日本統計学会誌
第36巻，第1号，2006年9月
25頁～36頁

潜在クラスモデルの局所独立性を利用した共変量調整法

酒 折 文 武*，山 口 和 範**

Covariate Adjustment Methods with Latent Class Model

Fumitake Sakaori* and Kazunori Yamaguchi**

観察研究や非無作為化実験において，共変量からの影響が処理群と対照群の間で異なることが予想される．したがって，処理群と対照群における目的変数の値を比較する際には共変量の偏りを考慮に入れて調整をおこなう必要がある．

共変量を調整する方法は多く提案されているが，傾向スコアを用いる方法が近年注目されており，さまざまな分野で実際に分析がおこなわれている．傾向スコア法では，適切な調整ができるかどうかは傾向スコアの推定の良し悪しに強く依存する．つまり傾向スコアを推定する際の共変量と割り付けとの間のモデル設定が重要である．モデル設定が適切でない場合，傾向スコアの推定値は信頼性がなく，適切な共変量調整がおこなえない可能性がある．

本論文では，潜在クラスモデルの局所独立性を利用した共変量調整法を提案する．そして，欠測値補完の分析例を通して，提案手法が傾向スコア法よりもモデルのずれに関して頑健であることを示す．

In observation studies and non-randomized experiments, treatment and control groups may differ with respect to certain covariates. The propensity score, defined as the conditional probability that an observation is assigned to the treatment group given the covariates, can be used to balance the covariates in the treatment and control groups. The estimation of the propensity score, however, strongly depends on the model between covariates and treatment assignment. Therefore, improper model can not adjust the covariates.

In this paper, we propose a new covariates adjustment method based on latent class models. We then show the effectiveness and robustness of the method relative to the propensity score method by using some artificial examples.

Key Words and Phrases: latent class model, corariate adjustment, propensity score, missing data

1. はじめに

無作為化実験での処理群と対照群の比較の問題では，共変量の影響を考慮することなく分析をおこなうことができる．目的変数に影響を与えていると考えられる共変量は，どちらの割り付けに関しても均等に影響を与えていることになり，共変量は考慮に入れる必要が無いからである．ところが，観察研究や非無作為化実験においては，共変量による影響が群間で異なるため，共変量の偏りを考慮に入れて分析をおこなわなくてはならない．そこで共変量を調整する方法論が古くから研究されている．たとえばもっとも単純な方法として，共変量の一致する個

* 立教大学社会学部，〒171-8501 東京都豊島区西池袋3-34-1
** 立教大学経営学部，〒171-8501 東京都豊島区西池袋3-34-1

体ごとに層別もしくはマッチングをおこなう方法がある．しかしながら共変量の数が多い場合には共変量のパターンが完全に一致するような層やペアを作るのが難しい．

近年注目されているのが，共変量が完全に一致する組を作るのではなく，共変量からの影響が同一である組を作成し，組の中で比較をおこなう手法である．中でも，傾向スコア（Rosenbaum and Rubin, 1983, 2002）を用いた手法はもっとも広く利用されている．傾向スコアは，共変量を所与として各個体が処理群に割り付けられる確率である．傾向スコアを用いた層別解析やマッチング（Rosenbaum and Rubin, 1983, 1984），また傾向スコアによる重み付け法（Rosenbaum, 1987）などにより，共変量を調整して偏りなく処理効果を推定することが可能となる（星野・繁桝，2004参照）．ところが，傾向スコアを用いて適切な調整ができるかどうかは，共変量を用いた傾向スコアの推定の良し悪しに強く依存する．つまり共変量と割り付けとの間のモデルの設定が重要である．共変量と割り付けとの間のモデルを適切に定めないと適切な調整ができない可能性がある．

本論文では潜在クラスモデル（Lazarsfeld and Henry, 1968; Vermunt and Magidson, 2002; 渡辺，2001）の局所独立性を利用した共変量調整法を提案し，その適用例として処理効果の推定の問題および欠測値の補完について扱う．分析例による検討から，本手法は傾向スコア法を用いた場合よりもモデルのずれに関して頑健であることが示唆された．

本論文の構成は以下の通りである．まず，傾向スコアを用いた共変量調整法およびそれを用いた処理効果の推定と欠測値の補完について2節で概説する．つぎに3節では，潜在クラスモデルを用いた共変量調整法を提案し，同様に処理効果の推定の問題と欠測値の補完の問題を扱う．そして4節では，具体的な例題を用いて傾向スコア法との比較をおこなう．最後に5節でこれらを受けて考察する．

2. 傾向スコアによる調整

本節では，傾向スコアを用いた共変量調整の方法と処理効果の推定の問題および欠測値の補完の問題への適用について概説する．

いまzを処理群と対照群への割り付け変数とし，$z=1$のとき処理群，$z=0$のとき対照群であるとする．目的変数を$\boldsymbol{y}=(y_1, y_0)$とし，$z=1$のとき$y_1$のみ，$z=0$のとき$y_0$のみが観測されるとする．目的変数$\boldsymbol{y}$に影響を与えていると考えられる共変量を$\boldsymbol{x}=(x_1, x_2,..., x_p)$とする．

