



図 7.6 就業状態の変化が女性の結婚意欲に及ぼす影響に関する構造モデル

今回の構造モデルは就業状態の変化が女性の結婚意欲に及ぼす影響についての構造モデルである。就業状態を非正規就業から正規就業への変化とした場合に、教育水準と所得の影響はその変化に正の影響を及ぼしている。正規雇用化は結婚意欲にやや弱いながらも正の影響を及ぼしており、所得の水準の方がやや強めの影響である。年齢は若い方で結婚意欲が高く、第1回調査時点の結婚意欲の自己相関も高い結果となっており、全体的な分散の多くを説明することがわかった。

ここまで共分散構造分析について、Rを用いた分析例を示してきた。一般的に共分散構造分析を行う場合はMplusやAmos、LISREL、CALISが有名であるが、Rによる推定結果はそれらとほぼ同様であることが指摘されており、フリーソフトならではの強みを活かす事が可能である。またRはオープンソースであるため、各分析で用いるパッケージは日々改善が行われているため、今後の共分散構造分析の発展に寄与することであろう。

例) graphviz における図表作成コード

graphviz における図表作成コード

```

digraph sem {
  rankdir=LR;
  size="15";
  node[fontname="Helvetica" fontsize=14 shape=box];
  edge[fontname="Helvetica" fontsize=10];
  center=1;
  [rank=same; w4edu wcomp w4inc w4age];
  [rank=min; w1cp];
  [rank=max; w4mrmt];
  graph[size="15",label="Structure Equation Model",
  labelloc=t,labeljust=c,rankdir=LR];
  w1mrmt [shape=box,filesize=true,width=1.0,height=0.5];
  w4mrmt [shape=box,filesize=true,width=1.0,height=0.5];
  wcomp [shape=box,filesize=true,width=1.0,height=0.5];
  w4edu [shape=box,filesize=true,width=1.0,height=0.5];
  w4inc [shape=box,filesize=true,width=1.0,height=0.5];
  w4age [shape=box,filesize=true,width=1.0,height=0.5];
  w1cp [shape=box,filesize=true,width=1.0,height=0.5];
  w1mrmt -> w4mrmt [label="0.623"];
  w1cp -> w1mrmt [label="0.095"];
  wcomp -> w4mrmt [label="0.041"];
  w4edu -> wcomp [label="0.303"];
  w4edu -> w4inc [label="0.109"];
  w4inc -> w4mrmt [label="0.060"];
  w4inc -> wcomp [label="0.121"];
  w4age -> w4mrmt [label="-0.231"];
  w1mrmt -> w1mrmt [label="0.991",dir=both];
  w4mrmt -> w4mrmt [label="0.514",dir=both];
  wcomp -> wcomp [label="0.885",dir=both];
  w4age -> w4age [label="1",dir=both];
  w4edu -> w4edu [label="1",dir=both];
  w4inc -> w4inc [label="0.988",dir=both];
  w1cp -> w1cp [label="1",dir=both];
}

```

## 参考文献

McArdle J. John and Fumiaki Hamagami, 2006. “*Structural Equation Modeling in Longitudinal Research*”, Longitudinal Research Institute Workshop. Section B, Section Q, Section S.

Hamagami Fumiaki, 2005. “Multiple Time Categorical Factor Models”, APA-LRI Workshop.

釜野さおり, 2010, 「21世紀成年者縦断調査（第1回～第5回）における男性票の脱落者・継続回答者の特性に関する分析」, 金子隆一編『パネル調査（縦断調査）に関する総合的高度統計分析システムに関する開発研究』（厚生労働科学研究費補助金）平成21年度総括報告書.

厚生労働省大臣官房統計情報部, 2006. 「第4回21世紀成年者縦断調査（国民の生活に関する継続調査）結果の概況」.

小林淑恵, 2006. 「結婚・就業に関する意識と家族形成ツ循環モデルによる検証ツ」, 『人口学研究』第33号, pp.1-18.

第一生命経済研究所, 2008. 『景気見通しと結婚・出産意欲の関係』調査結果.

竹内啓（監修）, 豊田秀樹（著者）, 1992. 『SASによる共分散構造分析SASで学ぶ統計的データ解析』, 東京大学出版会.

豊田秀樹, 2009, 『共分散構造分析実践編—構造方程式モデリング（統計ライブラリー）』, 朝倉書店.

永瀬伸子, 2002. 「若年層の雇用の非正規化と結婚行動」, 『人口問題研究』, 58-2, pp.22-35.

福田節也, 2006. 「独身者の結婚意欲ならびに有配偶者の希望子ども数に関する分析：『21世紀成年者縦断調査』を用いた分析事例」, 金子隆一編, 『パネル調査（縦断調査）のデータマネジメント方策及び分析に関する総合的システムの開発研究』厚生労働科学研究費, 平成16-17年度総合報告書, pp.343-367.

守泉理恵・釜野さおり（2009）「21世紀成年者縦断調査（第1回～第5回）における女性票の脱落者・継続回答者の特性に関する分析」, 金子隆一編『パネル調査（縦断調査）に関する総合的高度統計分析システムに関する開発研究』（厚生労働科学研究費補助金）平成20年度総括報告書.

## 第8章

# 傾向スコア・脱落サンプルバイアスの検定法

本章では、パネル調査における脱落の影響の度合いの測定方法について適用例を含めて紹介する。

### 8.1 はじめに

パネル調査においては、同一対象者を繰り返し調査する際に、回答を途中でやめてしまう脱落（パネルの摩耗）の問題が生じる。パネル調査における脱落を「データの欠損」という観点から考えると（岩崎 2002, 金子 2005, 星野 2009）、1) 完全にランダムな欠損（脱落）MCAR: missing completely at random、2) ランダムな欠損（脱落）MAR: missing at random、3) ランダムでない欠損（脱落）NMAR: not missing at random もしくは無視できない欠損（nonignorable missing）に分けることができる。脱落状況が完全にランダムである場合、統計モデルにおける推定値に影響しないと考えられるため、そのままのモデル推定を行ってよい。ただし、ランダムな脱落およびランダムでない脱落である場合は、選択バイアス等によって推定結果を歪める結果となるため、脱落および欠損値を適切に処理するか、推定モデルの修正を行い推定値の調整を行う必要が生じる。

本章においては、脱落の発生が初婚の発生に対する因果効果について、初婚を規定していると考えられる共変量を調整したうえで測定する方法についての検討を行う。共変量の調整には傾向スコアを用いたマッチングによる因果効果の推定ならびにサンプルバイアスの検定を行う。

#### 8.1.1 脱落データへの対処

岩崎（2002）はデータの欠損一般に対する対処として、1) 欠損のある個体を取り除き、完全データとして解析、2) 個体のどこかの変量に欠損があっても、他の得られた観測値を用いて解析、3) 欠損値に値を代入（補完・埋め込み）し、完全データとして解析、4) 欠損はそのままモデル化して解析といった方法があるとしている。金子（2005）では、1) 欠損値標本の削除、2) 欠損値変数の削除、3) ダミー変数法、すなわち「欠損値をカテゴリ変数における一つのカテゴリ変数と同様に扱



