

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も、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・繁柁 (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。

3. 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という (星野 2003)。

傾向スコアの数理モデルを星野・繁柁 (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、以下のように推定される。

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

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、以下のようなになる。

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

$\hat{E}(y_1)$ 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。

4. 脱落発生の中婚発生への因果効果

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 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。

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

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)。

本稿に即していうと、脱落の発生を欠測インディケータとした場合の、中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 をそのまま測定することができない。そこで、前節で説明した中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 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。

表 1 には、本モデルにおける記述統計量を示している。中婚は、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

¹ マッチングは R 2.10.1 “Matching” package 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。

(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2に示している。

表1 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 ² 乗	.034		.031		.038		.039	
Nagelkerke R ² 乗	.063		.047		.054		.054	

有意水準：0.1 + 0.05 * 0.01 **

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、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ということがある。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）の差の検定を一樣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。

表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4から表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。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うえでの脱落の初婚に対する因果効果が推定されていることを担保している。

表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	0.0085571	0.0089498	0.0088936	0.0096965
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・高専卒	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	0.56925	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・高専卒	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	0.89867		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	0.26045	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	0.26567	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 **

おわりに

本稿は、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。

IPW 推定量を用いた因果効果の測定や、多変量解析モデルへの適用などは今後の課題としたい。また、傾向スコアや IPW 推定量に限らず、脱落事象を初めとした欠測事象を扱うデータ分析では、2段階推定の手法が多くみられる通り、脱落事象の多い成年者調査においては、今後手法的に検討する意義が多くあると考える。

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.
- 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, 繁栞算男, 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.

【成年者調査】

10 縦断調査マイクロシミュレーション分析システムの設計・開発 － 標本個体クラスモジュールならびに日付処理クラスモジュールの開発 －

金子 隆一

1. はじめに

パネル調査(縦断調査)データの有効な分析・活用法の一つとして、マイクロシミュレーション分析がある。マイクロシミュレーションとは、各種属性を持った個人の集団をコンピュータ上に構成して、おのおのの行動や状態変化を発生させることにより、集団の変化を再現するシミュレーション技法である。対象集団の将来予測、行政制度・施策の効果の予見をはじめ、行動メカニズムの解明や統計手法の精度評価など、幅広く応用される。一方、パネル調査は、抽出された標本内の同一対象(個人、世帯)を追跡しながら継続的に調査し、対象者の変化とその要因を記録して行くものであり、その枠組みやデータ構造はマイクロシミュレーションにきわめて近く、両者はきわめて親和性が高い。実際、諸外国においては、社会政策、税制等の制度・施策の評価や検討のためにパネル調査に基づいたマイクロシミュレーション分析が行われている。

本事業では、21世紀縦断調査の結果を元に、その分析対象となる結婚、出生、就業などの事象の発生メカニズム、決定要因の解明や、介入(たとえば制度・施策の実施)の効果の評価・予測を行う際に有力な分析手段となるマイクロシミュレーション分析を行うこととしている。また、パネル調査の統計分析上の弱点ともいえる標本脱落や回答不詳・不整合の影響を評価する方法ともなりうるので、既存の統計モデルの検証に用いることで、より信頼性の高い分析結果を提供できると考えられる。したがって、それら既存の統計モデルによる分析と合わせてマイクロシミュレーション分析を行うことによって、縦断調査データの活用範囲を広げるとともに、提供する情報の信頼性向上に資することが期待できる。

本研究では、21世紀縦断調査データを活用して今後継続的なマイクロシミュレーション分析が行えるよう、その基礎としてエージェント型(agent-base)のマイクロシミュレーションモデルに必要な標本を生成するシステムを開発している。これはパネル調査データの管理情報を活用して、シミュレーション分析に必要な標本モデルを半自動的に生成するシステムであり、現行ではC++によるシミュレーションモデルを作成することができる。システムは、本事業で構築を行ったデータマネジメントシステムの一環として開発されており、統合的に扱うことができるものである。これまで、シミュレーションシステムに必要なクラスモジュールを順次開発して来たところであるが、本年度の事業では、本モデルシステムの中核に位置する、調査対象個人のモデル、すなわち標本個体クラスモジュールと、このクラスのプロパティとして重要な日付処理クラスモジュールの開発と調整を行った。

