

## JGSS 統計分析セミナー2009

### - 傾向スコアを用いた因果分析 -

三輪 加奈  
大阪商業大学 JGSS 研究センター

菅澤 貴之  
九州大学キャリア支援センター

JGSS Statistical Analysis Seminar:  
Causality Analysis based on the Propensity Score

Kana MIWA  
JGSS Research Center  
Osaka University of Commerce

Takayuki SUGASAWA  
Career Path Creation Center  
Kyushu University

JGSS Research Center hosted a statistical analysis seminar on the theme of causality analysis based on the propensity score. Propensity score is one methodology to analysis causal effect when a random sampling data is not available. Although causality analysis using the propensity score is known as to avoid the sample selection bias which we cannot deal with by normal multivariable analysis, there are few studies that employed this approach in Japan. This paper aims to explain how to do the causality analysis based on the propensity score, and show examples by using JGSS data.

Key Words: JGSS, causality analysis, propensity score, satisfaction of family life

JGSS 研究センターでは、「傾向スコアを用いた因果分析」をテーマに 2009 年度統計分析セミナーを開催した。傾向スコア (propensity score) とは、ランダムな割り当てが不可能な相関研究において、因果効果を分析する方法として生まれた概念である。傾向スコアを用いる因果分析は、今日注目され、かつ必要とされている分析手法であり、通常の変数解析では対応しきれない選択バイアスの問題を回避することができる。しかし、日本ではまだ適用頻度が低いといえる。本稿では、傾向スコアによる調整を含む因果分析とは、どのようなものなのか、またその分析の手順について、JGSS のデータを用いた実習例を示しながらまとめる。

キーワード：JGSS，因果分析，傾向スコア，家庭生活満足度

## 1. はじめに

JGSS 研究センター主催の統計分析セミナーは、2009年8月27日と28日の2日間、統計分析のスキルアップを目指す大学院生・研究者を対象として開催された。セミナーの講師には、2007年・2008年に引き続いて、シカゴ大学社会学部の山口一男教授を招聘した。2009年の統計分析セミナーでは、因果関係を分析する際に、有効であるとされ、近年その適用例が増加しつつある、「傾向スコア (propensity score)」を用いた因果分析についての講義をお願いした。

因果関係についての分析方法としては、多変量解析が広く用いられているが、多変量解析の分析手法では、標本選択バイアス (sample selection bias) の問題を完全に回避することができず、推定量に偏りが生じる場合がある。この選択バイアスの問題を解決する方法として、注目されているのが、「傾向スコア」を用いた因果分析である。傾向スコアによる因果分析は、因果関係分析を行う多くの場合に、その適用が望ましいとされている。

本稿では、統計分析セミナーで紹介された、傾向スコアを用いた因果分析の理論的枠組みと分析手法を解説する。セミナーでは、日本版総合的社会調査 (Japanese General Social Surveys、以下 JGSS) のデータを用いた分析演習も行われたので、この点も解説に含める。傾向スコアなどの新たな分析手法を取り入れた因果分析の手法は、その重要性は認識されているが、日本ではまだその適用頻度は低い。したがって、本稿において、その分析・推計手順を示すことは意義があるといえる。

本稿の構成は次の通りである。第2節では、傾向スコアを用いた因果関係分析の理論的枠組みを示す。第3節では、統計セミナーでの分析演習例で用いたデータの概要を述べ、続く第4節で、因果分析の手順について実習例を用いて説明する。最後に、統計セミナーで取り上げられた内容についてまとめる。

## 2. 傾向スコアの利用と因果分析

### 2.1 傾向スコアを用いた因果分析の枠組み

これまで、統計的分析での因果関係は、「結果  $Y$  に影響する他の条件をすべて固定したときに、ある説明変数 (処理変数)  $X$  の変化が  $Y$  を変化させるか否かの問題」と概念化していた。具体的には、以下の3つの条件を満たせば、 $X$  は  $Y$  に因果的に影響すると考えた。

(1) Temporality :  $X$  は  $Y$  に時間的に先行する

(2) Correlation :  $X$  と  $Y$  には有意な相関がある

(3) Non-spuriousness : 相関は疑似相関でない

( =  $X$  と  $Y$  に時間的に先行し、 $X$  と  $Y$  の双方に影響を与える変数—共変量、交絡要因—を多変量回帰モデルなどで制御しても相関は有意である )

ここで、 $X$  と  $Y$  の両方に影響する共変量 (交絡要因) の影響を完全に排除することは困難である。また、たとえ適切な共変量を見つけることができても、単に線形回帰で統制するだけではうまく処理できるとは限らないために、推計結果にバイアスが生じやすい。このような問題に対処するための方法として出てきたのが、「因果関係分析 (因果分析)」である。

因果関係分析では、ある「処理 (治療、treatment)」とその結果  $Y$  を考え、「反事實的 (counterfactual) 因果関係」の定義をおく。なお、 $Y_i$  は個人  $i$  の結果を表す。

**定義** : 処理 (treatment) が個人  $i$  に与える因果的処理効果 (平均的処理効果—average treatment effect—) は、次のように定義される ( $E$  は統計的期待値)。

$$d_i \equiv E\{Y_i | \text{treatment}\} - E\{Y_i | \text{no treatment}\}$$

この定義では、ある個人*i*に対しては、実際には、 $E\{Y_i | treatment\}$ と $E\{Y_i | no treatment\}$ のうちいずれか一方のケースのみしか観察できず、その差を直接求めることはできない。また、ある処理を受けることに対する「選択バイアス」の問題も生じうる。