い、ダミー変数をたてる方法」(p.60)、4) 代入による方法(「欠損値に何らかの統計的方法による推定値を代入する方法の総称」)(同上)、5) 最尤推定法(「統計モデルのパラメータの最尤推定の際に、欠損値の発生確率をもとにした尤度を組み込み、欠損値の発生を考慮した推定を行う方法」)の5種の方法が紹介されている。星野(2009)は、ランダムな脱落である場合、2つの条件において対処が異なるとし、1)「脱落する確率に影響を与える変数と従属変数との回帰関係が明確な場合」(p.166)、回帰分析を最尤推定し従属変数の周辺分布の母数を推定する。2)「脱落するかどうかのモデリングが明確な場合」(同上)、傾向スコアを利用して、IPW推定量や二重にロバストな推定量を用いることがその対処法であることを示している。また、モデルにおいてランダムでない脱落が生じる場合については、Diggle and Kenward(1994)の「ロジスティック回帰モデルの説明変数に『正規分布に従う変量効果』」(p.167)が入ったモデルや、近年では感度分析がよく使われていること、そして古典的な脱落の調整モデルとしてHeckman and Wise(1979)のプロビット選択モデルによる二段階推定が対処法として考えられるとしている。ただし、ヘックマンの二段階推定を脱落バイアスの調整に用いるには問題点があることが指摘されている(坂本2006, 直井2007)。第一に、モデルが非線形であるときに使用できない、第二に、対象サンプルと独立で、かつ従属変数に影響を与える外生変数が必要となるなど、分析手法の摘要には解決すべき点が多い。

### 8.1.2 21世紀成年者縦断調査における脱落の発生状況

21世紀縦断調査における脱落状況の把握とその要因分析についての多くの成果が報告されている。出生児縦断調査では、福田・金子(2005)が第1回と第2回の調査を用いて発生状況および脱落要因の分析を行っており、12点の知見を得ている。主なものでは、母親の属性に依存し、夫婦の年齢が若い、年齢差がある、婚前妊娠結婚である、方親家庭、非嫡出出生、外国籍を持つ親であること、有業女性であること等の状況において脱落が多い傾向にあることが指摘されている。ただし、モデリングが適切である場合は多変量解析による推定結果に有意な影響を与えないことを先行研究から指摘している。西野・金子(2005)も第1回と第2回の調査を用いて居住地移動と脱落の関係を考察しており、出産が転居を伴い、居住地変更を届出する層には大都市居住、核家族、高学歴等の影響が多いことを指摘し、脱落防止策に関する提言を行っている。西野(2006, 2007, 2008)はさらに、第3回(2006)、第4回(2007)、第5回(2008)までの脱落・居住地移動・復活サンプルの分析において、父母のどちらかが外国人、ひとり親世帯、育児への関わりが弱いケース、収入が少ないケース等、政策的な支援が必要な対象で脱落が生じていることを指摘している。また、調査を重ねるごとに累積して脱落が増えていることを反映して、影響は大きくなっているとしている。

成年者縦断調査では、福田(2008)が第4回までの結果を用い、脱落による推定値への影響について初婚を従属変数としたイベントヒストリー分析のモデルの一つであるネステッド・ロジットモデルによって検証している。ネステッド・ロジットモデルを用いることによって、1)初婚と脱落が独立に生起しているかを検証し、2)脱落を打ち切りとして扱った場合に独立変数の推定値にどの程度の影響があるかを検証できるとしている。第4回調査までの未婚女性の脱落がそれぞれ第1回から第2回で1162ケース(15.3%)、第2回から第3回で880(14.4%)、第3回から第4回で574(11.7%)となっているため、初婚を対象とした離散時間ロジット分析を行う場合、重大な

バイアスが生じていることを明らかにしている。その理由として、初婚と脱落が独立しておらず競合イベントとして生じていることが指摘されている。そのような競合イベントを用いる場合には、ネステッド・ロジットモデルが競合リスクをもつ要因（「非観察要因の相関」 p.109）を統制した上で推定を行う点が簡便かつ有用であるとしている。守泉・釜野（2009）では、女性票の脱落・継続回答者の属性分析を行っている。成年者調査の脱落率は出生児調査に比べて高く、第5回までで49.2%に達しており、無配偶、同棲している、低学歴、子どもなし、有業女性で脱落が多くみられることが指摘されている。平成19年に実施した第6回調査結果においては（厚生労働省2009）、回収客体数は男女計で第1回の調査客体数33689ケースから16444ケースと脱落が半数を超えている。調査ごとの回収率は約8割から9割であるが、継続率が低い状況にある。

21世紀縦断調査以外の脱落に関する知見として、Fitzgerald et al. (1998) は PSID (Panel Study of Income Dynamics) の脱落研究において、所得の減少、無配偶、移動したケースにおいて脱落がみられることを明らかにしているものの、対象者の属性と脱落が系統的であっても、推定値に与える影響は大きくないとしている。坂本（2006）は「消費生活に関するパネル調査」を用いた脱落の規定要因と推計バイアスの検証において、1) 転居+転居先不明、2) 多忙、3) 結婚・出産、4) 離婚・別居といったライフイベント前後で脱落が生じる傾向があること、無配偶では低収入、有配偶では収入の変化が大きいことを示している。また、脱落の影響の調整として IPW 推定量を用い、結婚したケースが脱落することによって選択バイアスが生じていること、結婚に影響を与える変数の推定値が過小になることを示している。直井（2007）は、「慶應義塾家計パネル調査」の脱落分析を行い、調査プロセス情報を除外変数としたヘックマンタイプのサンプルセレクションモデルの推定において、世帯の住居移動行動が脱落に対して選択バイアスを持っていることから、多くの変数の推定値を過大に推定していることが示されている。

以上の先行研究をもとに、本稿では、初脱落の発生が初婚の発生に対する因果効果について、初婚を規定していると考えられる共変量を調整したうえで測定する方法についての検討を行う。データは21世紀縦断調査の第1回から第5回の女性票を用い、共変量については福田（2008）を参考に構築し、共変量の調整指標である傾向スコアによる調整について星野（2003, 2009）、星野・繁柘（2004）、星野・前田（2006）、坂本（2006）を参考に行う。因果効果の推定にはマッチングという手法を用い、脱落の初婚に対する因果効果の測定を行う。

## 8.2 傾向スコアおよび IPW 推定量

傾向スコアは、もともと無作為割り当てを行う実験研究の因果効果を、大規模調査を用いた「相関研究」において実現するための手法として Rosenbaum and Rubin (1983) によって提案された概念である（星野2009）。傾向スコアは、従属変数と独立変数の間の因果関係を推定するために、それ以外の変数で従属変数に関連する要因=共変量を一つの変数に縮約した得点のことをいう。その共変量を基準として、無作為割り当て状況を作り出すというアイデアである。近年では、傾向スコアを用いた調整は Web 調査においても利用されてきている（星野2003）。

