

第2子出生の発生確率が低いことを示している。第1子出生時年齢も同様の結果がみられる。ただし、「1975-79年」・「1980-84年」は一般的に、産み始めの時期であり第1子出生が多く、加齢効果の影響も考慮に入れる必要がある。

学歴については、統計学的に有意な結果はみられなかった。これは予備推定において、「高校」をリファレンスカテゴリーにおいた相対確率においても同様の結果であった。初職については、「正規の職員」に比べ、モデル1では「パート・アルバイト」である場合、負の効果を示し、「無職・学生」である場合、正の効果を示している。モデル2では「無職・学生」の場合、正の効果がみられた。

モデル2とロジスティック回帰分析との比較においては、出生コーホートと第1子出生時年齢でロジスティック回帰分析の推定値が高めに推定されている。これは、ロジスティック回帰分析ではタイミング効果を考慮することができないため、加齢効果が推定値に大きく影響を与えているものと考えられる。学歴と初職については、ほぼ同水準であるといえる。

最後に、特定のモデルを用いない生存関数（サバイバル確率）の推定を行う。ここでは、一般的なカプラン・マイヤー法を用いた。図3から図6に生存関数の分布を示した。分布から得られるカテゴリーごとの差はCoxモデルによる推定結果に符合する。

7. おわりに

本稿では、パネル調査データのマイクロ分析として有力な分析手法であるイベントヒストリー分析を山口一男「イベントヒストリー分析」『統計』（2002-2003, [1]-[15]）を通じてレビューしてきた。解説の手引きとして、21世紀成年者調査や21世紀出生児調査を用いた分析を提示したが、両調査はまだ開始して間もないことから年次のデータの蓄積が多くなく、パネル調査の長所である各年の属性の変化を活かしきるにいたらなかった。とはいえ、今後継続的にデータの蓄積が可能になるにつれ、分析対象となるイベントのケース数が増加することによって、分析可能範囲は広がることは必至である。

参考文献

- Allison, Paul D. 1984. "Event History Analysis-Regression for Longitudinal Event Data, Sage Publications, Inc.
- 大橋泰雄・浜田知久馬, 1995. 『生存時間解析 SASによる生物統計』東京大学出版会.
- 山口一男, 2002-2003. 「イベントヒストリー分析(1)~(14)」『統計』52(9)~53(11).