選択バイアスについて、具体的に、「婚姻率と健康（罹患率・死亡率）」との関係を例にすると、次のようである。先行研究では、この関係について、「婚姻群（結婚している人の集合）と非婚姻群（結婚していない人の集合）を比較すると、結婚している人の方がその後の健康状態が良い」という結果が示されている。しかし、この関係には、そもそも健康でない人の方が離婚する確率が高いかもしれないし、病気がちな人が結婚する可能性が低いかもしれないという、選択バイアスが存在していると考えられる。このバイアスが生じる理由は、婚姻群と非婚姻群がランダムに振り分けられていないことによる。

ここで、仮に、標本を処理群（treatment group） $T$ と統制群（control group） $C$ とにランダムに割り当てるのが実験などにより可能であれば、平均的処理効果  $d$  を処理群についてグループの平均で測定できることになる。

$$d = E\{Y^T | treatment\} - E\{Y^T | no treatment\}$$

既述のように、上式の右辺第 1 項は観察することができるが、第 2 項は観察することができない。もしこれがランダムな割り当てであれば、

$$E\{Y^T | no treatment\} = E\{Y^C | no treatment\}$$

となり、 $d$  の推定値は、以下の通りとなる。

$$E\{Y^T | treatment\} - E\{Y^C | no treatment\}$$

しかし、標本をランダムに割り当てるという実験は、現実社会における観察では倫理的にすべきではなく、通常可能な自然観察の場合には、上述の例のように処理群と統制群の  $Y$  の平均値は異なるために、実際には処理群への選択バイアス  $d_s$  が生じることとなる。

$$\begin{aligned} d &= E\{Y^T | treatment\} - E\{Y^T | no treatment\} \\ &= E\{Y^T | treatment\} - E\{Y^C | no treatment\} - d_s \end{aligned}$$

このように、ランダムな割り当てが不可能な状況において、因果効果を推定する方法として Rosenbaum and Rubin (1983) は、「傾向スコア (propensity score)」という新しい概念を提案した。彼らが提案した方法では、「強く無視できる割り当て条件 (strongly ignorable treatment assignment)」を仮定している。この仮定は、完全にランダムではないが、『観察される複数の変数 (共変量)  $\{Z_i\}$  のみに  $X$  の (処理群と統制群への) 割り当て確率  $P$  が依存する』というものであり、この確率  $P$  が、傾向スコアである。傾向スコアを用いた、ある処理を受けることに対する平均的処理効果は、下式のように表すことができる。

$$E\{Y^T - Y^C\} = E\{E[Y^T | treatment, P(\mathbf{Z})] - E[Y^T | no treatment, P(\mathbf{Z})]\}$$

この傾向スコアを用いた因果分析においては、「傾向スコアを所与として (= 同じ傾向スコアの値のもとで) 各群から得られた  $Y^T, Y^C$  から  $(Y^T - Y^C)$  を計算し、それを傾向スコアの分布で期待値を取ると、元々の関心対象である平均的処理効果となる」ことがわかる (星野 2009)。つまり、この傾向スコアの分布が処理群と統制群で同じになるように調整すれば、 $X$  の  $Y$  への影響は、あたかも処理群と統制群が完全にランダムに割り当てられたようにみなせることになり、因果分析が可能となる。

その処理群と統制群の傾向スコアの分布を同じにする方法の一つが、Rubin (1985) や Imbens (2000) らによって発展された、逆確率処理推定 (inverse probability treatment estimation、以下 IPTE) であり、

傾向スコアの逆数によるウェイト付けである<sup>(1)</sup>。

IPTE 法では、「観察されない共変量（あるいは交絡要因）がない」と仮定しており、傾向スコア  $P$  を用いて、処理群の標本を  $1/P$  倍、統制群の標本を  $1/(1-P)$  倍するという、処理変数  $X$  の値に応じてそれぞれに  $P$  の逆確率をかけることで、処理群と統制群の標本がともに  $P$  の一様分布をもつことになる。その結果、ウェイト調整後は、 $P$  と処理変数  $X$  が独立になり、標本があたかもランダムに  $X$  に割り当てられたような状態が生まれる（なお、実際には、これらをさらに  $X$  の周辺分布が調整前と調整後で同じになるようにウェイト調整をする必要がある）。このウェイト調整された標本を用いることで、ある処理  $X$  を受けることの  $Y$  への影響（平均的処理効果）を推計することができる。次の小節では、IPTE 法による因果分析の推計手順を示す。