目的変数y_1とy_0とを比較する際，共変量$\boldsymbol{x}$からの影響に偏りがあることが考えられる．そこで従来用いられてきた手法として共変量パターンによる層別がある．共変量の値がすべて一致するような層別をして層内で目的変数の値を比較するのである．これにより，目的変数への共変量の影響を無視して比較することができる．とはいえ一般に共変量の次元は小さくなく，すべての共変量が一致するような層別は難しい．そこでRosenbaum and Rubin（1983）は傾向スコアを用いた層別を提唱している．

傾向スコアは，あたえられた共変量$\boldsymbol{x}$に対して$z=1$となる確率

$$p(\boldsymbol{x})=\Pr(z=1|\boldsymbol{x}) \qquad (1)$$

である．この傾向スコア$p(\boldsymbol{x})$は，それを与えたもとで$\boldsymbol{x}$とzが独立になるという性質がある．このことを，Dawid（1979）の記法に倣って

$$\boldsymbol{x} \perp\!\!\!\perp z|p(\boldsymbol{x}) \qquad (2)$$

と表現することとする．

いま，共変量を与えた上での強い意味での無視可能性，すなわち共変量を所与としたとき目

的変数と割り付け変数とが独立であること

$$(y_1, y_0) \perp\!\!\!\perp z|\boldsymbol{x}, \quad 0<\Pr(z=0|\boldsymbol{x})<1$$

を仮定する．この仮定は，共変量の値が同じ個体ならば，どちらの群に割り付けられたかを問わず (y_1, y_0) の分布は同じ，すなわち y_1 と y_0 を比較可能であることを意味している．強い意味での無視可能性を仮定すると，傾向スコア $p(\boldsymbol{x})$ を与えたもとでの (y_1, y_0) と z の独立性，すなわち

$$(y_1, y_0) \perp\!\!\!\perp z|p(\boldsymbol{x}) \tag{3}$$

を証明することができる（Rosenbaum and Rubin, 1983）．このことより，傾向スコアが等しい個体の間では共変量 $\boldsymbol{x}$ の影響を無視して目的変数 y の比較をおこなってよいことがわかる．

真の傾向スコアの値は実際にはわからないため，何らかの方法で推定する必要がある．そして推定された傾向スコアの値に応じて何層かに層別をし，層内では傾向スコアの値は等しい，つまり共変量の影響を無視できると考えるのである．適切な層の数や層別の方法を定めることは難しいが，5層以上のなるべく同じサイズに分けるのが望ましいといわれている．

重要なことは傾向スコアの推定である．共変量と割り付けとの関係を適切にモデル化しないと，傾向スコアの推定値が真の値とかけ離れてしまい，適切な共変量調整ができないことが予想される．一般に用いられるのは，共変量を説明変数，割り付け変数を目的変数としたロジスティック回帰モデル

$$\log\frac{p(\boldsymbol{x})}{1-p(\boldsymbol{x})}=\boldsymbol{\beta}'\boldsymbol{x} \tag{4}$$

である．他にも一般化ブースト回帰モデルを用いた方法（McCaffrey 他，2004）などいくつかの方法が提案されている．

しかし，モデルの妥当性の問題，モデルに含めるべき共変量および共変量間の交互作用効果の選別の問題がある．一般にモデル選択で用いられている情報量規準や修正適合度指標などは，あくまで予測モデルとしてのモデルの信頼性を測るものであるため，共変量調整を目的とした状況では妥当な規準であるとは限らない．したがって，単純な線形性を仮定したロジスティックモデルの妥当性が成立しない場合など，真のモデルがわからない限り傾向スコアの推定に信頼性は無く，適切な共変量調整でない可能性もある．

2.1 処理効果の推定

傾向スコアによる共変量調整を応用して，非無作為割付の場合に処理効果を推定する問題を考える．

処理効果は，処理群と対照群における目的変数 y の差である．各個体に関する処理効果 y_1-y_0 の期待値

$$E[y_1-y_0]=E[y_1]-E[y_0] \tag{5}$$

の推定が分析の目的である．しかしながら，同一個体に関しては y_1, y_0 のどちらかのみ観測されるため，実際に推定することができるのは

$$E[y_1|z=1]-E[y_0|z=0] \tag{6}$$

である．もし処理への割り付けが無作為ならば，割り付けと共変量とは独立であり，目的変数への共変量の影響は割り付けに関わらず等しいと考えられる．したがって

$$E[y_1|z=1]=E[y_1|z=0], \tag{7}$$
$$E[y_0|z=1]=E[y_0|z=0], \tag{8}$$

が成り立つ．このことより，（6）式の推定量

$$\Delta=\frac{1}{n_1}\sum_{i=1}^{n_1}y_{1i}-\frac{1}{n_0}\sum_{i=1}^{n_0}y_{0i} \tag{9}$$
$$=\frac{\sum_{i=1}^{n}z_iy_{1i}}{\sum_{i=1}^{n}z_i}-\frac{\sum_{i=1}^{n}(1-z_i)y_{0i}}{\sum_{i=1}^{n}(1-z_i)} \tag{10}$$