傾向スコアの数理モデルを星野・繁柘（2004）に準拠すると、第  $i$  被験者の共変量を  $x_i$ 、割り当て変数を  $z_i$  とするとき、群 1 へ割り当てられる確率  $e_i$  は

$$e_i = p(z_i = 1|x_i) \text{ ただし } (0 \leq e_i \leq 1)$$

と示すことができ、これを第*i*被験者の傾向スコアという。傾向スコアを算出するには、一般に割り当て変数を従属変数としたロジスティック回帰分析を行うことにより得られる。 $\hat{\alpha}$ を最尤推定量であるとすると、以下のように推定される。

$$\hat{e}_i = \frac{1}{1 + \exp\{-\hat{\alpha}^t x_i\}}$$

そうして算出した傾向スコアを基準として調整を行うことによって、バイアスを減少させることができる。Rosenbaum and Rubin (1983)は傾向スコアを用いた不偏推定を行う方法として、以下の3手法をあげている。

1. マッチング：傾向スコアが等しい群の被験者をペアにして、その差をもって因果効果とするという方法である。傾向スコアが完全に一致するペアだけを抽出する場合、分析対象数が極端に減少してしまう事態が生じることが多く、そういった場合には、「処置群の観測値に対して最小の距離になる対照群の観測値をマッチングさせる最近傍マッチングや、最近傍マッチングを行った場合に『ある特定の距離以上になるときはマッチングしない』キャリパーマッチング (caliper matching) などを行なう」(星野 2009 p.64)。
2. 層別解析：傾向スコアを5つ程度に分けて、それぞれの層別での処置群と対照群の平均の計算と、全体としての効果の推定量を算出する。

3) 共分散分析：割り当て変数と傾向スコアを独立変数として線形回帰分析を行う。

傾向スコアの利点としては、従属変数と共変量の回帰モデルを仮定する必要がないことにあり、傾向スコアは共変量を1変数に縮約しているため、処置群と対照群において外生要因・共変量の影響が少ない場合でも分析に用いることができる。共変量と従属変数のモデル設定を行わなくてもよい。モデルの誤設定に強いことが示されている(星野 2009)。

対して問題点としては、対象となるカテゴリが3群以上ある場合、2群ごとに傾向スコアを算出する必要があるため、因果効果を求めるための母集団が2群ごとに異なってしまう点、マッチング・層別解析を行う際には標準誤差が正確に計算できず、周辺期待値の推定もできない点、マッチングに使用する傾向スコアは連続変数であるため、ペアを作る際には最近傍マッチングなどの距離を使う必要があり、そこに恣意性が生まれる点、マッチングを行う際にペア化されなかったデータが無駄になる点、共分散分析は線形を仮定しているが、傾向スコアは0から1をとるため、仮定を満たさないといった点がある(星野 2009)。

そこで、このような問題点を克服する方法としてIPW推定量が提案されている(Rubin 1985)。IPW推定量は傾向スコアによる重み付け推定法であり、傾向スコアの逆数による重み付け平均を示す。割り当て変数  $z=0$  のときの従属変数を  $y_0$ 、 $z=1$  のとき  $y_1$  とすると、以下ようになる。

$$\hat{E}(y_1) = \sum_{i=1}^N \frac{z_i y_i}{e_i} \bigg/ \sum_{i=1}^N \frac{z_i}{e_i} \quad \hat{E}(y_0) = \sum_{i=1}^N \frac{(1-z_i) y_i}{1-e_i} \bigg/ \sum_{i=1}^N \frac{(1-z_i)}{1-e_i}$$

$\hat{E}(y_1)$  は、真の傾向スコアがわかっていて、 $y_0$  と  $y_1$  が割り当て変数  $z$  と独立であり、かつ共変量  $x$  の  $y_1$  に割り当てられる確率が0から1に収まる、いわゆる「強く無視できる割り当て状況」で

あれば、 $\hat{E}(y_1)$  は不偏推定量になる。よって、IPW 推定量を算出することによって、傾向スコアでは未知であった各群の周辺期待値である  $\hat{E}(y_1)\hat{E}(y_0)$  を計算することが可能となり、当指標をモデル推定に用いることにより、選択バイアスの影響を調整することが可能となる。

### 8.3 脱落発生が初婚発生へ及ぼす因果効果

脱落の発生が脱落と競合するイベントであることが福田 (2008) によって指摘されていることは上述した。本稿ではそのような競合イベントに対して共変量の影響を調整したうえでの因果効果を抽出することが目的であり、ここでいう因果効果とは星野 (2009) で示されている「処置群での平均介入効果または因果効果」average Treatment Effect on the Treated: TET (もしくは ATT: Average Treatment effect for the Treated) を指す。

処置群での平均介入効果は、無作為割り当てを前提とした実験法における処置群における潜在的な従属変数の差の期待値を意味し、以下のように示すことができる。

$$TET (ATT) = E(y_1 - y_0 | z = 1)$$

ここで、 $y_1 - y_0$  は因果効果、 $z$  は「欠測インディケータ」を示し、 $z = 1$  は処置群を示す ( $z = 0$  は対照群)。

本分析においては、脱落の発生を欠測インディケータとした場合の、初婚の発生に対する因果効果を測定することといえる。ただし、ここで用いるデータは調査データであるため、無作為割り当てによる実験法による処置群と対照群に分ける状況を作り出すことができない。そのため、調査データをそのまま用いるだけでは、上記の TET (ATT) をそのまま測定することができない。そこで、前節で説明した初婚の発生に影響する共変量の影響を調整した傾向スコアを算出し、用いることによって、調査データを用いて疑似的に実験法における処置群と対照群に振り分ける状況を作り出し、TET (ATT) の測定を行う。本稿では、Rosenbaum and Rubin (1983) が示した3つの方法のうち、マッチングの手法を用いて検証する\*1。データは21世紀縦断調査の第1回から第5回の女性票を用いる。

#### 8.3.1 データ・記述統計、傾向スコアの算出

表 8.1 には、本モデルにおける記述統計量を示している。初婚は、子どものいない独身者としており、第5回調査時における累積発生数をダミー変数化している。脱落の発生は第2回から5回調査までをそれぞれダミー変数化している。共変量は第1回調査における学歴 (中学卒、高校卒、専門学校卒、短大・高専卒、大学・大学院卒)、職業 (中小企業、大企業、官公庁、自営業・会社役員、非正規雇用、学生、無職、その他)、就業時間 (対数化)、本人の年収 (対数化)、親との同居 (両親と同居、方親と同居、両親と別居)、結婚意欲 (絶対したい、どちらかといえばしたい、どちらともいえない、あまりしたくない、絶対したくない) を用いて、傾向スコアを算出している。傾

\*1 マッチングは R 2.12.1 Matchingpackage 4.7-6 を用いて分析を行い、その詳細については、Sekhon(2007) を参考にした。

向スコアの算出は、各脱落の発生を従属変数としたロジスティック回帰分析による。傾向スコアを用いたマッチングには欠損値のないバランスドパネルが必要となるため、初婚、脱落、共変量が全て回答されている6209ケースが分析対象である。初婚に関する共変量による脱落の発生を従属変数としたロジスティック回帰分析による傾向スコアの推定結果は表8.2に示している。