図3 カプラン・マイヤー法による出生 cohorts の生存関数 (サバイバル関数) 分布

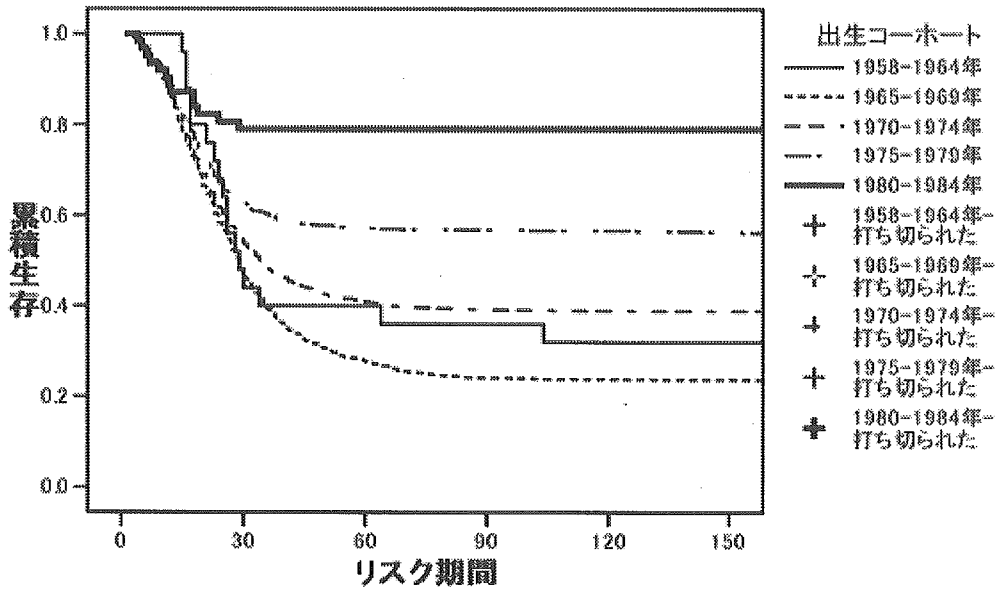


図4 カプラン・マイヤー法による第1子出生時年齢の生存関数 (サバイバル関数) 分布

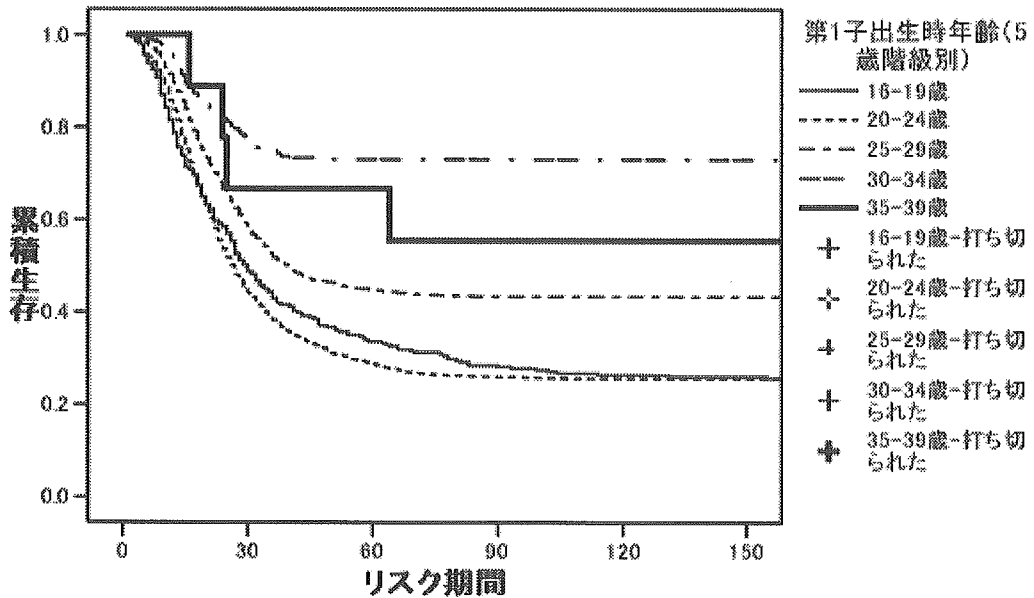


図5 カプラン・マイヤー法による学歴の生存関数（サバイバル関数）分布

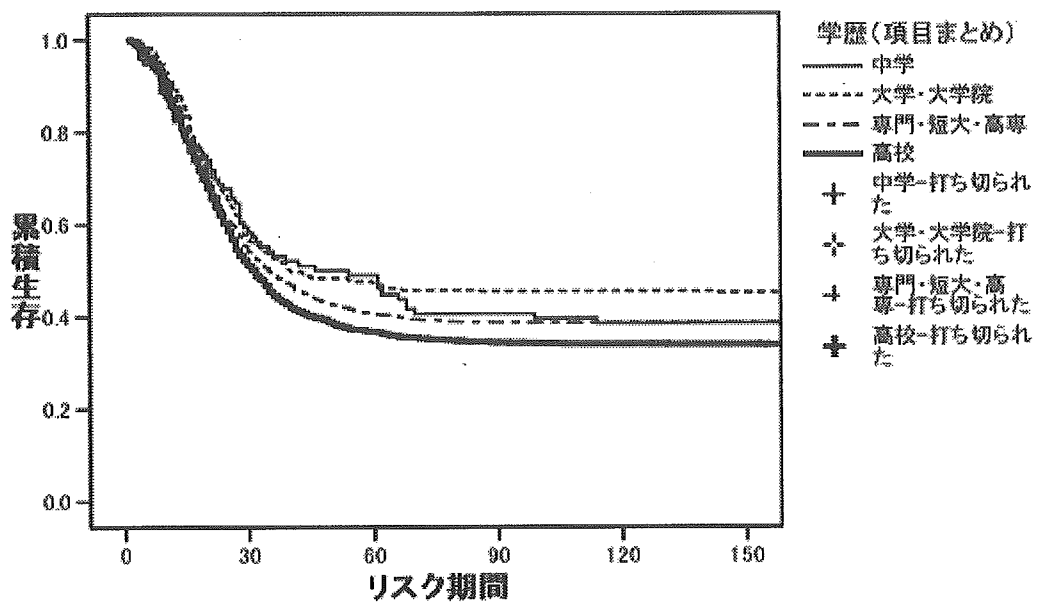
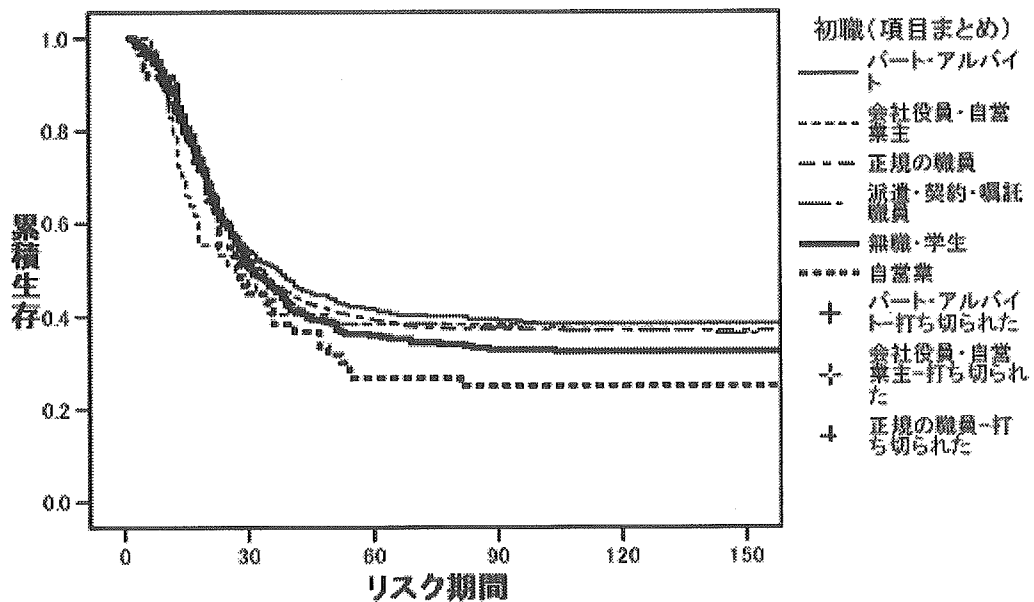


図6 カプラン・マイヤー法による初職の生存関数（サバイバル関数）分布



3 縦断調査マイクロシミュレーション分析の基礎システムの開発

金子 隆一

1. はじめに

パネル調査(縦断調査)の分析法の一つとして、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手法である。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、その分析手法として親和性が高いといえよう。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。

本研究では、諸外国の事例について検討を行った後、21世紀縦断調査を基にしたマイクロシミュレーション分析を行うための基礎フェーズに対する支援システムの開発を行った。すなわち、パネル調査データの管理情報を基に、シミュレーション分析に必要な標本モデルをシミュレーション言語(現行ではC++)と連携しながら生成するシステムを作成した。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。本システムによれば、縦断調査データに即したさまざまなタイプのマイクロシミュレーション分析を比較的簡単に展開することができる。

2. パネル調査とマイクロシミュレーション

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 micro-simulation 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。