により処理効果を推定することができる．

しかし割り付けが無作為でない場合には共変量の影響を無視することができないため，（6）式で処理効果を推定することはできない．このような状況で用いられる方法のひとつに，傾向スコアに基づいた共変量調整がある．

まず（4）式のロジスティック回帰モデル等により傾向スコア $p(\boldsymbol{x})$ を推定する．そしてその値により個体をいくつかの層にわける．このとき，この層内では共変量の値によらず y の値を比較することができる．すなわち各層で $E(y_1-y_0)$ を評価することができる．したがって，第 s 層のデータを $\boldsymbol{x}_i^{(s)}$, $y_i^{(s)}=(y_{1i}^{(s)}, y_{0i}^{(s)})$, $z_i^{(s)}$，各層の大きさを $n^{(s)}$ とすると，各層内での処理効果 $\Delta^{(s)}$ は

$$\Delta^{(s)}=\frac{\sum_{i=1}^{n^{(s)}}z_i^{(s)}y_{1i}^{(s)}}{\sum_{i=1}^{n^{(s)}}z_i^{(s)}}-\frac{\sum_{i=1}^{n^{(s)}}(1-z_i^{(s)})y_{0i}^{(s)}}{\sum_{i=1}^{n^{(s)}}(1-z_i^{(s)})} \tag{11}$$

で推定される．このことより，全体の処理効果 Δ_P は，層の大きさで重み付けをして

$$\Delta_P=\frac{1}{n}\sum_{s=1}^{S}n^{(s)}\Delta^{(s)}$$
$$=\frac{1}{n}\sum_{s=1}^{S}n^{(s)}\left\{\frac{\sum_{i=1}^{n^{(s)}}z_i^{(s)}y_{1i}^{(s)}}{\sum_{i=1}^{n^{(s)}}z_i^{(s)}}-\frac{\sum_{i=1}^{n^{(s)}}(1-z_i^{(s)})y_{0i}^{(s)}}{\sum_{i=1}^{n^{(s)}}(1-z_i^{(s)})}\right\} \tag{12}$$

により推定することができる．

2.2 欠測値の補完

共変量調整法は，欠測値を含む多変量データにおける欠測データの補完に応用することができる．本節は前節までと設定が異なるため記号を再定義する．いま，欠測はひとつの変数にしかおきていないとし，その変数を y とする．また，z を欠測指標変数とし，$z=1$ のとき観測，$z=0$ のとき欠測であるとする．すべての個体において観測されている変数を x とする．

欠測発生のメカニズムが「欠測が完全にランダム（Missing Completely At Random）」であれば，欠測部分 y を考慮に入れずに分析をおこなうことができる．たとえば，欠測データをケースごと除外して分析をおこなったり，欠測部分の推定値として全体の平均値などを入れる方法がある．ただし，前者の方法では標本の大きさが小さくなってしまい，後者の方法では分散を過小評価してしまうことには注意が必要である．

一方，欠測が完全にランダムではない場合，上記のような方法では平均が大きく変わり分析結果に重大な誤りが生ずる可能性がある．しかし「欠測がランダム（Missing At Random，以下 MAR）」であれば，適切な分析手法を用いることで偏りのない結果を得ることが可能となる．欠測がランダムであるとは，ある部分が欠測するかどうかは観測部分のみに依存し，欠測している部分には無関係ということである．このとき，x が与えられたもとで y と z は独立，すな

わち

$$y \perp\!\!\!\perp z | \boldsymbol{x}$$

となる．このことより，$\boldsymbol{x}$ が同じ個体ならば，y の値によって観測されるか欠測するかの確率は変わらない．MAR の仮定は強いように感じるが，欠測のある変数 y と関連のある共変量をできるだけ多く含めておくことでそれほど不自然な仮定ではなくなる．

このとき，欠測したケースをどのように処理するのかが問題となる．ケースごと除外する方法や平均値で補完する方法など様々な手法が提案されているが，なかでも傾向スコアを用いた Hot Deck 法による補完は，適切な層別が可能であれば計算結果に偏りを生じさせないため有効である．

まず，完全データ $\boldsymbol{x}$ を説明変数，欠測指標変数 z を目的変数としてロジスティック回帰モデル等を用い，傾向スコアを推定する．つぎに推定された傾向スコアの値によっていくつかの層にわける．そしてその層ごとに欠測値の補完をおこなえばよい．層ごとに補完をおこなう方法としては，単一値代入法が最も簡単である．たとえば層 s における欠測値は，$z=1$ に対する y の値，つまり観測されている部分の平均値をもって欠測部分の推定値とする方法である．たとえば，欠測があるひとつの変数のみに起こっているとすると，傾向スコア法により第 s 層に層別された欠測値を

$$\hat{y}^{(s)} = \frac{\sum_{i=1}^{n^{(s)}} z_i^{(s)} y_{1i}^{(s)}}{\sum_{i=1}^{n^{(s)}} z_i^{(s)}}$$