表 8.1 記述統計量

記述統計量					
	度数	平均値	標準偏差	最小値	最大値
初婚（第5回時点）	6209	0.159	0.366	0	1
脱落（第2回調査）	6209	0.136	0.343	0	1
脱落（第3回調査）	6209	0.231	0.422	0	1
脱落（第4回調査）	6209	0.292	0.455	0	1
脱落（第5回調査）	6209	0.353	0.478	0	1
学歴：中学卒	6209	0.018	0.133	0	1
学歴：高校卒	6209	0.275	0.446	0	1
学歴：専門学校卒	6209	0.186	0.389	0	1
学歴：短大・高専卒	6209	0.249	0.432	0	1
学歴：大学・大学院卒	6209	0.272	0.445	0	1
職業：中小企業（従業員数500人未満）	6209	0.315	0.464	0	1
職業：大企業（従業員数500人以上）	6209	0.134	0.341	0	1
職業：官公庁	6209	0.027	0.161	0	1
職業：自営業・会社役員	6209	0.030	0.170	0	1
職業：非正規雇用	6209	0.346	0.476	0	1
職業：学生	6209	0.030	0.170	0	1
職業：無職	6209	0.064	0.245	0	1
職業：その他	6209	0.055	0.228	0	1
就業時間（対数化）	6209	3.045	1.490	0.000	6.907
本人の年収（対数化）	6209	4.883	2.715	0.000	11.513
両親と同居	6209	0.741	0.438	0	1
片親と同居	6209	0.119	0.324	0	1
両親と別居	6209	0.140	0.347	0	1
結婚意欲：絶対したい	6209	0.343	0.475	0	1
結婚意欲：どちらかといえばしたい	6209	0.358	0.479	0	1
結婚意欲：どちらともいえない	6209	0.217	0.412	0	1
結婚意欲：あまりしたくない	6209	0.063	0.244	0	1
結婚意欲：絶対したくない	6209	0.018	0.134	0	1
傾向スコア（第2回調査における脱落）	6209	0.136	0.070	0.055	0.479
傾向スコア（第3回調査における脱落）	6209	0.231	0.078	0.108	0.537
傾向スコア（第4回調査における脱落）	6209	0.292	0.092	0.138	0.643
傾向スコア（第5回調査における脱落）	6209	0.353	0.097	0.139	0.722

表8.2の推定結果については、当モデルが脱落の規定要因の分析が目的ではなく、初婚の規定要因と考えられる共変量を脱落発生モデルで推定し、その予測値を傾向スコアとして作成するために推定している。よって、本モデルにおけるそれぞれの共変量の解釈については一定の留保が必要である。ただし、学歴、職業、本人の年収については一般的な社会経済的属性であるため、その差について結果を参照する。学歴は、第2回では高校卒を基準としたときの差はみられないが、第3回以降は短大・高専卒が負の関係を示しており、第5回では大学・大学院卒と負の関係を示してい

る。学歴が高校と同程度もしくは高い場合、脱落が生じにくいことを示している。次に職業についてであるが、中小企業就業者を基準としたときに、第2回調査では、大企業就業者、非正規雇用、その他で正の関係を示し、第3回調査以降では官公庁では負、第4回調査でのみ自営業・会社役員で正の関係を示している。本人の年収は対数化した数値において、正の方向で有意であった。

表 8.2 ロジスティック回帰分析による傾向スコアの推定結果

共変量	脱落：第2回調査		脱落：第3回調査		脱落：第4回調査		脱落：第5回調査	
	B	Exp(B)	B	Exp(B)	B	Exp(B)	B	Exp(B)
学歴：中学卒	.317	1.373	-.006	.994	-.256	.774	-.180	.835
学歴：高校卒								
学歴：専門学校卒	-.056	.946	.085	1.088	-.003	.997	-.094	.910
学歴：短大・高専卒	-.151	.860	-.170 +	.843	-.223 **	.800	-.305 **	.737
学歴：大学・大学院卒	-.107	.898	-.047	.954	-.072	.931	-.204 **	.815
職業：中小企業（従業員数500人未満）								
職業：大企業（従業員数500人以上）	.245 *	1.277	.177 +	1.193	.131	1.140	.232 **	1.261
職業：官公庁	-.321	.726	-.352 +	.703	-.431 *	.650	-.489 **	.613
職業：自営業・会社役員	.203	1.226	.285	1.329	.284 +	1.329	.250	1.283
職業：非正規雇用	.186 +	1.204	.210 **	1.233	.170 *	1.185	.205 **	1.227
職業：学生	-.011	.990	.204	1.227	.208	1.232	.319	1.376
職業：無職	.246	1.278	.103	1.108	.182	1.200	.206	1.229
職業：その他	.507 *	1.660	.376 +	1.456	.434 *	1.543	.422 *	1.525
就業時間（対数化）	-.012	.988	-.012	.988	.005	1.005	.016	1.016
本人の年収（対数化）	.034 *	1.035	.022 +	1.022	.020 +	1.020	.030 **	1.031
両親と同居								
片親と同居	.332 **	1.394	.126	1.134	.303 **	1.354	.171 *	1.186
両親と別居	1.288 **	3.627	1.029 **	2.798	1.116 **	3.054	1.074 **	2.926
結婚意欲：絶対したい	-.037	.964	.026	1.026	-.055	.946	-.070	.932
結婚意欲：どちらかといえばしたい	-.124	.884	-.155 +	.856	-.102	.903	-.132 +	.877
結婚意欲：どちらともいえない								
結婚意欲：あまりしたくない	-.296 +	.744	.024	1.025	-.124	.883	-.194	.823
結婚意欲：絶対したくない	-.036	.965	-.174	.840	-.280	.755	-.479 *	.619
定数項	-2.274 **	.103	-1.511 **	.221	-1.185 **	.306	-.892 **	.410
サンプル数	6209		6209		6209		6209	
脱落なし	5362		4773		4394		4020	
脱落あり	847		1436		1815		2189	
カイ2乗	217.167 **		193.867 **		241.360 **		249.785 **	
-2 対数尤度	4730.185		6521.973		7261.767		7809.653	
Cox-Snell R <sup>2</sup> 乗	.034		.031		.038		.039	
Nagelkerke R <sup>2</sup> 乗	.063		.047		.054		.054	

有意水準： 0.1 + 0.05 \* 0.01 \*\*

### 8.3.2 マッチングによる因果効果の測定

次に、推定された傾向スコアを用いてマッチングを行い、脱落の有無が初婚の発生に与える因果効果を測定する。分析は、第5回までの累積した初婚の発生に対して、第2回から第5回までの脱落の発生における脱落と非脱落の差が、共変量の影響を調整したうえでも、処置群での平均介入効果または因果効果 (TET) として統計的に有意になるかどうかについての検証となる。分析モデルは第2回調査～第5回調査までの脱落/非脱落の影響をみた4つである。さらに各モデルに対し、各共変量に対して調整前および調整後での処置群と対照群の平均値の差の計算や検定を行う。