2. パネル調査とマイクロシミュレーション

パネル調査データの分析法としてのマイクロシミュレーションは、調査対象集団の変化の将来予測、行政制度・施策の効果の予見をはじめ、行動メカニズムの解明や統計手法の精度評価など、既存の統計分析に止まらない多くの応用と可能性を持っている。すなわち、パネル調査で捉えられた標本をシミュレーションモデルとして再現すれば、さまざまな仮想的条件や仮定の下での標本の変化を観察することが可能であり、それらを実際の変化と比較すれば、仮定の現実的な妥当性を評価することができる。

実際、諸外国においては、社会政策、税制等の制度・施策の評価や検討のためにパネル調査に基づいたマイクロシミュレーション分析が盛んに行われている。カナダでは早くから統計局においていくつかのモデルが開発され、長年にわたって政策シミュレーションに用いられている。そのうち SPSD/M と呼ばれるものは、さまざまな横断調査や行政情報を組み合わせて構築された標本データベースを基にしたシミュレーションモデルであり、主として税制や所得分析に用いられている。また、縦断型のモデルとしては、LifePaths と呼ばれるモデルがある。これは国民を代表する標本について、ライフコース全体をシミュレートする能力があり、個人や世帯を対象とした政策の評価や世代間公平性などの分析に用いられている。POHEM は、健康・疾病に関する縦断型のシミュレーションモデルである。さらに、汎用的なシミュレーションを構築するシステムとして、Modgen という言語が開発されている。これらはすべて統計局のインターネットサイト上に説明書と共に公開されている。アメリカ政府によって実施されているシミュレーション分析とともにこれら进行评估した論文集が見られる(Lewis and Michel (eds.) 1990)。アメリカ政府からはマイクロシミュレーションの実施に関する説明資料が公刊されている(Citro and Hanushek (eds.) 1991)。また、この他にも欧米各国(イギリス、ドイツ、オランダ、オーストリア、フィンランド、スウェーデン、デンマーク、ノルウェー、カナダなど)の社会政策、税制等をテーマにしたマイクロシミュレーションの実施に関する個別論文を含んだ論文集が見られる(Harding 1996)。

縦断型マイクロシミュレーションは、21世紀縦断調査についても、その主要なテーマである結婚・出生・子育てなどの発生メカニズムと決定要因の解明や、制度・施策効果の評価を行う有力な手法となるほか、脱落をはじめとするパネル調査特有の統計分析上の困難に対して、さまざまな条件下におけるそれら統計手法の妥当性や精度を検証する有効な手段を与えると考えられる。

3. 縦断調査用マイクロシミュレーション分析システムの開発

マイクロシミュレーションとは、各種属性を持った個人の集団をコンピュータ上に構成して、おのおのの行動や状態変化を発生させることにより、集団の変化を再現するシミュレーション手法である。とくに縦断型マイクロシミュレーション longitudinal microsimulation と呼ばれるものは、個人の経時的変化を模擬するもので、パネル調査データとの親和性が高く、対象集団の変化の将来予測、行政制度・施策の効果の予見をはじめ、行動

メカニズムの解明や統計手法の精度評価など、既存の統計分析に止まらない多くの応用と可能性を持っている。パネル調査で捉えられた標本をシミュレーションモデルとして再現すれば、さまざまな仮想的条件や仮定の下での標本の変化を観察することが可能であり、それらを実際の変化と比較すれば、仮定やモデルの妥当性を評価することができる。