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。

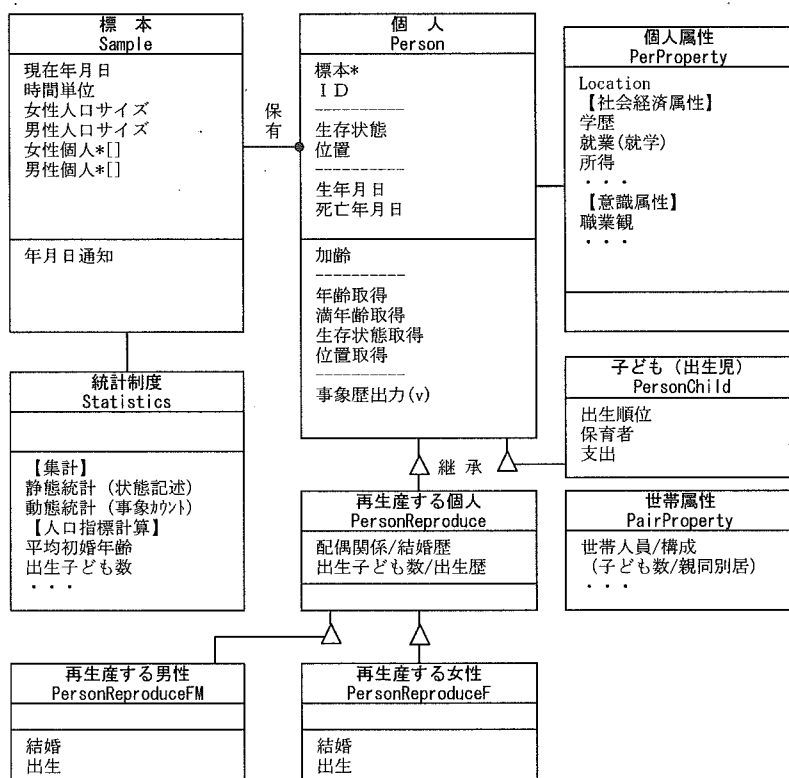
3. システムの開発とプロトタイプモデル

(1) マイクロシミュレーションモデルの概要

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。そ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 プロトタイプモデルのクラス図



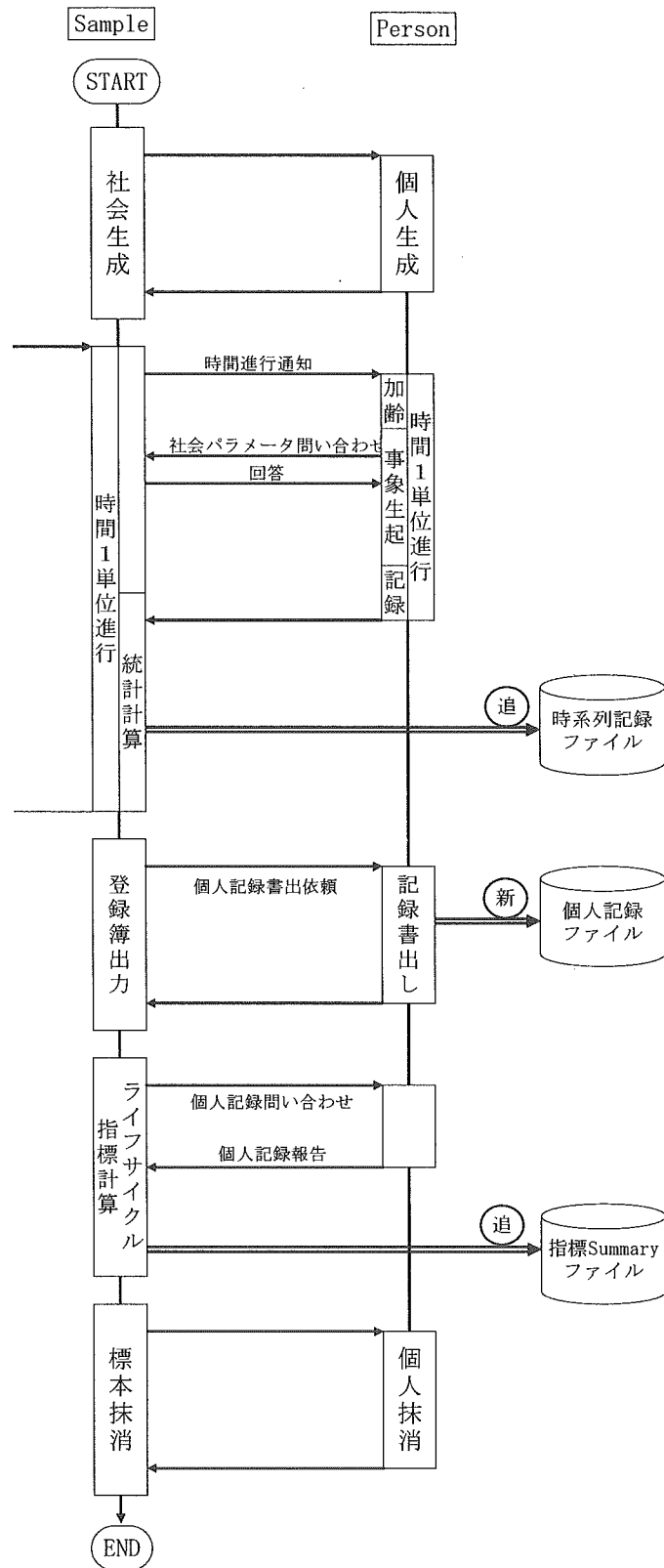
本システムでは、この個人クラスをパネル調査データ情報から自動的に生成することとする。これにより、作成されたクラスに対して、縦断調査の実データを実装することによって、シミュレーションの対象となる標本を生成することができる。この標本を、以下に示す経時的に過程の中におくことによって、個人ならびに集合単位（夫婦、家族、社会等）の行動が模擬され、シミュレーションが行われる。