## 2.2 傾向スコアを用いた因果分析の推計手順：逆確率処理推定法

傾向スコアを用いる逆確率処理推定法（IPTE 法）による、因果分析の推計手順は下表に示す通りである（なお、本稿の推計手順は、処理変数  $X$  が 2 値のケースを示している）。

|    |                                                                                                                                                                                   |
|----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. | 処理変数 $X_i$ を、ロジスティック回帰分析により推計し、傾向スコア $P$ のモデルを求める。<br>$X_i = \alpha + \beta \cdot Z_i + v_i$                                                                                      |
| 2. | 傾向スコア $P_i$ を、各標本について計算する。<br>$P_i = \exp[\hat{\alpha} + \hat{\beta} \cdot Z_i] / (1 + \exp[\hat{\alpha} + \hat{\beta} \cdot Z_i])$ （ $\hat{\alpha}$ , $\hat{\beta}$ はそれぞれ推定値を表す） |
| 3. | $X_i$ の値に応じて、標本に逆確率である、 $1/P_i$ または $1/(1-P_i)$ をウェイトとしてかける。                                                                                                                      |
| 4. | 頻度を $X$ の周辺分布がウェイト前とウェイト後で同じになるように再度調整する。<br>このとき、手順 3 の調整の収束度をチェックする。                                                                                                            |
| 5. | ウェイトをかけた標本を用いて、結果 $Y_i$ について回帰分析を行う（回帰モデルは、従属変数の性質による）。<br>$Y_i = \delta + \text{Weighted}[\eta \cdot X_i + \gamma \cdot Z_i] + u_i$                                              |

以上の推計の手順において、注意すべき点は、（特に欠測データが含まれるデータセットを用いる場合には）手順 1・4・5 がすべて同じ標本にもとづいているかどうかということであり、それを確認しながら推計を行う必要がある。

手順 3 において、傾向スコアを用いたウェイト調整後には、処理変数  $X$  とその予測変数群（共変量） $\{Z\}$  は独立になり、 $Y_i$  への回帰式に  $\{Z_i\}$  を含めなくても、 $X$  の影響は理論的には変わらないはずである。

しかし、実際の推計においては、傾向スコア  $P_i$  の推定は完全ではないかもしれないし、IPTE 法で  $P_i$  と  $X_i$  の独立を達成する十分な情報がないかもしれない。このような状況下で、傾向スコアによる調整を含む因果分析の推定結果をより頑強なものにするためには、 $Y_i$  への回帰分析の際に、傾向スコア調整後であっても、共変量  $\{Z_i\}$  を説明変数として加えることが望ましい。その場合、(a)  $Y$  への  $X$  と  $\{Z\}$  の線形の影響の仮定か、(b) 傾向スコアについての仮定（傾向スコアは正しく推定され、また IPTE 法において  $P$  と  $X$  の分布はほぼ独立となる）の、いずれか一方が成立していれば、 $X$  の  $Y$  への影響は正しく推定されることとなる。この場合の処理変数  $X$  の係数の推定値は、「二重にロバストな推定量（doubly robust estimator）」であるといわれる。

この二重にロバストな推定量は、仮定 (a) が成立しているときには「共変量で従属変数を説明する回帰関数」のみを利用した推定量よりは推定量の分散が大きくなる。しかし、仮定 (a)・(b) のどちらかが成立していれば一致推定量が得られるという利点は、推定量の効率より重視される場合が多いことから、近年応用研究で利用されるようになってきている（星野 2009）。

以下では、傾向スコアによるウェイト調整を含む因果分析について、JGSS のデータを用いてその演習例を示す。

### 3. データの概要

JGSS 統計セミナーでは、統計解析ソフト SPSS のプログラムを用いて、アメリカの General Social Survey (GSS) と日本版総合的社会調査 (JGSS) の、それぞれのデータセットを用いた分析演習が示されたが、本稿では、そのうちでも特に、JGSS のデータによる分析演習例を取り上げる。

分析には、2000 年～2003 年にかけて日本全国で実施された、JGSS-2000・JGSS-2001・JGSS-2002・JGSS-2003 の 4 年間の累積データを用いる。分析対象は、20～69 歳の男女個人で、推計に用いるすべての変数について欠損値のない 10,063 人とする。なお、各年の調査の概要は表 1 の通りである。

表 1 JGSS-2000～2003 の調査概要

|       | JGSS-2000                              | JGSS-2001                             | JGSS-2002                             | JGSS-2003                             |
|-------|----------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|
| 調査時期  | 2000 年 10～11 月                         | 2001 年 10～11 月                        | 2002 年 10～11 月                        | 2003 年 10～11 月                        |
| 調査対象  | 2000 年 6 月 25 日時点で満 20 歳以上 89 歳以下の男女個人 | 2001 年 9 月 1 日時点で満 20 歳以上 89 歳以下の男女個人 | 2002 年 9 月 1 日時点で満 20 歳以上 89 歳以下の男女個人 | 2003 年 9 月 1 日時点で満 20 歳以上 89 歳以下の男女個人 |
| 標本規模  | 4,500 人                                | 4,500 人                               | 5,000 人                               | 7,000 人                               |
| 標本抽出  | 層化 2 段無作為抽出法                           | 層化 2 段無作為抽出法                          | 層化 2 段無作為抽出法                          | 層化 2 段無作為抽出法                          |
| 調査方法  | 面接および留置調査                              | 面接および留置調査                             | 面接および留置調査                             | 面接および留置調査<br>(留置は A・B の 2 種類)         |
| 有効回答数 | 2,893 人                                | 2,790 人                               | 2,953 人                               | 留置 A 票：1,957 人<br>留置 B 票：1,706 人      |
| 回収率   | 64.9%                                  | 63.1%                                 | 62.3%                                 | 留置 A 票：55.0%<br>留置 B 票：48.0%          |

出所：JGSS ウェブサイトより筆者作成

### 4. 傾向スコアによる調整を含む因果分析の演習例 - JGSS データを用いて -

本節では、統計セミナーで行われた分析演習のうち、JGSS-2000～2003 の 4 年間の累積データを用いて、傾向スコアによるウェイト調整を含めた、因果分析の方法を示す。