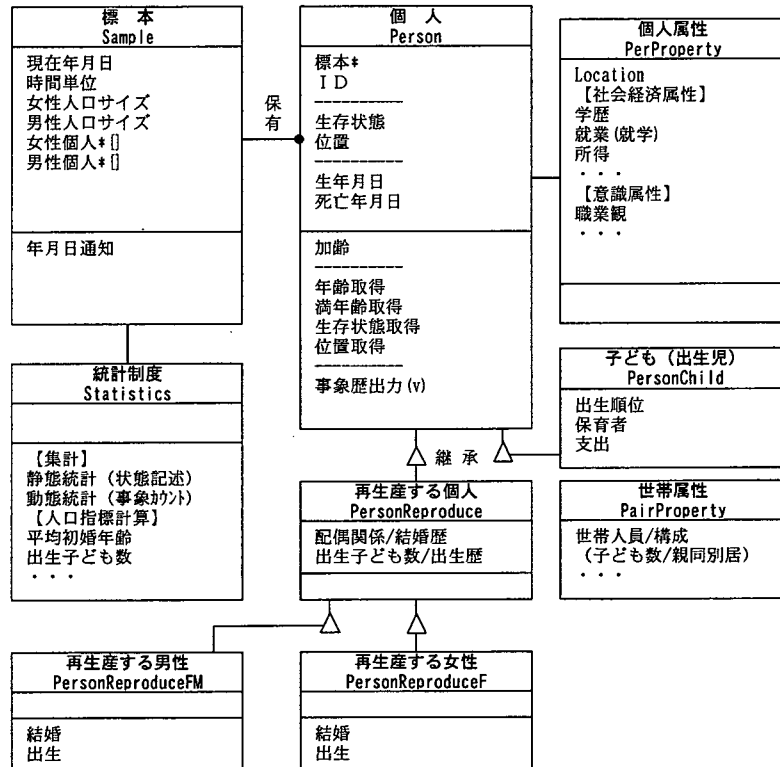
縦断型マイクロシミュレーションは、21世紀縦断調査についても、その主要なテーマである結婚・出生・子育てなどの発生メカニズムと決定要因の解明や、制度・施策効果の評価を行う有力な手法となるほか、脱落をはじめとするパネル調査特有の統計分析上の困難に対して、さまざまな条件下におけるそれら統計手法の妥当性や精度を検証する有効な手段を与えると考えられる。

本研究では、21世紀縦断調査データを活用して今後継続的なマイクロシミュレーション分析が行えるよう、その基礎としてエージェント型(agent-base)のマイクロシミュレーションモデルに必要な標本を生成するシステムを開発した。これはパネル調査データの管理情報を活用して、シミュレーション分析に必要な標本モデルを半自動的に生成するシステムであり、現行ではC++によるシミュレーションモデルを作成することができる。システムは、本事業で構築を行ったデータマネジメントシステムの一環として開発されており、統合的に扱うことができるものである。