(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



(3) シミュレーション用標本生成システムの開発

以上のプロトタイプで示された枠組みを持つシミュレーション分析に対して、21世紀縦断調査データを用いた標本を生成するシステムを開発した。システムは、データマネジメントシステムとして開発されたコード表変換システムを利用し、シミュレーション言語（現行では C++）によって記述されたシミュレーションのフレーム（上記プロトタイプで示される）に対して、標本を自動生成して供給する。シミュレーションのフレームには、この標本を読み込む機能が用意されており、これらの連携によりシミュレーションの基礎部分が形成されることになる。さらにシミュレーションのフレームは、図2のモデルのシーケンスで示された経時変化がシミュレートされ、これに対して目的に応じた個人、夫婦、家族ならびに集団の振る舞いを定式化して与えることにより、シミュレーション分析が行われる。

図3には、本システムによって自動生成された標本記述の例（一部）を示した。この記述にしたがってシミュレーション言語は標本を読み込む。その際に、脱落や回答不詳に対しては、目的にあった補充 implementation が行われる。

図3 標本記述の例（成年者調査データの一部）

```
//*****  
//*** C++ Class : ***  
//*** 第1回成年者縦断調査（女性票） ***  
//*** [ test_code_adult_female01.xls (AF01) ] ***  
//*** <Lexis> 2006/02/16 ***  
//*****  
  
[Data]  
Name = 第1回成年者縦断調査（女性票）  
DataFileName = D:\ProjectR\Panel\00_Data\sample\baby\af01.dat  
  
[Variable]  
調査票番号<1,1>{ }  
KEY番号（世帯情報） 地区番号<2,5>{  
    [01002-47090]地区番号}  
KEY番号（世帯情報） 単位区番号<7,2>{ }  
KEY番号（世帯情報） 世帯番号<9,2>{ }  
KEY番号（世帯情報） 該当者番号<11,1>{ }  
KEY番号（世帯情報） 配偶者番号<12,1>{  
    [1-6]1-6,  
    [.]非該当}  
出生年月 年<13,2>{  
    [42-57]昭和42-57年,  
    [.]不詳}  
出生年月 月<15,2>{  
    [01-12]01-12月,  
    [.]不詳}  
問1 最終学歴<17,1>{  
    [1]中学,  
    [2]高校,  
    [3]専門学校,  
    [4]短大・高専,  
    [5]大学,  
    [6]大学院,  
    [7]その他,  
    [9]不詳,  
    [.]不詳}  
問1 卒業・在学の別<18,1>{  
    [1]卒業,  
    [2]在学中（休学等を含む）  
    [9]不詳,  
    [.]不詳
```

まとめ

本研究では、諸外国の事例について検討を行った後、21世紀縦断調査を基にしたマイクロシミュレーション分析を行うための基礎的なシステムの開発を行った。本システムは、パネルデータの管理情報を基に、シミュレーション分析に必要な標本モデルをシミュレーション言語（現行では C++）と連携しながら半自動的に生成するシステムである。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。本システムによれば、縦断調査データに即した各種マイクロシミュレーション分析を比較的簡単に展開することができるが、本年の研究ではその基本機能を実現することができた。

マイクロシミュレーションは、既存の統計分析に止まらない分析手法として、結婚・出生・子育てなど分析対象となる事象の発生メカニズム、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。また、既存の統計モデルと合わせて用いることで、それらの信頼性を検証することができるので、パネル調査における統計分析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ことで、より信頼性の高い分析結果の獲得に資することが期待される。ここでは21縦断調査データから、シミュレーションの標本を自動生成するシステムを開発し、今後の分析の基礎を与えた。これにより、諸外国で行われているタイプの政策関連のマイクロシミュレーションをはじめ、今後、本縦断調査に即した多様なモデルが開発されることが期待される。

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.