マッチングとは、傾向スコアが等しい群の被験者をペアにして（ここでは脱落の有無）、その差をもって因果効果とするという方法である。本分析においては、傾向スコアのマッチング基準として、ケースは1対1でのペアを用いて検定を行い、傾向スコアのマッチングの際の許容範囲として、キャリパーマッチングを用い、最近傍マッチングにおける標準偏差 0.25 以上でペアを作成しないという基準でのマッチングを行った。また、マッチングの問題点の一つである、処置群および対照群一方のケース数に依存し、どちらかにケースが偏っている場合、多くのケース数の情報を失ってしまうことがある。本データにおいても、第2回から第3回調査までは脱落ケースの方が非脱落ケースよりも少ないために偏った配分となっている。そのため、情報をできるだけ分析結果に反映するために、対照群については、繰り返しケースを用いるという方法でマッチングを行った。すなわち、ある傾向スコアをもつ処置群（脱落ケース）と同じ傾向スコアをもつ対照群（非脱落ケース）の差の検定を一様に行うということである（ただし、それぞれの検定自体は1対1での検定となる。分析モデルによっては1対nの検定も行うことも可能である）。さらに、マッチングによる調整の前後での共変量の処置群（脱落ケース）と対照群（非脱落ケース）の間での平均値の差の測定および検定を行う。マッチング調整後に共変量の差がなくなっていることで、脱落の初婚発生に対する共変量の変動を調整し、処置群における因果効果 TET を推定することが可能となる。以下は、R (2.12.1) におけるコードである。従属変数 (Y) と脱落変数 (tre) を定義し、第2回から第5回までの脱落の有無を tre2 から tre5 とした。次に、Match() 関数の引数は、Y=に従属変数（初婚の有無）、Tr=に処置変数（脱落の有無）、X=に傾向スコア（変数 PRE.1）を設定する。M=1 はマッチングに用いるスカラーを示し、数値が1であるときは1対1のマッチングを行うことを指定する。caliper=0.25 はマッチング対象の選択において、標準偏差 0.25 までの範囲でマッチング対象を決めるという意味である。傾向スコアがまったく一致することは少ないため、ある程度の幅をもたせてマッチングを行うときに指定する。replace=T はマッチングの際にマッチング対象の交換をするかどうかについてであり、T(true) であれば行うという意味である。estimated=ATT は、推定する因果効果を ATT (the sample Average Treatment effect for the Treated) とするということを意味し、処置群の平均処置効果を推定する。その他のオプションには、ATE (the sample Average Treatment Effect)、ATC (the sample Average Treatment effect for the Controls) がある。Match() にはこれらの他にも多くの引数がある (sekhon 2007)。

## パッケージのインストールならびに傾向スコアの推定

```

library(Matching)
library(stats)

#従属変数 (Y) および脱落変数 (tre)
Y <- logit$mard
tre2 <- logit$ATR2
tre3 <- logit$ATR3
tre4 <- logit$ATR4
tre5 <- logit$ATR5

#第 2 回調査での脱落が第 5 回調査までの結婚選択に与える因果効果
#Y に従属変数, Tr に独立変数, X に傾向スコアの推定値を推定

mo1a <- Match(Y=Y, Tr=tre2, X=logit$PRE_1, M=1, caliper=.25, replace=T,
  estimand="ATT")
summary(mo1a)

```

例) Match( ) Summary ( ) アウトプット (第 2 回～第 5 回調査)

```

Estimate... -0.16838
AI SE..... 0.0085571
T-stat..... -19.677
p.val..... < 2.22e-16

Original number of observations..... 6209
Original number of treated obs..... 847
Matched number of observations..... 847
Matched number of observations (unweighted). 18262

Caliper (SDs)..... 0.25
Number of obs dropped by 'exact' or 'caliper' 0

```

表 8.3 はマッチングによる脱落の初婚に対する因果効果の推定結果を示している。第 2 回から第 5 回までの結果をみると、-0.15 から -0.16 といった推定値で安定しており、脱落が初婚に対する因果効果は脱落数が急増している割には安定した値となっている。推定値が示す処置群における因果効果 TET は、脱落の発生は初婚の発生に対して負の因果効果をもつことを示している。すなわち、脱落の発生は初婚の発生を過小にするという意味で、福田 (2008) で示されている競合性を裏付ける結果となっている。



表 8.3 マッチングによる脱落の初婚に対する因果効果の推定結果

	第2回調査	第3回調査	第4回調査	第5回調査
推定値	-0.16838	-0.15267	-0.16034	-0.16457
標準誤差	0.008557	0.00895	0.008894	0.009697
t値	-19.677 **	-17.059 **	-18.029 **	-16.972 **
観測値	6209	6209	6209	6209
脱落ケース数	847	1436	1815	2189
マッチング数	847	1436	1815	2189
マッチング数 (非加重)	18262	24650	29987	26079
キャリパー値 (標準偏差)	0.25	0.25	0.25	0.25

有意水準： 0.1 + 0.05 \* 0.01 \*\*

表 8.4 から表 8.7 はマッチングによる調整の前後での共変量の処置群 (脱落ケース) と対照群 (非脱落ケース) の間での平均値の差の測定および検定を行っている。いずれの結果においても共変量の平均の差は、調整を行うことで調整前にあった差はなくなっているカテゴリがほとんどである。よって、表 3 で示された処置群における因果効果 TET は共変量で調整した上で脱落の初婚に対する因果効果が推定されているといえることができる。

マッチング前後での共変量の平均の差の測定および検定は、MatchBalance() 関数を用いる。第 5 回までの脱落に対するマッチングの検定の例を以下に示すと、傾向スコアの算出に用いた各カテゴリを条件式の右辺に記す。match.out はマッチングのアウトプット (ここでは mo4a という名前で第 5 回目までの脱落の因果効果を示している) を示し、nboots=1000 はブートストラップサンプルの数を指定するもので、Kolmogorov-Smirnov test によってカテゴリカルな変数を用いた場合でも正確な共変量比較を可能にする。nboots=500 以上の設定をするのが望ましいとされている。アウトプットからは、脱落の有無によって生じているだろう差がマッチング後にはなくなっていることの確認を行う。例ではマッチング前とマッチング後で、学歴差がどのように変化したのかを確認でき、マッチング前は脱落の有無によって中卒は差があったのに対して、マッチング後はその差がなくなっていることがわかる。このような調整によって、脱落が初婚に対する真の因果効果を測定することが可能となる。

#### マッチング前後での共変量の平均の差の検定

```
mb4a <- MatchBalance(ATR5 ~ e1jhd + e1hd + e1sd + e1tkd + e1ud + j1c + j1d
+ j1p + j1j + j1hs + j1t + j1k + j1o + lnltm + lnincome + dk1rd + dk1kd +
dk1rb + mi1 + mi2 + mi3 + mi4 + mi5
,match.out=mo4a, nboots=1000, data=logit)
```

