表資料.
- 白川俊之, 2021, 「子どもによる社会経済的背景の報告とその妥当性」『社会と調査』27 : 64-75.
- 林岳彦・黒木学, 2016, 「相関と因果と丸と矢印のはなし：はじめてのバックドア基準」岩波データサイエンス刊行委員会『岩波データサイエンス3』岩波書店.

Appendix

p-bias の範囲は以下のように計算される。

$p - bias = K \cdot \frac{std(Y)}{std(A)}$ とすると、次の式が成立する。

$$\begin{aligned} \frac{(Z - XY) \cdot q^2 - \sqrt{UX} \cdot q + Z - K[1 - X^2] \cdot q^2 - K}{q^2 \cdot [1 - X^2] + 1} &= 0 \\ \Rightarrow (Z - XY - (1 - X^2)K) \cdot q^2 - \sqrt{UX} \cdot q + Z - K &= 0 \quad A - ① \end{aligned}$$

式 A-①は q に関する二次方程式となっている。 $q = \sqrt{\frac{1 - \rho_{PU}}{\rho_{PU}}}$ という本稿の設定を念頭に、

p-bias の範囲内の K なら、式 A-①にすくなくとも一つの非負解が存在する。

そのために、実数解が存在し、そしてより大きい解が非負であるという 2 つの条件を満たさなければならない。

まず、判別式が式 A-②となる

$$\begin{aligned} \Delta &= UX^2 - 4(Z - XY - (1 - X^2)K)(Z - K) \\ &= UX^2 - 4Z(Z - K) + 4XY(Z - K) + 4(1 - X^2)K(Z - K) \\ &= UX^2 - 4Z^2 + 4ZK + 4XYZ - 4XYK + 4(1 - X^2)ZK - 4(1 - X^2)K^2 \\ &= UX^2 - 4Z^2 + 4XYZ + [4Z - 4XY + 4(1 - X^2)Z]K - 4(1 - X^2)K^2 \quad A - ② \end{aligned}$$

式 A-②もまた K に関する二次方程式となっており、二次項係数 $-4(1 - X^2)$ は、交絡が複数存在するモデルにおいて常にマイナスである (R=1 且つ A の生成式に誤差項がないという場合のみ X が 1 となる)。

つまり、式 A-①の実数解が存在するという条件 (即ち式⑬が非負) を成立させる K の範囲は次の不等式になる (式 A-③) ¹。

$$\begin{aligned} \frac{(2Z - XY - X^2Z) - X\sqrt{U(1 - X^2) + (XZ - Y)^2}}{2(1 - X^2)} &\leq K \\ &\leq \frac{(2Z - XY - X^2Z) + X\sqrt{U(1 - X^2) + (XZ - Y)^2}}{2(1 - X^2)} \quad A - ③ \end{aligned}$$

もう一つの条件、式 A-①に非負解が存在する場合、式 A-④が満たされる。

$$\max \left\{ \frac{\sqrt{UX} + \sqrt{\Delta}}{2(Z - XY - (1 - X^2)K)}, \frac{\sqrt{UX} - \sqrt{\Delta}}{2(Z - XY - (1 - X^2)K)} \right\} \geq 0 \quad A - ④$$

¹ 式 A-②が 0 になるときの解を求めるための判別式を計算すると、次のように、必ず正であることが確認できた。

$$\begin{aligned} \Delta_k &= [4Z - 4XY + 4(1 - X^2)Z]^2 + 4(UX^2 - 4Z^2 + 4XYZ)4(1 - X^2) \\ &= 16[(2Z - XY - X^2Z)^2 + (UX^2 - 4Z^2 + 4XYZ)(1 - X^2)] \\ &= 16X^2(U - UX^2 + X^2Z^2 - 2XYZ + Y^2) = 16X^2[U(1 - X^2) + (XZ - Y)^2] > 0 \end{aligned}$$

式 A-④の分母の符号によって、上式は次の二つの場合に分けられる。

$$2(Z - XY - (1 - X^2)K) > 0 \Rightarrow \frac{Z - XY}{1 - X^2} > K \text{ の場合に, } \frac{\sqrt{U}X + \sqrt{\Delta}}{2(Z - XY - (1 - X^2)K)} \text{ が必ずプラスなので, 式}$$

A-④が常に満たされる。

一方, $\frac{Z - XY}{1 - X^2} < K$ の場合に, 式②の一次項 $4Z - 4XY + 4(1 - X^2)Z < 0$ 。二次関数 Δ の対称軸がマイナスであることが分かる。そうすると, 非負解が存在することは, 即ち二つの解の積がマイナスであることを意味する。根と係数の関係に基づいて, 次の関係が得られる。

$$\begin{aligned} \frac{Z - K}{Z - XY - (1 - X^2)K} &\leq 0 \Rightarrow K \leq Z \\ &\Rightarrow \frac{Z - XY}{1 - X^2} < K \leq Z \end{aligned}$$

以上のように二つの場合を検討した結果, $\frac{Z - XY}{1 - X^2} > K$, もしくは $\frac{Z - XY}{1 - X^2} < K \leq Z$ のいずれの場合でも式 A-④が満たされることが分かった。二つの場合を合わせると, 式 A-⑤となる。

$$K \leq \max\left\{\frac{Z - XY}{1 - X^2}, Z\right\} \quad \text{A - ⑤}$$

実は, 次の計算で確認できるように, 式 A-⑤が式 A-③が示している K の上限の十分条件である。

$$\begin{aligned} &\frac{(2Z - XY - X^2Z) + X\sqrt{U(1 - X^2) + (XZ - Y)^2}}{2(1 - X^2)} - \frac{Z - XY}{1 - X^2} \\ &= \frac{(XY - X^2Z) + X\sqrt{U(1 - X^2) + (XZ - Y)^2}}{2(1 - X^2)} > \frac{X(Y - XZ) + X|XZ - Y|}{2(1 - X^2)} \geq 0 \\ &\frac{(2Z - XY - X^2Z) + X\sqrt{U(1 - X^2) + (XZ - Y)^2}}{2(1 - X^2)} - Z \\ &= \frac{(-XY + X^2Z) + X\sqrt{U(1 - X^2) + (XZ - Y)^2}}{2(1 - X^2)} > \frac{-X(Y - XZ) + X|XZ - Y|}{2(1 - X^2)} \geq 0 \end{aligned}$$

そこで, 式 A-③と式 A-⑤を合わせると, p-bias の範囲は式⑦のようになる。

$$\frac{(2Z - XY - X^2Z) - X\sqrt{U(1 - X^2) + (XZ - Y)^2}}{2(1 - X^2)} \leq K \leq \max\left\{\frac{Z - XY}{1 - X^2}, Z\right\}$$