分析では、処理変数  $X$  が 2 値をとる場合を考える。処理変数以外の (共変量としての) 説明変数については、それが連続変数 (または 3 値以上の値をとる変数) である場合には、それぞれの変数について、2 値のダミー変数として定義 (リコード) する必要がある。これは、2 値のダミー変数とした方が、傾向スコア等を推計する際に収束しやすく、より望ましいといわれているためである。

以下では、実際の例を示しながら、傾向スコアによるウェイト調整をとまなう因果分析の方法を、順を追って説明する。

#### 4.1 分析例 1：家庭生活満足度 - 有配偶者 vs 無配偶者 -

はじめに、配偶者の有無 (有配偶であるか、無配偶であるか) が「家庭生活満足度」に与える影響に関する因果分析の例を取り上げる。

従属変数  $Y1$  となる家庭生活満足度は、「家庭生活について、あなたはどのくらい満足していますか」と尋ね、1 = 不満～5 = 満足の 5 段階の尺度で測られる指標である。また、処理変数  $X1$  としては、ある人に配偶者がいる (有配偶) 場合には 1、配偶者がいない (無配偶：婚姻経験のない人、および婚姻経験はあるが配偶者と離別または死別した人を含む) 場合には 0 をとる 2 値変数を用いる。この分析に用いる変数は、表 2 の通りである。

第 2.2 節で示した分析の手順にしたがって、第 1 段階として、処理変数である有配偶ダミーについて、共変量を説明変数として、ロジスティック回帰を行う。このロジスティック回帰の推定結果は、表 3 に示している。次に、このロジスティック回帰の推定結果にもとづき、各個人  $i$  に対する、有配偶者である確率 (有配偶者確率) を表す傾向スコア  $PX1_i$  を求める。この例では、表 3 の推定結果の係数の値を用いて、(1) 式のように求められる<sup>(2)</sup>。

$$\begin{aligned}
 PX1_i &= \frac{\exp[\hat{\alpha} + \hat{\beta} \cdot \mathbf{Z}_i]}{1 + \exp[\hat{\alpha} + \hat{\beta} \cdot \mathbf{Z}_i]} \\
 &= \frac{\exp[-0.769 - 0.399 \cdot \text{性別}_i + 2.056 \cdot 30代_i + 2.770 \cdot 40代_i + \dots + 1.548 \cdot 60代男性_i]}{1 + \exp[-0.769 - 0.399 \cdot \text{性別}_i + 2.056 \cdot 30代_i + 2.770 \cdot 40代_i + \dots + 1.548 \cdot 60代男性_i]}
 \end{aligned}
 \tag{1}$$

第3段階として、(1)式により求められた傾向スコア  $PX1$  を用いて次の式の通りに、ウェイト  $WT1$  を得る。

$$WT1 = \frac{X1}{PX1} + \frac{(1 - X1)}{(1 - PX1)} \tag{2}$$

表2 変数の定義

| 変数名                 | 変数の定義                       |
|---------------------|-----------------------------|
| 従属変数                |                             |
| 家庭生活満足度             | 1 = 不満 ~ 5 = 満足             |
| 処理変数                |                             |
| 有配偶                 | 有配偶者 = 1、無配偶者（離別・死別者含む） = 0 |
| 共変量（説明変数）           |                             |
| 性別                  | 男性 = 1、女性 = 0               |
| 年齢の年代ダミー（20代がベース）   |                             |
| 30代                 | 年齢が30代 = 1、それ以外 = 0         |
| 40代                 | 年齢が40代 = 1、それ以外 = 0         |
| 50代                 | 年齢が50代 = 1、それ以外 = 0         |
| 60代                 | 年齢が60代 = 1、それ以外 = 0         |
| 学歴ダミー               | （高卒未満がベース）                  |
| 高卒                  | 学歴が高校卒業 = 1、それ以外 = 0        |
| 大卒                  | 学歴が大学・大学院卒業 = 1、それ以外 = 0    |
| 市郡規模ダミー（町村がベース）     |                             |
| 大都市                 | 居住地が大都市 = 1、それ以外 = 0        |
| その他の市               | 居住地が大都市以外の市 = 1、それ以外 = 0    |
| 調査年ダミー（2000年調査がベース） |                             |
| 2001年               | 調査を2001年に実施 = 1、それ以外 = 0    |
| 2002年               | 調査を2002年に実施 = 1、それ以外 = 0    |
| 2003年               | 調査を2003年に実施 = 1、それ以外 = 0    |
| 交差項                 |                             |
| 30代男性               | 30代の男性 = 1、それ以外 = 0         |
| 40代男性               | 40代の男性 = 1、それ以外 = 0         |
| 50代男性               | 50代の男性 = 1、それ以外 = 0         |
| 60代男性               | 60代の男性 = 1、それ以外 = 0         |