例) MatchBalance( ) のアウトプット：学歴 (中学卒) の場合

```

***** (V1) eljhd *****
                                Before Matching          After Matching
mean treatment.....           0.024793                0.024793
mean control.....             0.016785                0.027056
std mean diff.....            5.1474                  -1.4541

mean raw eQQ diff.....         0.0070838              0.00087614
med raw eQQ diff.....           0                      0
max raw eQQ diff.....           1                      1

mean eCDF diff.....            0.0040043              0.00043807
med eCDF diff.....              0.0040043              0.00043807
max eCDF diff.....              0.0080086              0.00087614

var ratio (Tr/Co).....          1.4666                  0.91851
T-test p-value.....             0.15493                  0.7542

```

表 8.4 第 2 回調査における脱落に対する共変量の調整前後の平均差の検定

第2回調査における脱落	調整前				調整後			
	平均 (脱落)	平均 (非脱落)	平均比 (脱落/ 非脱落)	t検定	平均 (脱落)	平均 (非脱落)	平均比 (脱落/ 非脱落)	t検定
学歴：中学卒	0.02479	0.01679	1.46660		0.02479	0.02706	0.91851	
学歴：高校卒	0.28689	0.27280	1.03220		0.28689	0.30270	0.96928	
学歴：専門学校卒	0.19126	0.18557	1.02450		0.19126	0.18170	1.04030	
学歴：短大・高专卒	0.21606	0.25401	0.89475	*	0.21606	0.22001	0.98700	
学歴：大学・大学院卒	0.28099	0.27079	1.02420		0.28099	0.26853	1.02860	
職業：中小企業 (従業員数500人未満)	0.27981	0.32022	0.92668	*	0.27981	0.28227	0.99468	
職業：大企業 (従業員数500人以上)	0.14876	0.13148	1.11000		0.14876	0.14849	1.00150	
職業：官公庁	0.02361	0.02704	0.87713		0.02361	0.02486	0.95089	
職業：自営業・会社役員	0.03188	0.02965	1.07360		0.03188	0.03756	0.85366	
職業：非正規雇用	0.34238	0.34670	0.99507		0.34238	0.33970	1.00380	
職業：学生	0.02361	0.03059	0.77835		0.02361	0.02483	0.95206	
職業：無職	0.08028	0.06173	1.27610	+	0.08028	0.06977	1.13770	
職業：その他	0.06966	0.05259	1.30190	+	0.06966	0.07250	0.96376	
就業時間 (対数化)	2.94970	3.05980	1.17220	+	2.94970	2.98100	1.06950	
本人の年収 (対数化)	5.04790	4.85640	1.25950	+	5.04790	5.09290	1.10560	
両親と同居	0.57261	0.76762	1.37330	**	0.57261	0.66925	0.99806	
片親と同居	0.12515	0.11805	1.05260		0.12515	0.13146	0.95890	
両親と別居	0.30224	0.11414	2.08790	**	0.30224	0.29929	1.00560	
結婚意欲：絶対したい	0.34711	0.34278	1.00700		0.34711	0.34331	1.00520	
結婚意欲：どちらかといえばしたい	0.33884	0.36069	0.97250		0.33884	0.35312	0.98075	
結婚意欲：どちらともいえない	0.23731	0.21410	1.07670		0.23731	0.22945	1.02370	
結婚意欲：あまりしたくない	0.05549	0.06472	0.86677		0.05549	0.05143	1.07430	
結婚意欲：絶対したくない	0.02125	0.01772	1.19640		0.02125	0.02269	0.93812	

有意水準： 0.1 \* 0.05 \*\* 0.01 \*\*

表 8.5 第3回調査における脱落に対する共変量の調整前後の平均差の検定

第3回調査における脱落	調整前				調整後			
	平均 (脱落)	平均 (非脱落)	平均比 (脱落/ 非脱落)	t検定	平均 (脱落)	平均 (非脱落)	平均比 (脱落/ 非脱落)	t検定
学歴：中学卒	0.01880	0.01760	1.06760		0.01880	0.01635	1.14740	
学歴：高校卒	0.28203	0.27257	1.02170		0.28203	0.28464	0.99446	
学歴：専門学校卒	0.20474	0.18081	1.09980	*	0.20474	0.20209	1.00970	
学歴：短大・高专卒	0.21448	0.25917	0.87794	**	0.21448	0.22137	0.97745	
学歴：大学・大学院卒	0.27994	0.26985	1.02360		0.27994	0.27555	1.00980	
職業：中小企業（従業員数500人未満）	0.28552	0.32349	0.93261	**	0.28552	0.28082	1.01010	
職業：大企業（従業員数500人以上）	0.13928	0.13220	1.04540		0.13928	0.14832	0.94900	
職業：官公庁	0.02228	0.02787	0.80470		0.02228	0.02346	0.95096	
職業：自営業・会社役員	0.03343	0.02891	1.15130		0.03343	0.03374	0.99090	
職業：非正規雇用	0.35515	0.34339	1.01620		0.35515	0.34789	1.00950	
職業：学生	0.02925	0.02975	0.98409		0.02925	0.03091	0.94778	
職業：無職	0.07173	0.06202	1.14520		0.07173	0.06657	1.07150	
職業：その他	0.06337	0.05238	1.19640		0.06337	0.06828	0.93299	
就業時間（対数化）	2.97630	3.06540	1.10120	+	2.97630	2.98010	1.01450	
本人の年収（対数化）	4.95910	4.85950	1.16690		4.95910	4.99010	1.05390	
両親と同居	0.63997	0.77142	1.30730	**	0.63997	0.63956	0.99950	
片親と同居	0.11490	0.12026	0.96174		0.11490	0.11495	0.99963	
両親と別居	0.24513	0.10811	1.92000	**	0.24513	0.24549	0.99900	
結婚意欲：絶対したい	0.35655	0.33941	1.02370		0.35655	0.35277	1.00480	
結婚意欲：どちらかといえばしたい	0.32730	0.36686	0.94837	**	0.32730	0.33500	0.98833	
結婚意欲：どちらともいえない	0.22911	0.21370	1.05160		0.22911	0.22269	1.02030	
結婚意欲：あまりしたくない	0.06894	0.06181	1.10750		0.06894	0.07115	0.97124	
結婚意欲：絶対したくない	0.01811	0.01823	0.99393		0.01811	0.01839	0.98487	