II. 個別研究報告（データマネジメント）

4 21世紀パネル(縦断)調査データベースシステム(PDB21)の開発

金子 隆一
三田 房美

はじめに — システムの目的と概要 —

本事業では、縦断調査によって毎年継続して蓄積されて行く統計データに対し、その集計・分析法をも考慮した有効で具体的なデータ管理・分析システムを検討し、開発することを目指している。パネル調査(縦断調査)は同一対象(個人)を追跡しながら継続的に調査を実施するものであり、その有効性を十分に引き出すためには横断調査とは異なるデータ管理、および分析が要求される。最も異なる点は、逐次累積される複数の調査回データセットを個人をキーとしながら連係させて集計・分析を行わなくてはならない点であろう。成年者縦断調査においては、男性票・女性票データセットを用いて夫婦単位の集計分析なども行われるため、データ操作はさらに複雑となる。

本事業におけるこれまでの研究では、縦断調査のデータ管理・分析システムとしての基本的課題とその対処するためのシステム要件について検討を行いエクセル・コード表を中心としたデータ管理・分析システムを開発した。ここではこれらを当初の構想にしたがってさらに発展させ、現在広く用いられている標準的リレーショナル・データベース・システムの枠組みを応用することにより、各回および各調査票に対応する多数のデータセットを連携させるシステムを開発した。これによれば、特定のテーマに対する集計ならびに統計分析の際に、必要な変数値を複数のデータセットから自動的に集約し、集計ならびに統計分析用のソフトウェアに受け渡すことができるので、分析者において毎回多数の指定を行う負担を軽減し、また多様かつ多重な変数の扱いにおける混乱や誤りを防ぐことができると期待される。

本システムは、21世紀パネル(縦断)調査データベースシステム(PDB21)と称する。その基本的な機能は、21世紀縦断調査によって毎年継続して蓄積されていく調査回データセットを管理し、集計・分析の支援を行うことである。ただし、現行システムでは、出生児縦断調査および成年者縦断調査を対象とし、各回調査データの管理と分析時の連携を図り、とくに成年者縦断調査においては、女性票・男性票・配偶者女性票・配偶者男性票を用いた配偶関係による連関を考慮した分析に対処する。すなわち、本システムは、複数の調査回ならびに調査票に対応したデータセットを、個人をキーとしながら各調査項目(変数)の条件指定を行なうことによりデータを抽出し、滞りなく各種の集計・分析ソフトウェアに受け渡す機能を有している。また、そのような多数のデータセットをデータベース化するためのデータ作成機能を有している。

システムは、リレーショナル・データベース・システム PostgreSQL(Ver.8.0.4)をデータベースエンジンとして使用し、DICS-IV(または ADAM)、SPSS、SAS、Perl その他の一般の集計または統計分析ソフトウェアと連携する。ただし、基本的ユーザインターフェースは、すべて EXCEL を用いることでユーザに対する操作習熟に対する要求は最低限度となるように配慮した。

システムの詳細は、次ページ以降のパネル(縦断)調査データベースシステムの説明を参照されたい。また、使用方法についてはこれに続くマニュアル編を参照のこと。

パネル(縦断)調査データベースシステムの説明

《 目 次 》

I. システム概要

1. システム概要	1
1-1. システム名称	1
1-2. システムの概要	1
1-3. システム構成図	2
1-4. ハードウェア/ソフトウェア仕様	3

II. システム機能

1. システム機能一覧	1
1-1. システム機能一覧	1
2. システム機能仕様書	2
2-1. パネル調査データ抽出機能	2
2-2. パネル調査DBデータ作成機能	7

III. データベース定義

1. データベース定義書	1
1-1. データベース名	1
1-2. データベース属性	1
1-3. グループ	1
1-4. ユーザー	1
1-5. E-R図	2
2. テーブル/ビュー一覧表	3
3. テーブル定義書	4
3-1. 調査票一覧テーブル	4
3-2. 調査票情報テーブル	8
3-3. コードマスタテーブル	12
3-4. 出生児調査テーブル	16
3-5. 成年者調査テーブル	20