表 3 有配偶者確率の推定結果（ロジスティック）

| 説明変数                       | 係数         | 標準誤差  |
|----------------------------|------------|-------|
| 切片                         | - 0.769*** | 0.125 |
| 性別                         | - 0.399*** | 0.117 |
| 年齢の年代ダミー（20代がベース）          |            |       |
| 30代                        | 2.056***   | 0.111 |
| 40代                        | 2.770***   | 0.125 |
| 50代                        | 2.475***   | 0.113 |
| 60代                        | 1.847***   | 0.110 |
| 学歴ダミー（高卒未満がベース）            |            |       |
| 高卒                         | 0.332***   | 0.078 |
| 大卒                         | 0.146*     | 0.086 |
| 市郡規模ダミー（町村がベース）            |            |       |
| 大都市                        | - 0.312*** | 0.082 |
| その他の市                      | - 0.113*   | 0.067 |
| 調査年ダミー（2000年調査がベース）        |            |       |
| 2001年                      | 0.022      | 0.077 |
| 2002年                      | - 0.013    | 0.075 |
| 2003年                      | - 0.012    | 0.072 |
| 交差項                        |            |       |
| 30代男性                      | - 0.257    | 0.162 |
| 40代男性                      | - 0.096    | 0.177 |
| 50代男性                      | 0.668***   | 0.168 |
| 60代男性                      | 1.548***   | 0.169 |
| - 2 対数尤度                   | 9241.075   |       |
| Cox & Snell R <sup>2</sup> | 0.177      |       |
| Nagelkerke R <sup>2</sup>  | 0.263      |       |
| 標本数                        | 10063      |       |

注：\*は10%、\*\*は5%、\*\*\*は1%水準で統計的に有意

この傾向スコアにより算出されたウェイトを、処理変数  $X1$  にかけることで求められる、ウェイト付き処理変数  $WTX1_T$  の記述統計量を示したのが、表 4 である。表 4 から、以上の手続きにより処理群  $WTX1_T$  と統制群  $WTX1_C$  の頻度が変わり、またその平均値の差が 0.001 と非常に小さいことから、これらが一様分布となっていることが確認できる。

表 4 有配偶者確率の記述統計量（標本数 = 10,063）

|                                | 最小値  | 最大値   | 合計       | 平均値   | 標準偏差  |
|--------------------------------|------|-------|----------|-------|-------|
| $X1$ （有配偶 = 1、無配偶 = 0）         | 0    | 1     | 7601     | 0.755 | 0.430 |
| $WTX1_T \equiv WT1 * X1$       | 0.00 | 4.85  | 10056.42 | 0.999 | 0.729 |
| $WTX1_C \equiv WT1 * (1 - X1)$ | 0.00 | 13.92 | 10040.09 | 0.998 | 2.269 |

しかし、この傾向スコアの逆確率によるウェイトをかけただけでは、周辺分布が大きく変化してしまうことがある。そのため、実際の推計においては、表 4 で示した、ウェイト調整後の有配偶者確率の平均を、ウェイトの合計で割るといふ、さらなるウェイトの調整の手順を踏む必要がある。計算式は以下の通りである。

$$ADJ\_WT1 = \frac{WTX1_T * 7601}{10056.42} + \frac{WTX1_C * (10063 - 7601)}{10040.09}$$

この調整後のウェイトの値は、表 5 の通りである。最小値と最大値の差がより小さいほど、また平均値が 1 に近いほど (0.95 ~ 1.05 の間が許容範囲) よいとされることから、ここでのウェイトの調整はうまくいったといえる。

表 5 記述統計量 (標本数 = 10,063)

|         | 最小値  | 最大値  | 合計       | 平均値   | 標準偏差  |
|---------|------|------|----------|-------|-------|
| ADJ_WT1 | 0.30 | 3.66 | 10063.00 | 1.000 | 0.493 |

実際、表 6 に示しているように、処理変数の観測値 (ウェイト付けをしていない値) と傾向スコアの値との相関係数は有意であるのに対し、この調整されたウェイトにより、ウェイト付けした処理変数の値は、傾向スコアとの有意な相関関係はなくなっている。したがって、標本が  $X$  にランダムに割り当てられたような状態が生まれていることとなる。

表 6 処理変数と傾向スコアの相関係数

|               | 観測値<br>(ウェイト調整前) | ウェイト調整後 |
|---------------|------------------|---------|
| Pearson の相関係数 | 0.464***         | 0.002   |
| 有意確率 (両側)     | 0.000            | 0.820   |
| 標本数           | 10063            | 10063   |

注：\*\*\*は 1%水準で統計的に有意 (両側) を示す

以上の手順で求められたウェイト ADJ\_WT1 により、ケースに重み付けをしたデータを用いて、「家庭生活満足度」に配偶者の有無が与える影響を検証していく。家庭生活満足度の推計には、最小二乗法 (OLS) を用いる。推定結果は、表 7 の通りである。比較のために傾向スコアを用いない、単純な OLS による多変量解析の結果も示している。

表 7 より、配偶者がいる人の方が、いない人と比べて家庭生活の満足度が有意に高いことがわかる。また、ウェイト調整を行わない単純な多変量解析の結果と、傾向スコアを用いたウェイトによる調整を行った因果分析の推定結果とでは、この例では、有配偶かどうかは家庭生活満足度に与える影響に、大きな違いはみられないと結論づけられる。

他の説明変数について、多変量解析と傾向スコアを用いた因果分析の推定結果では、性別や交差項などいくつかの変数で係数の値や有意性が異なる場合がみられる。このような違いは、因果分析において、傾向スコアを用いてサンプルにウェイト調整した結果、起こりうることだろう。その違いを解釈することは困難であり、また、意味があることとはいえない。しかし、傾向スコアによる因果分析で、いくつかの説明変数が有意になるということは、第 2.2 節で示した「二重にロバストな推定量」を得るための仮定の一つである、「 $Y$  への  $X$  と  $\{Z\}$  の線形の影響」が成立している可能性が考えられ、その場合には、より頑強な推定結果が得られたこととなる。