(1) マイクロシミュレーションにおける標本モデル

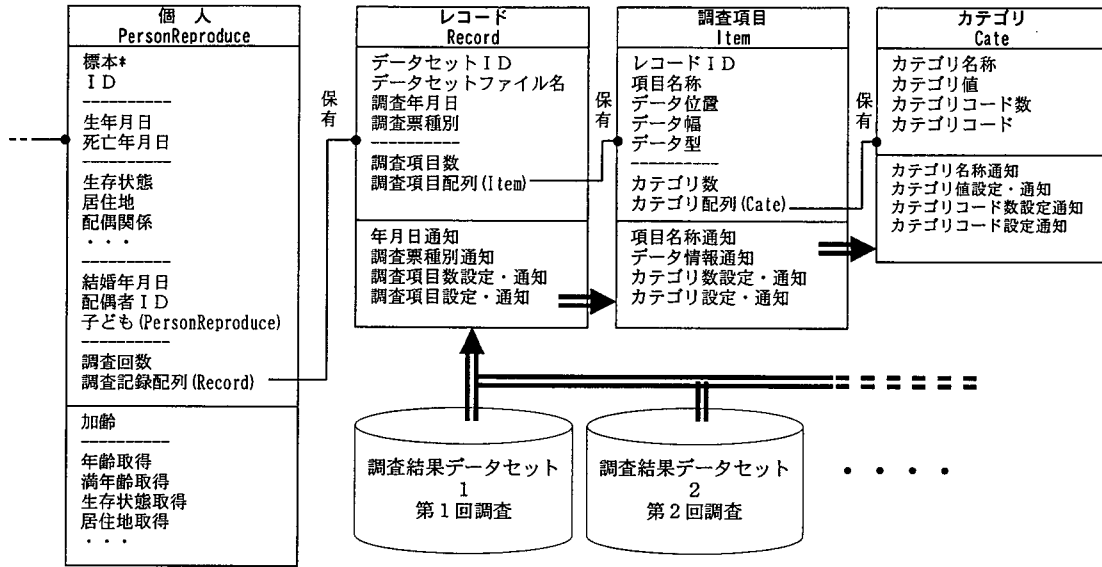
ここで想定する縦断型マイクロシミュレーションは、エージェント型(agent-base)のマイクロシミュレーションモデルを基礎とするものである。ここでは個人のモデルは、自律性を備えたオブジェクト、すなわちエージェントとして実装される。図1には、本シミュレーションのベースモデルとなるプロトタイプモデルのクラス図を示した。これは観察単位（エージェント）の時間的変化・行動を継続的に発生するタイプのクラスの定義である（クラスとは、エージェントまたはオブジェクトのシミュレーション言語上の定義のことである）。21 縦断調査の対象者に対応するエージェント・クラスを中心として、その属性や家族などの関係者、さらには標本集団とその統計的特性を集合的に計測、記録、出力する統計制度のエージェント・クラスを配置している。これらを基本とし、出生児調査、成年者調査など各調査ごとに、また分析テーマごとに、必要なエージェント・クラスを追加して分析モデルを構築することとなる。

図1 標本モデルのクラス図



このエージェントモデルと、縦断調査結果のデータセットを具体的に結びつけるデータクラスを図2に示した。図1における個人 (Person) (あるいはその継承クラスの「再生産する個人」 PersonReproduce など) は、実はほとんどの属性を各調査回に対応する調査票イメージのデータセットの調査記録 (Record) として保有している。調査記録 (Record) は、たとえば、最終学歴、職業などの多数の調査項目 (Item) の集合によって構成されている。さらに、各調査項目 (Item) は、その中身としてカテゴリ (Cate) によって構成されており、各カテゴリ (Cate) は基本的に名前と値を持つ。これらは調査結果データを個人別に格納するが、調査記録 (Record) のメソッドを介して、調査結果データセットのファイルから実際のデータを読み込むことになる。これらのメソッドによって、縦断調査結果データとシミュレーションの標本モデルがユーザの見かけ上単純な操作によって、直接に結び付けられる。あとは、個人の振る舞いに関する加齢モジュールに結婚、出生、あるいは就業などの行動モデルを記述することによってライフコース事象が属性や環境に依存しながら発生する様子をシミュレートすることができる。

図2 調査結果データのクラス図



エージェント型シミュレーションモデルの優れた点は、各種のライフコース事象をそのエージェントが置かれた環境やエージェントどうしの相互作用に依存するモデルを構成し、検証することができる点である。またその依存関係についても境界値による事象の制御など、非線形な関係を記述することができる。これらは通常の統計モデルでは、ほぼ不可能である。実際の個人のライフコース選択においては、環境からの影響やオールオアナッシングの判断などが重要な役割を果たしている可能性があることから、こうした機構を持つモデルについて検証することは、少子化などの現象のメカニズム解明に対して大きな貢献が期待される場所である。また、こうした非線形現象は複雑系現象として知られるが、人間行動が複雑系現象であるとの指摘がなされており、これに対して、21世紀縦断調査という優れた現実の事象データを用いることができることから、こうした分野の発展にも寄与することが大いに期待できる。