有意水準：0.1 + 0.05 \* 0.01 \*\*

表 8.6 第4回調査における脱落に対する共変量の調整前後の平均差の検定

第4回調査における脱落	調整前				調整後			
	平均 (脱落)	平均 (非脱落)	平均比 (脱落/ 非脱落)	t検定	平均 (脱落)	平均 (非脱落)	平均比 (脱落/ 非脱落)	t検定
学歴：中学卒	0.01653	0.01843	0.89887		0.01653	0.01591	1.03820	
学歴：高校卒	0.28981	0.26855	1.04810	+	0.28981	0.29525	0.98915	
学歴：専門学校卒	0.19780	0.18161	1.06790		0.19780	0.18916	1.03450	
学歴：短大・高专卒	0.21433	0.26309	0.86884	**	0.21433	0.21669	0.99208	
学歴：大学・大学院卒	0.28154	0.26832	1.03060		0.28154	0.28300	0.99688	
職業：中小企業（従業員数500人未満）	0.29256	0.32385	0.94550	*	0.29256	0.30482	0.97671	
職業：大企業（従業員数500人以上）	0.13719	0.13245	1.03040		0.13719	0.12480	1.08370	
職業：官公庁	0.02204	0.02845	0.78006		0.02204	0.02390	0.92383	
職業：自営業・会社役員	0.03361	0.02845	1.17550		0.03361	0.03319	1.01220	
職業：非正規雇用	0.35041	0.34433	1.00850		0.35041	0.35059	0.99976	
職業：学生	0.02865	0.03004	0.95538		0.02865	0.02903	0.98743	
職業：無職	0.07273	0.06077	1.18200	+	0.07273	0.07123	1.01940	
職業：その他	0.06281	0.05166	1.20190	+	0.06281	0.06245	1.00540	
就業時間（対数化）	2.98550	3.06920	1.11600	*	2.98550	2.97650	1.02580	
本人の年収（対数化）	4.94520	4.85660	1.18330		4.94520	4.99630	1.09520	
両親と同居	0.63416	0.78516	1.37580	**	0.63416	0.63096	0.99636	
片親と同居	0.12727	0.11561	1.08670		0.12727	0.13014	0.98118	
両親と別居	0.23857	0.09900	2.03720	**	0.23857	0.23890	0.99905	
結婚意欲：絶対したい	0.34105	0.34433	0.99574		0.34105	0.34450	0.99519	
結婚意欲：どちらかといえばしたい	0.34711	0.36208	0.98146		0.34711	0.33605	1.01570	
結婚意欲：どちらともいえない	0.23140	0.21142	1.06710	+	0.23140	0.23744	0.98230	
結婚意欲：あまりしたくない	0.06336	0.06350	0.99834		0.06336	0.06706	0.94853	
結婚意欲：絶対したくない	0.01708	0.01866	0.91700		0.01708	0.01495	1.14030	

有意水準：0.1 + 0.05 \* 0.01 \*\*

表 8.7 第 5 回調査における脱落に対する共変量の調整前後の平均差の検定

第5回調査における脱落	調整前				調整後			
	平均 (脱落)	平均 (非脱 落)	平均比 (脱落/ 非脱落)	t検定	平均 (脱落)	平均 (非脱 落)	平均比 (脱落/ 非脱落)	t検定
学歴：中学卒	0.01782	0.01791	0.99505		0.01782	0.01347	1.31800	
学歴：高校卒	0.30105	0.28045	1.09270	**	0.30073	0.30917	0.98459	
学歴：専門学校卒	0.19370	0.18234	1.04780		0.19378	0.19549	0.99340	
学歴：短大・高专卒	0.21791	0.28567	0.87375	**	0.21801	0.22341	0.98261	
学歴：大学・大学院卒	0.26953	0.27363	0.99078		0.26965	0.25847	1.02750	
職業：中小企業（従業員数500人未満）	0.29328	0.32637	0.94296	**	0.29342	0.29926	0.98865	
職業：大企業（従業員数500人以上）	0.14390	0.12836	1.10130	+	0.14397	0.13571	1.05070	
職業：官公庁	0.02101	0.02960	0.71632	*	0.02102	0.02046	1.02680	
職業：自営業・会社役員	0.03244	0.02861	1.12960		0.03245	0.02958	1.09370	
職業：非正規雇用	0.35267	0.34254	1.01390		0.35283	0.35900	0.99228	
職業：学生	0.02969	0.02960	1.00320		0.02971	0.02896	1.02510	
職業：無職	0.06715	0.06269	1.06640		0.06719	0.06250	1.06950	
職業：その他	0.05985	0.05224	1.13660		0.05942	0.06452	0.92584	
就業時間（対数化）	3.02090	3.05770	1.10830		3.02230	3.00120	1.06590	
本人の年収（対数化）	4.99890	4.81920	1.23900	*	4.99590	4.98780	1.21760	
両親と同居	0.65692	0.78682	1.34390	**	0.65722	0.65407	0.99566	
片親と同居	0.12015	0.11841	1.01290		0.12020	0.12548	0.96371	
両親と別居	0.22248	0.09478	2.01670	**	0.22212	0.22045	1.00540	
結婚意欲：絶対したい	0.34354	0.34328	1.00060		0.34369	0.34865	0.99328	
結婚意欲：どちらかといえばしたい	0.34628	0.36393	0.97810		0.34644	0.34962	0.99575	
結婚意欲：どちらともいえない	0.23390	0.20821	1.08720	*	0.23355	0.22844	1.01560	
結婚意欲：あまりしたくない	0.06122	0.06468	0.95018		0.06124	0.05972	1.02390	
結婚意欲：絶対したくない	0.01508	0.01990	0.76143		0.01508	0.01357	1.11010	

有意水準： 0.1 + 0.05 \* 0.01 \*\*

## 8.4 おわりに

本章は、共変量の影響を1変数に縮約した傾向スコアという指標を用いることによって、脱落が初婚に与える因果効果を推定することを試みた。傾向スコアによって初婚に影響を与える共変量を調整したうえでの処置群における因果効果 TET は-0.15 から-0.16 といった推定値となっており、脱落が初婚に対する因果効果を測定することができた。このことは脱落の発生は初婚の発生に対して負の因果効果をもつことを示している。すなわち、脱落の発生は初婚の発生を過小にするという意味で、福田（2008）で示されている競合性を裏付ける結果となっている。

傾向スコアは因果効果を測定するのに適した指標であるが、対象となるカテゴリが3群以上ある場合、因果効果を求めるための母集団が2群ごとに異なってしまう点、マッチング・層別解析を行う際には標準誤差が正確に計算できず、周辺期待値の推定もできない点、マッチングに使用する傾向スコアは連続変数であるため、ペアを作る際には最近傍マッチングなどの距離を使う必要があり、そこに恣意性が生まれる点、マッチングを行う際にペア化されなかったデータが無駄になる点、共分散分析は線形を仮定しているが、傾向スコアは0から1をとるため仮定を満たさない点などいくつかの問題点が指摘されており（星野 2009）、それらを克服するための指標として傾向スコアの逆数の重み付け指標である IPW 推定量の開発が進んでいる。

また、傾向スコアや IPW 推定量を用いて、多変量解析におけるモデル推定の選択バイアスの除去を行っている坂本（2006）や、Web 調査と既存の調査のデータ融合を行うことによって、安価

なコストで正確な母集団を推定しようと試みている星野（2003, 2009）の研究など、脱落を含む欠測データ分析における対処の仕方に対しては多くの研究がある。本文で引用した直井（2007）は調査環境（調査員に関する情報）についての変数を第1段階のモデルに加えることで、脱落研究でも古典的な手法であるヘックマンの2段階推定の適用可能性についての知見を示しており、選択バイアス除去のために使用できる可能性があるため、本調査の今後の調査設計にも参考になると考えられる。

## 参考文献

Diggle, P. and Kenward, M. G., 1994. "Informative Drop-out in Longitudinal Data Analysis", *Applied Statistics*, Vol. 43, pp. 49-93.