で補完すればよい．より複雑な方法としては，層内でのリサンプリングによる近似ベイズブートストラップ法を用いたノンパラメトリックな多重代入法（岩崎，2002 参照）などが考えられる．

なお，ここで述べた層別は共変量調整における層別と本質的には同等であることに注意されたい．なぜなら MAR の仮定は 2.1 節で説明した“強い意味での無視可能性”と同義であり，したがって，MAR の仮定で観測部分のデータから層別することは，“強い意味での無視可能性”の仮定で共変量の値から層別をおこなうことと同義であることになる．

3. 潜在クラスモデルによる調整

本節では，傾向スコア法による層別と同じように共変量のバランスがとれている層（クラス）を構成する方法を提案する．傾向スコアを用いた共変量調整では，傾向スコアを与えたもとで $\boldsymbol{x}$ と z が独立となることを用い，層内での (y_1, y_0) と z の独立性を導いた（図 1 参照）．した

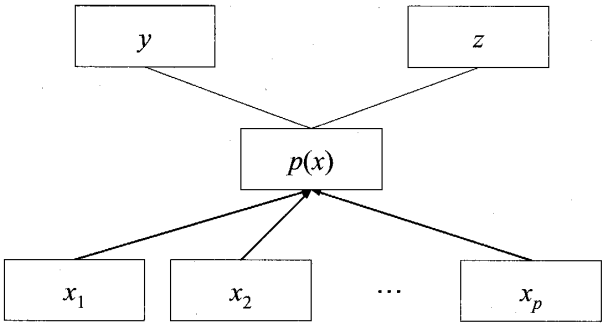

図 1　傾向スコア法のパス図

がって，傾向スコアに限らず $\boldsymbol{x}$ と z が独立となるような層別をおこなうことにより層内での (y_1, y_0) と z の独立性が示され，y_1 と y_0 の比較をおこなうことが可能となる．

所与のクラス数を C とし，クラス内での z と各 x_i, $i=1, 2,..., p$ との独立性（局所独立性）を仮定したモデル

$$P(z, \boldsymbol{x})=\sum_{c=1}^{C}P(c)P(z|c)P(\boldsymbol{x}|c) \tag{13}$$

を考える．ここで P は確率関数もしくは確率密度関数とする．条件付き確率（分布）は，$P(z|c)$ にはロジットモデル

$$P(z|c)=\frac{\exp(\alpha_z+\beta_{cz})}{\exp(\beta_0+\beta_{c0})+\exp(\beta_1+\beta_{c1})},$$

$P(\boldsymbol{x}|c)$ には $\boldsymbol{x}$ がカテゴリカルのとき $(x_1, x_2,..., x_p)$ のパターンによる多項ロジットモデル，$\boldsymbol{x}$ が量的のとき多変量正規分布

$$P(\boldsymbol{x}|c)=(2\pi)^{-p/2}|\boldsymbol{\Sigma}_c|^{-1/2}\exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_c)'\boldsymbol{\Sigma}_c^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_c)\right\}$$

を仮定することが考えられる．

また，モデル（13）に対し，共変量 $x_1, x_2,..., x_p$ 間に独立性を仮定することも可能である．その場合，モデルは

$$P(z, \boldsymbol{x})=\sum_{c=1}^{C}P(c)P(z|c)\prod_{i=1}^{p}P(x_i|c) \tag{14}$$

となる．この場合も同様に $P(x_i|c)$ にはロジットモデルもしくは1変量の正規分布を考えればよい．ただし適切なクラス分けをおこなうためには，モデル（13）よりも多くのクラスが必要なことが予想される．

さらに，共変量 x_i を

$$P(z|\boldsymbol{x})=\sum_{c=1}^{C}P(c|\boldsymbol{x})P(z|c) \tag{15}$$

のように入れ込むモデルも考えられる．つまり，各クラスへの分類は共変量のみに依存し，共変量から割り付けへの影響はクラスを通じてのみということである．この場合も，$P(c|\boldsymbol{x})$ には多項ロジットモデルなどを考えればよい．

モデル（14）は二値変数 z とカテゴリカルあるいは量的変数 $\boldsymbol{x}$ を indicator とした Finite Mixture Model もしくは潜在クラスモデル（Lazarsfeld and Henry, 1968; Vermunt and Magidson, 2002; 渡辺，2001）に他ならない．また，モデル（13）はその局所独立性の制約を緩めたモデル（Hagenaars, 1988）である．モデル（15）は，外生変数として $\boldsymbol{x}$ をもち割り付け変数 z を indicator とした潜在クラスモデル（Clogg, 1981）である．