図2に示した調査結果データのクラスの実装モデル (C++) については、本事業の先行事業である「パネル調査 (縦断調査) に関する総合的分析システムの開発研究」(厚生労働科学研究費補助金 (統計情報総合研究事業)) 平成 18 年度総括報告書、ならびに平成 18~19 年総合報告書に所収したところである。

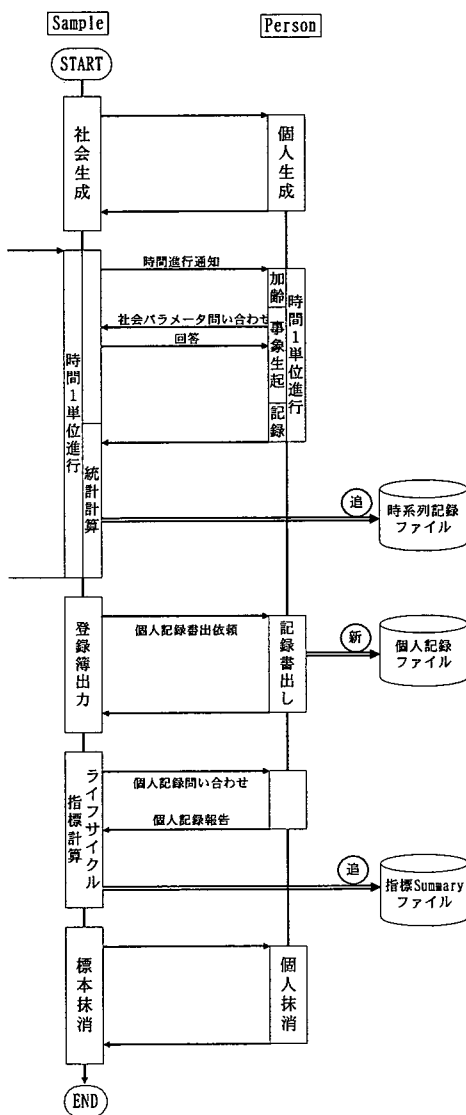
(2) モデルのシーケンス

次に図2には、本シミュレーションの基本的なシーケンス図を示した。調査対象者の個人を順に生成することによって、標本を再現する。その際、個人には実際の調査から得られた各種の個人属性が割り当てられ、時間経過にしたがって検証するモデル (規則) によって行動が発生することになる。その間に、必要であれば、個人間における相互作用が発生する。これら相互作用の過程は、通常の統計分析法では推定がほぼ不可能と考えられる領域である。出生等による新たな個人の参入については、新たなエージェントの発生と

して扱うことができる。これにより現実に近い状況を実現できるが、系の進行状況はきわめて複雑なものとなりうる。これを避けるためには、新たな個人を増やすのではなく新たな関係の発生と考え、属性の変化として扱うことも可能である。その方が状況の見通しはよくなると考えられる。

マイクロシミュレーションは、既存の統計手法の当該データ（21世紀縦断調査データ）への適用妥当性の検討や、選択的脱落・不詳の効果の推定や将来的な帰結についての予見に用いることができる。そのためには、対象標本について詳細な統計指標を算出することが必要となる。統計指標の算出は、時間単位ごとに行われるものと、ライフサイクル指標のように一定の期間終了後に算出されるものがある。いずれにせよ、それら指標はすべて記録され、標本に対する検討対象の統計手法の適用結果と詳細に比較されることによって、手法の妥当性が調べられることになる。