表7 家庭生活満足度の推定結果（分析例1）

| 従属変数                   | 家庭生活満足度             |       | 家庭生活満足度                    |       |
|------------------------|---------------------|-------|----------------------------|-------|
|                        | 多変量解析<br>(ウェイト調整なし) |       | 傾向スコアによる因果分析<br>(ウェイト調整あり) |       |
| 説明変数                   | 係数                  | 標準誤差  | 係数                         | 標準誤差  |
| 切片                     | 3.250***            | 0.051 | 3.128***                   | 0.053 |
| 有配偶 = 1 / 無配偶 = 0      | 0.401***            | 0.025 | 0.413***                   | 0.023 |
| 性別                     | 0.022               | 0.052 | 0.277***                   | 0.052 |
| 年齢の年代ダミー（20代がベース）      |                     |       |                            |       |
| 30代                    | -0.275***           | 0.049 | -0.137***                  | 0.048 |
| 40代                    | -0.364***           | 0.049 | -0.208***                  | 0.047 |
| 50代                    | -0.266***           | 0.047 | -0.123***                  | 0.045 |
| 60代                    | -0.145***           | 0.049 | -0.025                     | 0.047 |
| 学歴ダミー（高卒未満がベース）        |                     |       |                            |       |
| 高卒                     | 0.137***            | 0.028 | 0.101***                   | 0.028 |
| 大卒                     | 0.305***            | 0.032 | 0.259***                   | 0.031 |
| 市郡規模ダミー（町村がベース）        |                     |       |                            |       |
| 大都市                    | 0.032               | 0.031 | 0.059*                     | 0.031 |
| その他の市                  | 0.048**             | 0.024 | 0.066***                   | 0.024 |
| 調査年ダミー（2000年調査がベース）    |                     |       |                            |       |
| 2001年                  | -0.024              | 0.028 | -0.016                     | 0.029 |
| 2002年                  | -0.029              | 0.028 | -0.013                     | 0.028 |
| 2003年                  | 0.050               | 0.027 | 0.037                      | 0.027 |
| 交差項                    |                     |       |                            |       |
| 30代男性                  | 0.080               | 0.070 | -0.154**                   | 0.070 |
| 40代男性                  | 0.047               | 0.068 | -0.269***                  | 0.068 |
| 50代男性                  | -0.051              | 0.065 | -0.355***                  | 0.065 |
| 60代男性                  | -0.121*             | 0.066 | -0.037***                  | 0.066 |
| 自由度修正済み R <sup>2</sup> | 0.037               |       | 0.057                      |       |
| 標本数                    | 10063               |       | 10063                      |       |

注：\*は10%、\*\*は5%、\*\*\*は1%水準で統計的に有意

4.2 分析例2：家庭生活満足度 - 有配偶者 vs 離別・死別者 -

次に、未婚者を分析から除き、結婚経験のある人のみを分析の対象として、現在、配偶者がいるかどうか、家庭生活満足度に与える影響について考察する。分析対象は、8,360人である。従属変数と説明変数には、分析例1と同じ変数を用いる。処理変数 X2 については、「結婚経験はあるが、離別か死別により現在は配偶者がいない人（離死別者）は1、結婚の経験がありかつ現在配偶者がいる人（有配偶者）は0」をとる2値変数とする。

第4.1節で示した、分析例1の場合と同じ推計手順により、傾向スコアを用いたウェイト調整を含む因果分析を行う。はじめに、処理変数である離死別者のダミー変数について、表2に示した説明変数によりロジスティック回帰を行う。離死別者ダミーのロジスティック回帰の推定結果を示した表8より、離死別者確率としての傾向スコア PX2 を、(3)式の通りに求める。

$$PX2_i = \frac{\exp[-2.832 - 0.300 \cdot \text{性別}_i + 0.385 \cdot 30\text{代}_i + 0.460 \cdot 40\text{代}_i + \dots - 1.034 \cdot 60\text{代男性}_i]}{1 + \exp[-2.832 - 0.300 \cdot \text{性別}_i + 0.385 \cdot 30\text{代}_i + 0.460 \cdot 40\text{代}_i + \dots - 1.034 \cdot 60\text{代男性}_i]} \quad (3)$$

表 8 離死別者確率の推定結果（ロジスティック）

| 説明変数                       | 係数         | 標準誤差  |
|----------------------------|------------|-------|
| 切片                         | - 2.832*** | 0.333 |
| 性別                         | - 0.300    | 0.517 |
| 年齢の年代ダミー（20代がベース）          |            |       |
| 30代                        | 0.385      | 0.340 |
| 40代                        | 0.460      | 0.332 |
| 50代                        | 1.019***   | 0.321 |
| 60代                        | 1.727***   | 0.319 |
| 学歴ダミー（高卒未満がベース）            |            |       |
| 高卒                         | - 0.192**  | 0.124 |
| 大卒                         | - 0.537*** | 0.096 |
| 市郡規模ダミー（町村がベース）            |            |       |
| 大都市                        | 0.169      | 0.082 |
| その他の市                      | 0.092      | 0.067 |
| 調査年ダミー（2000年調査がベース）        |            |       |
| 2001年                      | - 0.120    | 0.116 |
| 2002年                      | - 0.055    | 0.112 |
| 2003年                      | - 0.063    | 0.106 |
| 交差項                        |            |       |
| 30代男性                      | - 0.483    | 0.597 |
| 40代男性                      | - 0.248    | 0.565 |
| 50代男性                      | - 0.476    | 0.540 |
| 60代男性                      | - 1.034*   | 0.535 |
| - 2 対数尤度                   | 4728.159   |       |
| Cox & Snell R <sup>2</sup> | 0.042      |       |
| Nagelkerke R <sup>2</sup>  | 0.093      |       |
| 標本数                        | 8360       |       |