(13)(14)(15) のいずれのモデルにおいても，傾向スコアによる層別と同様にクラス内での z と x_i との独立性が成立する（図2，3参照）．したがって，割り付けごとの共変量の偏りは調整されていることになる．なお，各クラスへの分類は，ベイズの定理より帰属確率という形で表現される．すなわち，モデル (13)(14) においては

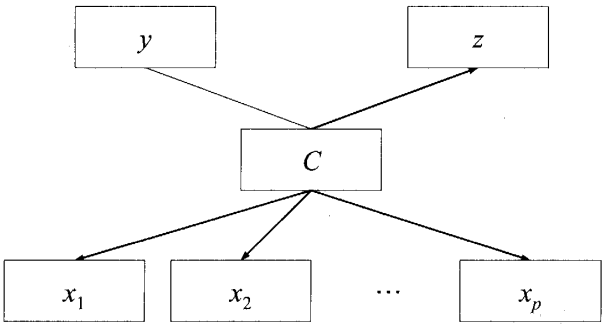

図2　潜在クラス法のパス図

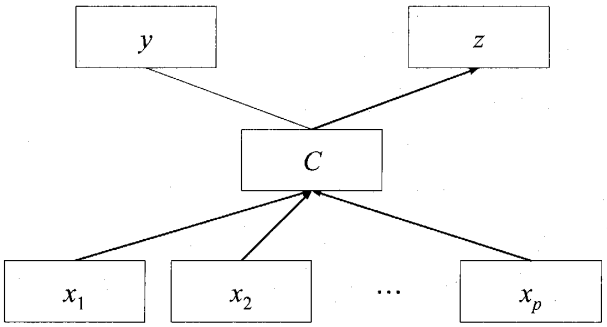

図3　外生変数を使用した潜在クラス法のパス図

$$P(c|\boldsymbol{x}, z)=\frac{P(c)P(z|c)P(\boldsymbol{x}|c)}{P(z, \boldsymbol{x})}, \tag{16}$$

モデル（15）においては

$$P(c|\boldsymbol{x}, z)=\frac{P(c|\boldsymbol{x})P(z|c)}{P(z|\boldsymbol{x})} \tag{17}$$

によって各クラスへの帰属確率を求めることができる．

なおこれらのモデルにおけるパラメータはEMアルゴリズムにより推定することができる．EMアルゴリズムの詳細はDempster, Laird and Rubin（1977），渡辺・山口（2000）を参照せよ．

3.1　処理効果の推定

分析の目的が処理効果の推定の場合には，各クラス内での目的変数の比較を行い，それをひとつに要約することで割り付けの効果を推定することになる．ただし傾向スコアを用いた場合と異なるのは，各個体がどこかひとつの層に属すのではなく，すべてのクラスへの帰属確率を持っているということである．

各クラスへの帰属確率の推定値をw_i^cと書くこととする．潜在クラスモデルにおいては，第cクラスのサイズを全データに関する帰属確率の和

$$n^{(c)}=\sum_{i=1}^{n} w_i^{(c)}$$

で考える．これにしたがうと，第cクラスにおける処理効果$\Delta_L^{(c)}$は

$$\Delta_L^{(c)}=\frac{\sum_{i=1}^n w_i^{(c)}z_i y_{1i}}{\sum_{i=1}^n w_i^{(c)}z_i}-\frac{\sum_{i=1}^n w_i^{(c)}(1-z_i)y_{0i}}{\sum_{i=1}^n w_i^{(c)}(1-z_i)} \tag{18}$$

で推定することができる．したがって全体での処理効果は，クラスの大きさで重み付けをして

$$\begin{aligned}\Delta_L&=\frac{1}{n}\sum_{c=1}^C n^{(c)}\Delta^{(c)}\\&=\frac{1}{n}\sum_{c=1}^C n^{(c)}\left\{\frac{\sum_{i=1}^n w_i^{(c)}z_i y_{1i}}{\sum_{i=1}^n w_i^{(c)}z_i}-\frac{\sum_{i=1}^n w_i^{(c)}(1-z_i)y_{0i}}{\sum_{i=1}^n w_i^{(c)}(1-z_i)}\right\}\end{aligned} \tag{19}$$

で推定できる．

3.2 欠測値の補完

傾向スコアの場合と同様，欠測がランダムである場合を考える．このとき，まず潜在クラスモデルにより各クラスへの帰属確率を推定する．つぎに，各クラス内において欠測値の補完をおこなう．単純な方法としては，帰属確率を重みとして各クラスにおいて観測されている目的変数の値の平均を算出し，その平均を欠測部分に代入する単一値代入法が考えられる．より複雑な方法としては，帰属確率を重みとした経験分布に基づいた近似ベイズブートストラップ法を使用することができる．いずれにしても，各クラスで欠測値の補完をおこなった後，帰属確率を重みとして全クラスに関して合計をとり欠測値を補完することとする．

たとえば，ひとつの変数のみに欠測が起きているとすると，単一値代入法による第cクラスでの欠測値は

$$\hat{y}^{(c)}=\frac{\sum_{i=1}^n w_i^{(c)}z_i y_i}{\sum_{i=1}^n w_i^{(c)}z_i} \tag{20}$$