図2 プロトタイプモデルのシーケンス図



(2) 標本個体クラスモジュールの開発

図1 標本モデルのクラス図に見られるように、本シミュレーションにおいて個人クラス `Person` は中心的な役割を持つものである。個人クラスは継承によって再生産する個人クラス `PersonReproduce` や、さらにこれを継承した女性バージョンとなる `PersonReproduceF`、ならびに男性バージョンとなる `PersonReproduceM` が派生される。また、`PersonReproduce` からは、別の派生クラスである子どもクラス `PersonChild` のインスタンスが生成される。

今回、上記のようなすべての標本個体クラスの基底となる `Person` のクラスモジュールの開発を行った。個人クラス `Person` は、図1、図2に示すとおり、個人属性クラスや調査記録配列として調査個票 `Record` クラスを保有する。また、`Person` は集合して標本クラス `Sample` に保有される。`Person` クラスは機能（メソッド）として、誕生、加齢、死亡などの基本的なライフコースの形成を行うほか、別クラスとして用意される社会経済基本属性クラス `PerProperty` を維持管理する。結婚や出産といったライフイベントは上述のとおり、本クラスから派生された女性クラス `PersonReproduceF`、男性クラス `PersonReproduceM` において追加される機能である。

(3) 日付処理クラスモジュールの開発

マイクロシミュレーションモデル分析においては、コンピュータ内にシミュレートする人口（標本）の年月・時間の進行を独自に管理する必要がある。すなわち、人口内の擬似的な個人（エージェント）は、単位計算サイクルごとにたとえば、1ヶ月加齢し、結婚、出生、就業、死亡等のライフコース事象を経験するので、加齢や事象発生時間の記録、管理、事象間の時間の計算などは、シミュレーション過程においてきわめて重要な機能である。また、シミュレーションに用いられる暦は、うるう年などを用いず、1ヶ月を常に30日とするなど、日常用いられている暦（グレゴリオ暦）と異なる日・時間体系を操作できる必要がある。このため、先行研究に引き続き、本シミュレーションモデルに親和する日付処理クラスモジュールの開発を行った。

本クラスモジュールでは、通常の日付、時間計算の管理をベースとしながらも、シミュレーション時間モードを導入し、仮想的な時間進行が可能となるように設計された。また、各種時間計算の勘弁に行えるようにするため、加減演算の演算子を定義し、日付・時刻と期間・時間との間で、直感的な加減式によって演算が行えるようにした。

まとめ

本研究では、21世紀縦断調査データを用いたライフコース事象のマイクロシミュレーション分析を行うための基礎的システムの検討・設計を行った。マイクロシミュレーション分析は、パネル調査との親和性が強く、既存の統計分析ではできない非線形現象としての事象メカニズムの分析や、脱落等の評価が行えるため、統計分析との併用によって縦断調査データの活用範囲を広げるとともに、提供する情報の信頼性向上に資することが期待できる。本年度は、標本個体クラスモジュール、日付処理クラスモジュールの開発を行った。

参考文献

Citro, C. F. and Hanushek, E. A. (eds.) 1991, *The Uses of Microsimulation Modelling. Vol. 1: Review and Recommendations*. National Academy Press, Washington, DC.

Harding, A.(ed.), 1991, *Microsimulation and Public Policy*, Contributions to Economic Analysis, vol.232, Elsevier, Amsterdam.

Lewis, G. H. and Michel, R. C. (eds.) 1990, *Microsimulation Techniques for Tax and Transfer Analysis*. Urban Institute Press, Washington, DC.