注：\*は 10%、\*\*は 5%、\*\*\*は 1%水準で統計的に有意

次の段階として、その傾向スコアの逆確率を、(2)式のように、処理変数  $X_2$  の値に応じてかけることで、ウェイトを算出する。そのウェイトを用いて調整をした処理変数の記述統計量を示したのが、表 9 の第 2・3 行目である（処理群： $WTX2_T \equiv WT2 * X_2$ 、統制群： $WTX2_C \equiv WT2 * (1 - X_2)$ ）。これより、 $WTX2_T$  と  $WTX2_C$  が一様分布となっていることが確認できるため、次のステップに進むことができる。

このウェイト付けした離死別者確率の平均を用いて、さらに以下のようにウェイトを調整する。

$$ADJ\_WT2 = \frac{WTX2_T * 759}{8289.80} + \frac{WTX2_C * (8360 - 759)}{8360.67}$$

表 9 の  $ADJ\_WT2$  の値を見ると、最小値と最大値の差はより小さくなり、かつ平均値も 1.000 となっていることから、このウェイト調整は適切であるといえる。

表 9 離死別者確率の記述統計量（標本数 = 8,360）

|                                 | 最小値  | 最大値   | 合計      | 平均値   | 標準偏差  |
|---------------------------------|------|-------|---------|-------|-------|
| $X_2$ （離死別 = 1、有配偶 = 0）         | 0    | 1     | 759     | 0.090 | 0.287 |
| $WTX2_T \equiv WT2 * X_2$       | 0.00 | 45.53 | 8289.80 | 0.992 | 3.817 |
| $WTX2_C \equiv WT2 * (1 - X_2)$ | 0.00 | 1.39  | 8360.67 | 1.000 | 0.326 |
| $ADJ\_WT2$                      | 0.32 | 4.17  | 8360.00 | 1.000 | 0.211 |

ADJ\_WT2のウェイトを用いてケースに重み付けしたデータにより、結婚経験のある人のなかで、現在の配偶者の有無による家庭生活満足度への影響を、OLSにより推計した結果を示したのが、表10である。

ウェイトによる調整をしていない多変量解析による結果と、傾向スコアを用いてウェイト調整を行った因果分析の推定結果は、ともに、結婚経験はあるが、配偶者とは離別または死別している人の方が、有意に家庭生活の満足度が低いことを示している。しかしながら、その係数の値を比較すると、多変量解析の場合には 0.350、傾向スコアを用いた因果分析の場合には 0.476 と、その影響の大きさに違いがみられることがわかる。

このように、多変量解析と傾向スコアを用いたウェイトにより調整した推計の結果（係数の値）が異なる場合には、通常多変量解析では、処理を受ける（離死別者となる）ことに対する選択バイアスをうまく取り除くことができず、傾向スコアを用いた因果分析の推定結果の方が、より適切であると結論づけることができる。

他の説明変数を見ると、この分析例では係数の値に大きな違いはみられない。ただし、有意な変数にも変化がないことから、傾向スコアによる因果分析の推定においても、家庭生活満足度への説明変数の線形の関係は存在しており、この推定結果はより頑強なものであると考えられる。

表10 家庭生活満足度の推定結果（分析例2）

| 従属変数                     | 家庭生活満足度    |       | 家庭生活満足度      |       |
|--------------------------|------------|-------|--------------|-------|
|                          | 多変量解析      |       | 傾向スコアによる因果分析 |       |
|                          | （ウェイト調整なし） |       | （ウェイト調整あり）   |       |
| 説明変数                     | 係数         | 標準誤差  | 係数           | 標準誤差  |
| 切片                       | 3.448***   | 0.071 | 3.466***     | 0.071 |
| <b>離死別 = 1 / 有配偶 = 0</b> | - 0.350*** | 0.038 | - 0.476***   | 0.037 |
| 性別                       | 0.373***   | 0.094 | 0.332***     | 0.094 |
| 年齢の年代ダミー（20代がベース）        |            |       |              |       |
| 30代                      | - 0.083    | 0.069 | - 0.084      | 0.069 |
| 40代                      | - 0.169**  | 0.067 | - 0.172**    | 0.067 |
| 50代                      | - 0.070    | 0.066 | - 0.081      | 0.066 |
| 60代                      | 0.048      | 0.068 | 0.022        | 0.068 |
| 学歴ダミー（高卒未満がベース）          |            |       |              |       |
| 高卒                       | 0.140***   | 0.030 | 0.135***     | 0.030 |
| 大卒                       | 0.289***   | 0.034 | 0.282***     | 0.034 |
| 市郡規模ダミー（町村がベース）          |            |       |              |       |
| 大都市                      | 0.041      | 0.034 | 0.053        | 0.034 |
| その他の市                    | 0.056**    | 0.026 | 0.060**      | 0.026 |
| 調査年ダミー（2000年調査がベース）      |            |       |              |       |
| 2001年                    | - 0.032    | 0.031 | - 0.044      | 0.031 |
| 2002年                    | - 0.037    | 0.031 | - 0.034      | 0.031 |
| 2003年                    | 0.049*     | 0.029 | 0.050*       | 0.029 |
| 交差項                      |            |       |              |       |
| 30代男性                    | - 0.203*   | 0.108 | - 0.218**    | 0.108 |
| 40代男性                    | - 0.240**  | 0.105 | - 0.226**    | 0.105 |
| 50代男性                    | - 0.390*** | 0.102 | - 0.356***   | 0.102 |
| 60代男性                    | - 0.475*** | 0.102 | - 0.421***   | 0.102 |
| 自由度修正済み R <sup>2</sup>   | 0.031      |       | 0.036        |       |
| 標本数                      | 8360       |       | 8360         |       |