となる．クラスの帰属確率で加重平均をとると，

$$\hat{y}_i=\sum_{c=1}^C w_i^{(c)}\hat{y}^{(c)} \tag{21}$$

が各欠測値に対する補完値ということになる．

4. 分析例

本節では，分析例として欠測値の補完の問題をとりあげる．欠測値の問題をとりあげる理由は，欠測の問題と処理効果の問題は共変量あるいは観測部分の調整という意味で本質的に同様であること，目的変数がひとつしか存在しない欠測の問題のほうがシンプルであること，完全データに対して人工的に欠測を与えることで補完値の精度を確認できること，などである．

分析用のデータとしてはFisherのアイリスデータを用い，乱数によって欠測を人工的に発生させることとした．品種は変数として使用せず，Petal Width（V_1），Petal Length（V_2），Sepal Width（V_3），Sepal Length（V_4）の4変数を使用した．V_1にのみ欠測を発生させ，V_2，V_3，V_4は完全データとした．欠測がランダムとなるよう，以下の3通りのメカニズムで欠測を発生させ，分析をおこなった．

Case 1：$V_2\geq 50$のとき，V_1は0.5の確率で欠測

Case 2：$V_2\geq 50$もしくは$V_2\leq 15$のとき，V_1は0.5の確率で欠測

Case 3：$V_2\geq 30$かつ$V_3<30$のとき，もしくは$V_2<30$かつ$V_4\geq 50$のとき，V_1は0.5の確率で欠測

いずれのケースもMARが成り立っているが，交互作用のない線形のロジスティック回帰モデルでは欠測の確率を正しく表現できない．Case 1はその中でも比較的線形に近く，Case 2は線形では適切に表現できず，Case 3は交互作用を含むため線形では表現できない．一様乱数によって欠測を人工的に発生させた結果，全ケース（$n=150$）中，Case 1では26ケース，Case 1では40ケース，Case 3では37ケースにおいてV_1が欠測した．この3つのデータを用いて分析をおこなっていく．

共変量調整の方法としては，共変量の主効果のみを考慮した線形のロジスティック回帰モデルに基づいた傾向スコアによる層別と，潜在クラスモデルのモデル（14）による層別をおこなった．傾向スコア法では5層にわけ，潜在クラス法ではどのCaseにおいてもBICで最適と判断された4クラスモデルを用いた．

表1は，欠測を補完した値と真の値との誤差平方和を記したものである．線形に近いCase 1の場合，若干傾向スコアによる方法のほうが誤差が小さいもののその差は小さく，またどちらの手法でも真の値に近い値が補完できていることがわかる．Case 2とCase 3の場合，線形なロジスティック回帰では表現できないため，傾向スコアによる方法では誤差が非常に大きい，つまり補完した値が真の値から非常に遠い．それに対して潜在クラスモデルを用いた場合には，どのケースでもほとんど誤差の値は変わらず，補完した値の精度が高いといえる．これらのことは，真の値と補完値との散布図である図4，図5，図6をみても確認することができる．なおこれらの散布図においては，位置が重なっている個体が存在することに注意されたい．

表2は，真の平均値（全体の平均値），観測部分のみの平均値（欠測のある個体ごと除外して計算したもの），および両手法で補完した場合の標本平均の値を求めたものである．どのケースでも，観測部分のみの平均は真の平均と差がある．そしてそれぞれの手法で補完をしたあとの平均をみると，真の平均に近づいていることがわかる．とくに，傾向スコアを用いた場合は真の値にかなり近い値になっていることが確認できる．

これらのことより，傾向スコアを用いた場合個々の補正値は真の値から遠く離れていても，全体としては真の平均に近い値になることがわかる．しかし，個々の値が真の値から遠く離れているということは，他の変数との関係性は正しく補正できていないと考えられ，他の変数との相関係数は真のものと大きく離れていることが想定される．そこで表3に，それぞれのケースでのV_1と他の変数との相関係数をまとめた．

Case 1においては，不完全データを完全に除外しても真の相関係数とそれほど値が変わらず，傾向スコアや潜在クラスを用いて補正をした結果もほとんど変わらない．しかしCase 2

表1 真の値からの誤差平方和

	Case 1	Case 2	Case 3
傾向スコア補正	292.028	2455.262	2257.252
潜在クラスモデル	457.140	465.066	368.116

表2 真の平均と補完後の平均

	Case 1	Case 2	Case 3
真の値	11.9	11.9	11.9
除外	10.2	11.2	12.4
傾向スコア補正	11.6	11.8	11.8
潜在クラスモデル	11.5	11.6	11.9