Fitzgerald, J., P. Gottschalk, and R. Moffitt, 1998. "An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics", *The Journal of Human Resources*, 33, pp. 251-299.

Guo, Shenyang and Fraser W. Mark, 2010. "*Propensity score analysis*", SAGE Publications, Inc.

Hausman, J. A. and Wise, D. A., 1979. "Attrition Bias in Experimental and Panel Data: The Gary Income Maintenance Experiment", *Econometrica*, Vol. 47, pp. 679-694.

Rosenbaum, P. R. and Rubin, D. B., 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects", *Biometrika*, Vol. 70, pp. 41-55.

Rubin, 1985, "The Use of Propensity Scores in Applied Bayesian Inference", J. M. Bernardo, M. H. De Groot, D. V. Lindley and A. F. Smith (eds), *Bayesian Statistics*, Vol.2, pp. 463-472.

Sekhon S. Jasjeet, 2007, "Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching package for R", *Journal of Statistical Software*, Vol. VV, Issue II.

岩崎学, 2002. 『不完全データの統計解析』, エコノミスト社.

金子隆一, 2005. 「パネル調査における統計分析モデル」, 厚生労働科学研究費補助金『パネル調査(縦断調査)のデータマネジメント方策及び分析に関する総合的システムの開発研究』平成16年度報告書.

厚生労働省, 2009. 「第6回21世紀成年者縦断調査(国民の生活に関する継続調査)結果の概況」, 平成21年3月11日発表.

坂本和靖, 2006, 「サンプル脱落に関する分析—「消費生活に関するパネル調査」を用いた脱落の規定要因と推計バイアスの検証」, 『日本労働研究雑誌』, Vol. 48 (6) (551), pp.55-70.

直井道生, 2006, 「家計の住居移動行動とサンプル脱落」, KUMQRP ディスカッションペーパー, pp.1-11.

西野淑美, 金子隆一, 2005. 「21世紀出生児縦断調査における居住地移動サンプルの考察」, 厚生

労働科学研究費補助金『パネル調査（縦断調査）のデータマネジメント方策及び分析に関する総合的システムの開発研究』平成16年度報告書。

西野淑美, 2006. 「21世紀出生児縦断調査における脱落・居住地移動・復活サンプルの分析」, 厚生労働科学研究費補助金『パネル調査（縦断調査）のデータマネジメント方策及び分析に関する総合的システムの開発研究』平成17年度報告書。

西野淑美, 2007. 「第1回～第4回21世紀出生児縦断調査の脱落・移動の動向」, 厚生労働科学研究費補助金統『パネル調査（縦断調査）のデータマネジメント方策及び分析に関する総合的システムの開発研究』平成18年度報告書。西野淑美, 2008. 「第1回～第5回21世紀出生児縦断調査の脱落・移動の動向」, 厚生労働科学研究費補助金統『パネル調査（縦断調査）のデータマネジメント方策及び分析に関する総合的システムの開発研究』平成19年度報告書。

福田節也, 金子隆一, 2005. 「21世紀出生児縦断調査における脱落要因の分析」, 厚生労働科学研究費補助金統『パネル調査（縦断調査）のデータマネジメント方策及び分析に関する総合的システムの開発研究』平成16年度報告書。

福田節也, 2008. 「『21世紀成年者縦断調査』を用いた分析—初婚の要因分析：ネステッド・ロジットモデルによる初婚と脱落の競合ハザード分析」, 厚生労働科学研究費補助金統『パネル調査（縦断調査）のデータマネジメント方策及び分析に関する総合的システムの開発研究』平成19年度報告書。

星野崇宏, 2003. 「調査データに対する傾向スコアの適用」, 『品質』, Vol. 33, No.3.

星野崇宏, 繁杵算男, 2004. 「傾向スコア解析法による因果効果の推定と調査データの調整について」, 『行動計量学』, 第31巻, 第1号, pp.43-61.

星野崇宏, 前田忠彦, 2006. 「傾向スコアを用いた補正法の有意抽出による標本調査への応用と共変量の選択法の提案」, 『統計数理』, 第54巻, 第1号, pp.191-206.

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』, 岩波書店。

守泉理恵, 釜野さおり, 2009. 「21世紀成年者縦断調査（第1回～第5回）における女性票の脱落者・継続回答者の特性に関する分析」, 厚生労働科学研究費補助金統『パネル調査（縦断調査）のデータマネジメント方策及び分析に関する総合的システムの開発研究』平成20年度報告書。

## 索引

A	
ATC .....	150
ATE .....	150
ATT .....	147, 150
C	
Cox 比例ハザードモデル .....	79
D	
Durbin-Wu-Hausman 検定 .....	60
F	
F Test .....	29
G	
Gauss-Markov の定理 .....	7
graphviz .....	138
H	
Hausman Test .....	29
hetcor .....	131
I	
IIA の仮定 .....	102
IPW 推定量 .....	146
L	
lm .....	11
LM Test .....	29
M	
MAR .....	143
Match .....	150
MatchBalance .....	152
MCAR .....	143
N	
NMAR .....	143
P	
plm .....	22
polycor .....	130, 131
psych .....	130, 133
S	
sem .....	130
summary .....	11
SURF モデル .....	101
T	
TET .....	147
tetrachoric .....	133
W	
Wu-Hausman 検定 .....	60
あ	
一元配置誤差構成要素回帰モデル .....	14
一般化3段階最小2乗法 .....	55
一般化積率法推定 .....	40
イベントヒストリー分析 .....	87
ウィズイン推定 .....	18
オッズ .....	88
か	
カプラン・マイヤ推定量 .....	76
完全情報最小距離推定 .....	55
完全情報推定法 .....	55
競合するイベント .....	101
共分散構造分析 .....	119
共分散分析 .....	146
共変量 .....	87, 145
傾向スコア .....	143, 145
系列相関モデル .....	35
構造方程式モデル .....	119
固定効果モデル .....	15
さ	
最小距離推定法 .....	40
最小二乗法 .....	7
最尤法 .....	87
最尤法推定 .....	37
残差 .....	13
3段階最小2乗法 .....	55
時間依存性共変量 .....	87
時間独立共変量 .....	88
四分相関 .....	119, 126, 133
弱相関操作変数 .....	64
出生児縦断調査 .....	30, 45, 65, 81
状態依存モデル .....	36
人-期間別データ .....	91
スプライン関数 .....	93
制限情報推定法 .....	55
生存関数 .....	73
生存時間分析 .....	73
成年者縦断調査 .....	108, 144
SEM .....	119
線形回帰モデル .....	7
センサリング .....	88
操作変数法推定 .....	37
層別解析 .....	146
た	
ダイナミック・パネルデータ .....	33
多分系列相関 .....	119, 127