付[1] 標本個体クラスモジュール

```

/*****
/*****      Person  Ver 1.0 : 個人クラス      *****/
/*****      Kaneko <09/02/18>      *****/
/*****
//((person 定義 file ))
// person.h
#if !defined(person_h)    //多重インクルード防止
#define    person_h

#include "..\NkDate\NkDate.h"
#include "..\Society\Society.h"

/--[ マクロ ]-----
// 一様乱数の発生
// (RAND_MAX = 0x7FFFU = 32,767) : 正規乱数発生のため rand0=0 を除外
#define UNIFORM    ((double)(rand0+1)/(double)(RAND_MAX+1))

//こちらは、0 から LONG_RAND_MAX(=2^31-1 =2,147,483,647)までの一様乱数
#define LONG_RAND_MAX    2147483647
#define UNIFORM    ((double)(_rand0)/(double)(LONG_RAND_MAX))
// 一様事象の生起
#define TRIAL_U(p)    (UNIFORM<(p)?1:0)

#define LOCATION_ORIGIN (0)    // 座標の原点
#define LOCATION_CYCLE    (360)    // 座標の周期

/===[ クラス定義 : Person ]=====
class Person {

public://公開関数-----

// クラス初期化、終了処理
static void    ClassInitialize(Society* s);    // クラス全体の初期化
static void    ClassFinalize(void);    // クラス全体の終了処理
static void    ClassCommonParameter(void);    // クラス全体の設定 (各期)

// コンストラクタ
Person(void);
Person(long idn);
Person(long idn, NkDate birthdate );
Person(long idn, NkDate birthdate, double location);

// デストラクタ
~Person(void);

// 行動操作
long    proceed(void);    // 単位時間前へ
long    retreat(void);    // 単位時間後へ

// 所属
```

```

long      ID(void){ return id; }          // 個人 I D 番号取得

// 生存関連の属性操作
bool  Alive(void){return alive;}         // 生存状態取得(true=生存, false=死亡)
NkDate BirthDate(void){return birth_date;} // 誕生日取得
NkDate DeathDate(void){return death_date;} // 死亡日取得
double Age(void);                        // 年齢取得
double Age(NkDate t);                    // 指定年月年齢(算出のみ設定はしない)
int    CompletedAge(void);               // 満年齢
int    CompletedAge(NkDate t);           // 指定年月満年齢(算出のみ設定はしない)

// ロケーション
double Location(void){return location;} // 位置座標取得

// 出力
static void OutputHead(FILE*);           // 出力表のヘッダを出力(クラス関数)
void      output(FILE*);                 // 出力表の個人データを出力
void      ProcessOut(FILE*,int);         // 途中状態の出力
void      showProperties(void);

double    Frailty(void)      { return frailty; }
double    Frailty(double f)  { return frailty=f; }
void      baptizeFrailty(double* p);

protected://内部関数-----
long  ID(long idn){ return id=idn; }     // 個人 I D 番号設定

// 生存関係
void  Birth(NkDate bd);                  // 出生発生：形式 1
void  Birth(unsigned y,                  // 出生発生：形式 2
        unsigned m,
        unsigned d);
void  Death(NkDate bd);                  // 死亡発生：形式 1
void  Death(unsigned y,                  // 死亡発生：形式 2
        unsigned m,
        unsigned d);
bool  Alive(bool a){ return (alive=a); } // 生存状態設定(true=生存, false=死亡)
NkDate BirthDate(NkDate bd);             // 誕生日設定：形式 1
NkDate BirthDate(long ld);              // 誕生日設定：形式 2
NkDate BirthDate(unsigned y,             // 誕生日設定：形式 3
        unsigned m,
        unsigned d);

NkDate DeathDate(NkDate dd);             // 死亡日設定：形式 1
NkDate DeathDate(long ld);              // 死亡日設定：形式 2
NkDate DeathDate(unsigned y,             // 死亡日設定：形式 3
        unsigned m,
        unsigned d);

// ロケーション
double Location(double l);

// 属性操作
void  aging(void);                       // 加齢

// 事象確率
bool  die();                              // 死亡判定
double probDeath(void);                   // 死亡確率

// ユーティリティ
NkDate  now(void){return home_society->Now0;} // 現在時刻
static Society* home(void){return home_society;} // 所属社会取得
static SimOpt*  Option(void){return home_society->Option0;} // オプション取得

```