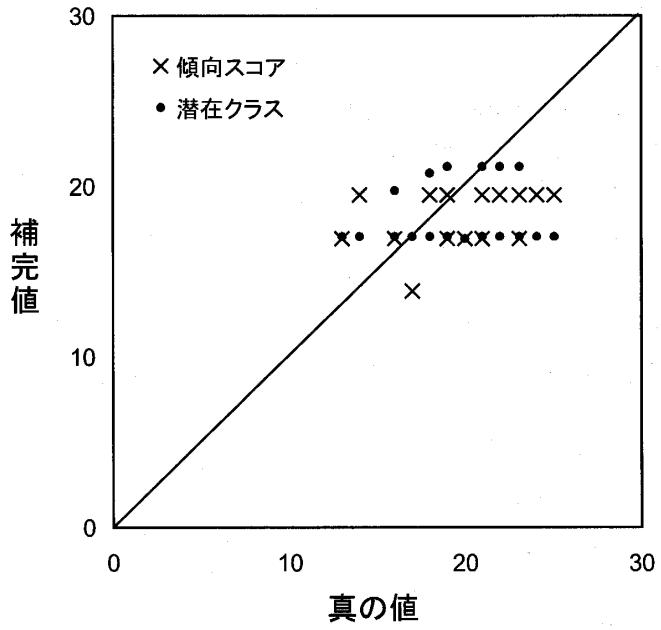

図4 Case 1における欠測値の真の値と補完値の散布図（欠測した個体数：26）

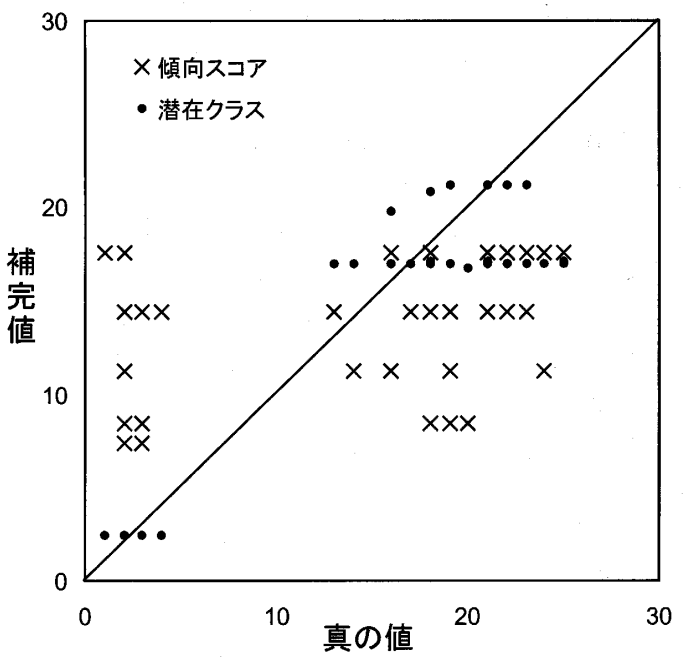

図5 Case 2における欠測値の真の値と補完値の散布図（欠測した個体数：40）

と3に関しては，大きく結果が異なってくる．除外した場合や潜在クラスを用いた場合にはほとんど相関係数は変わらないが，傾向スコアを用いた場合にはその大きさを過小評価することになる．これは，個々の補完の精度がわるいことを反映している．

なお，このデータはアイリスの3品種の混合集団である．したがって多変量正規分布の仮定は妥当ではなく，単に多変量正規分布を仮定したEMアルゴリズムによる最尤推定は適切ではない．そういった点からも本手法の有効性がわかる．

5. 議　　論

本論文では，潜在クラスモデルを用いて共変量を調整する手法を提案した．そして欠測の補

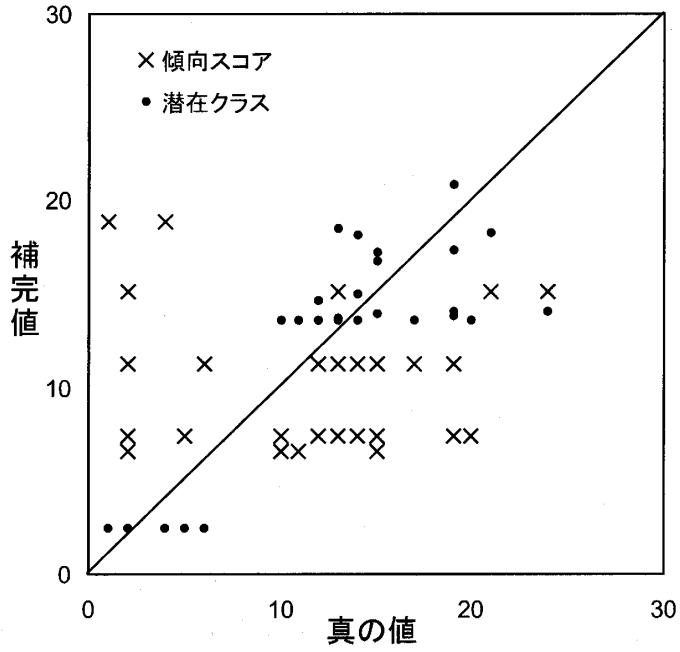

図6　Case 3 における欠測値の真の値と補完値の散布図（欠測した個体数：37）

表3　相関係数

	V_2	V_3	V_4
Case 1			
真の値	0.96	−0.37	0.81
除外	0.96	−0.43	0.82
傾向スコア補正	0.97	−0.40	0.82
潜在クラスモデル	0.97	−0.40	0.84
Case 2			
真の値	0.96	−0.37	0.81
除外	0.96	−0.36	0.82
傾向スコア補正	0.82	−0.19	0.76
潜在クラスモデル	0.97	−0.39	0.84
Case 3			
全体	0.96	−0.37	0.81
除外	0.96	−0.34	0.82
傾向スコア補正	0.82	−0.11	0.77
潜在クラスモデル	0.96	−0.37	0.84