注：\*は10%、\*\*は5%、\*\*\*は1%水準で統計的に有意

## 5. おわりに

2009年の統計分析セミナーでは、標本のランダムな割り当てが不可能な相関研究において、因果効果を分析する方法として生まれた概念である、「傾向スコア」が取り上げられた。ある処理(treatment)を受けることの影響を、因果関係分析により検証する際に、傾向スコアを用いて、各ケースにウェイトによる重み付けをすることで、処理を受けることに対する選択バイアスの問題を回避することができ、より信憑性の高い推計結果が得られることになる。

ただし、本稿で取り上げた、家庭生活満足度への婚姻状態の影響に関する分析演習例のように、通常の多変量解析と、傾向スコアによる調整を含む因果関係分析とでは、その推定結果にほとんど違いがみられない(係数の大きさや有意かどうかなど)場合が多くある。しかしながら、検証しようとしている事柄に、明らかに選択バイアス(標本がランダムに抽出されていない等)の問題がある場合には、多変量解析のみでは推定量に偏りが生じることが知られているため、傾向スコアを用いた調整を含む因果分析を試みるのが、今日の研究において求められているといえる。

傾向スコアを用いた因果関係分析は、日本ではまだ適用頻度が低いものの、それを用いた研究分析の論文などは、徐々に増えつつある。本稿では、傾向スコアを用いた因果分析の基礎的な部分しか取り上げていない。より詳しい理論的枠組みの学習や、今後の研究に傾向スコアを用いることを考える際には、それらの文献を参考にし、より深く理解されることをお勧めする。

### [Acknowledgement]

日本版 General Social Surveys (JGSS) は、大阪商業大学比較地域研究所が、文部科学省から学術フロンティア推進拠点としての指定を受けて(1999-2003年度)、東京大学社会科学研究所と共同で実施している研究プロジェクトである(研究代表: 谷岡一郎・仁田道夫、代表幹事: 佐藤博樹・岩井紀子、事務局長: 大澤美苗)。東京大学社会科学研究所附属日本社会研究情報センターSSJ データアーカイブがデータの作成と配布を行っている。

### [注]

- (1) IPTE法は、JGSSデータのように標本数が多い場合(情報量が多い場合)には有効であるため、セミナーではこの方法が取り上げられた。なお、標本数が少ない場合には、マッチング法が有効であるとされる。マッチング法について、日本語の文献では星野(2009)および菅(2009)が詳しい。
- (2) SPSSのシンタックスでは、ロジスティック回帰のコマンドに、“/SAVE=PRED”というオプションを付けることで、回帰式の被説明変数の推定値(ロジスティック回帰の場合には、被説明変数  $X_i = 1$  となる確率)を簡単に得る(保存する)ことができる(この推定値は、pre\_1という名前で保存される)。

### [参考文献]

- 星野崇宏, 2009, 『調査観察データの統計科学 因果推論・選択バイアス・データ融合』(シリーズ 確率と情報の科学), 岩波書店.
- Imbens, G. W., 2000, “The Role of the Propensity Score in Estimating Dose-response Functions,” *Biometrika*, 87(3), 706 - 710.
- 菅万里, 2009, 「母親の就労が思春期の子供の行動・学業に及ぼす効果: Propensity Score Matching による検証」, 東京大学社会科学研究所パネル調査プロジェクト ディスカッションペーパーシリーズ, No. 28. ([http://ssjda.iss.u-tokyo.ac.jp/panel/youthandmiddle/DP/PanelDP\\_028Kan.pdf](http://ssjda.iss.u-tokyo.ac.jp/panel/youthandmiddle/DP/PanelDP_028Kan.pdf), 最終閲覧日: 2010年2月1日)
- Rosenbaum, P. R. and Rubin, D. B., 1983, “The Central Role of the Propensity Score in Observational Studies for Causal Effects,” *Biometrika*, 70(1), 41-55.
- Rubin, D. B., 1985, “The Use of Propensity Score in Applied Bayesian Inference,” Bernardo, J. M., De Groot, M. H., Lindley, D. V., and Smith, A. F. M. (eds.), *Bayesian Statistics*, Vol.2, 463-472, North-Holland: Elsevier Science Publisher B. V.