完の例を用いて，傾向スコア法とどのように結果が異なるかを検討した．

傾向スコア法を用いた場合，傾向スコアを推定するときに用いたモデルが適切であればよい補正をおこなうことができる．しかしそうでない場合には，補正後の平均値すなわち一次のモーメントこそ真の値に近くなるが，個々の欠測値の補完はあてにならず，相関係数など高次のモーメントは真の値と大きく異なってしまう可能性がある．これは欠測値の補完の問題に限らず，処理効果の推定などの問題でも同様である．したがって，傾向スコア法を用いる場合には傾向スコアの推定のためのモデルを適切に選択することが重要となる．そのためには，データ収集の事前にモデルを探索しておくことが必要である．

それに対して潜在クラスモデルを用いた場合には，共変量と割り付け（欠測）の間の複雑な

構造にも対応可能である．したがって，共変量調整のためのモデルが明らかな場合には傾向スコア法が適切であるが，そうでない場合には潜在クラスモデルによる方法のほうが適切であるといえる．

なお残された問題として，クラス数の選択の問題がある．ロジスティック回帰モデルの場合と同様，潜在クラスモデルの当てはまりのよさの指標であるAIC，BICなどの情報量規準は目的変数を考慮に入れずに共変量と割付変数とで求めたモデルの当てはまりのみに着目したものである．したがって，分析の目的である共変量の調整もしくは目的変数の補正という観点からこれらの規準による選択が適切であるとは限らない．むしろクラス数はある程度多いほうが共変量のバランスのとれたクラス分けができる可能性があるが，多すぎると各クラスのサイズが小さくなり推定の精度が低くなることが予想される．しかし逆にクラス数が少なすぎると，局所独立性の仮定に無理が生じ，適切な共変量調整ができないであろう．実際，分析例の各ケースにおいてクラス数を変更してみると，クラス数がある程度多い場合のほうが分析結果がよいことが多かった．ただし多すぎる場合には，サイズのごく小さいクラスができてしまうこともあった．どのようにクラス数を決めればよいか，十分な検証が必要であろう．また提案した潜在クラスモデルのうちどのモデルを適用するのが妥当であるかについても検討の余地が残されている．

参考文献

[1] Clogg, C. C. (1981). New developments in latent structure analysis. D. J. Jackson and E. F. Borgotta (eds.), *Factor Analysis and Measurement in Sociological Research*, 215-246. Beverly Hills: Sage Publications.

[2] Dawid, A. P. (1979). Conditional independence in statistical theory (with discussion). *Journal of the Royal Statistical Society*, B, **41**, 1-31.

[3] Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society*, B, **39**, 1-38.

[4] Hagenaars, J. A. (1988). Latent structure models with direct effects between indicators: local dependence models. *Sociological Methods and Research*, **16**, 379-405.

[5] 星野崇宏，繁桝算男．(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，**31**, 43-61.

[6] 岩崎学．(2002)．不完全データの統計解析．エコノミスト社.

[7] Lazarsfeld, P. F., and Henry, N. W. (1968). *Latent Structure Analysis*. Boston: Houghton Mill.

[8] McCaffrey, D. F., Ridgeway, G., and Morral, A. R. (2004). propensity score Estimation with boosted regression for evaluating causal effects in observational studies. *Psychological Methods*, **9**, 403-425.

[9] Rosenbaum, P. R. (1987). Model-Based Direct Adjustment. *Journal of the American Statistical Association*, **82**, 387-394.

[10] Rosenbaum, P. R. (2002). *Observational Studies, 2nd Ed.*, New York: Springer-Verlag.

[11] Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, **70**, 41-55.

[12] Rosenbaum, P. R. and Rubin, D. B. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association*, **79**, 516-524.

[13] Vermunt, J. K., and Magidson, J. (2002). Latent class cluster analysis. J. A. Hagenaars and A. L. McCutcheon (eds.), *Applied Latent Class Analysis*, 89-106. Cambridge: Cambridge University Press.

[14] 渡辺美智子．(2001)．因果関係と構造を把握するための統計手法—潜在クラス分析法—，岡太彬訓・木島正明・守口剛（編著），マーケティングの数理モデル，73-115.

[15] 渡辺美智子，山口和範．(2000)．*EM* アルゴリズムと不完全データの諸問題，多賀出版.