IV. ファイル定義

1. ファイル一覧表	1
1-1. ファイル一覧表	1

付属資料

I. システム概要

1-1. システム名称

21世紀パネル（縦断）調査データベースシステム（PDB21）

1-2. システムの概要

パネル調査データベースは、パネル調査（縦断調査）によって毎年継続して蓄積されていく調査回データセットを管理するためのシステムである。

パネル調査には出生児縦断調査および成年者縦断調査があり、成年者縦断調査は女性票・男性票・配偶者女性票・配偶者男性票から成っている。

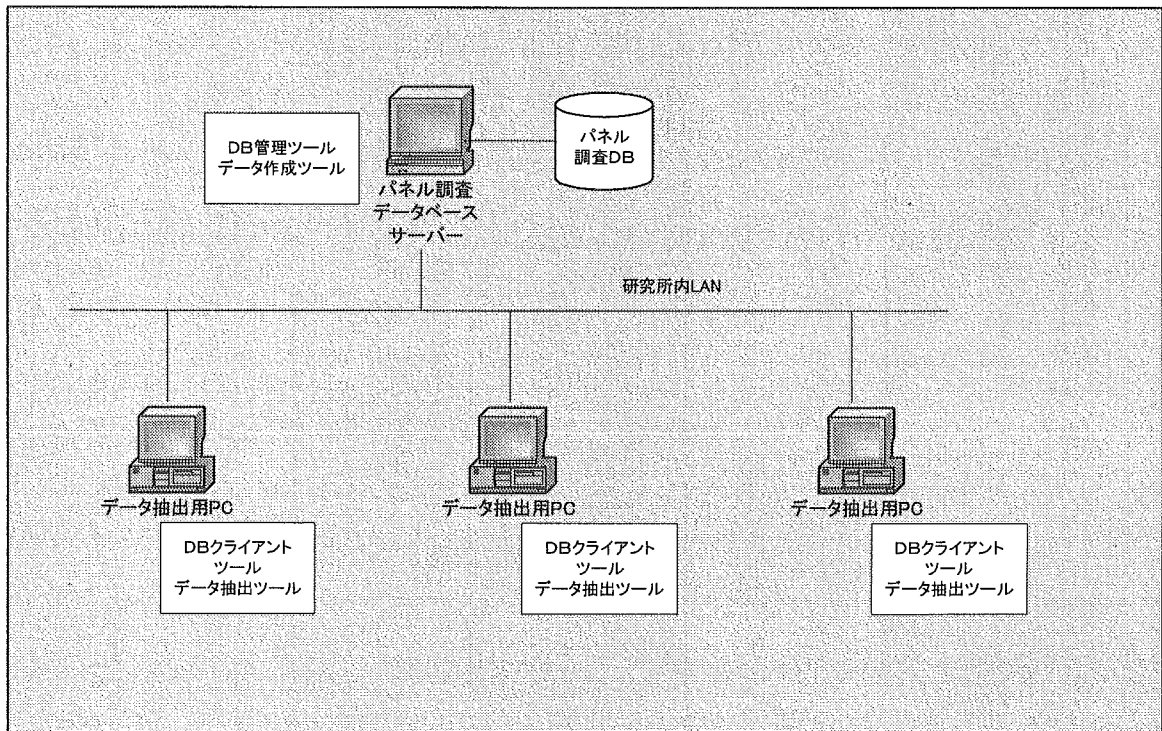
パネル調査データベースでは、逐次蓄積される複数の調査回データセットを個人をキーとしながら各調査項目の条件指定を行なうことによりデータを抽出する機能を有している。

また、そのような複数の調査回データセットをデータベース化するためのデータ作成機能を有している。

本システムは、以下の機能をもつシステムより構成されている。

- (1) データ抽出SQL文作成機能
- (2) パネル調査DBデータ作成機能
- (3) データベース管理機能

1-3. システム構成図



1-4. ハードウェア/ソフトウェア仕様

(1) ハードウェア仕様

H/W 番号	ハードウェア 名称	メーカー /型番	OS	主要ソフトウェア	詳細説明 (スペック、用途、設置場所など)
1			Windows XP Windows2000	PostgreSQL Ver.8.0.4	用途：DB サーバー CPU： メモリ： ディスク：
2			Windows XP Windows2000	MS Access 2000 MS Excel 2000 PostgreSQL Ver.8.0.4 (クライアント機能のみ)	用途：クライアントコンピュータ

(2) ソフトウェア仕様

S/W 番号	ソフトウェア名称	メーカー/型番	対応 OS	ライセンス数	詳細説明
1	PostgreSQL Ver.8.0.4		Windows XP Windows 2000	ライセンスフリー	リレーショナルデータベースシステム
2	ODBC		Windows XP Windows 2000	ライセンスフリー	データベース接続ミドルウェア
3	MS Access 2000	MicroSoft	Windows XP Windows 2000		
4	MS Excel 2000	MicroSoft	Windows XP Windows 2000		

II. システム機能

1. システム機能一覧

No.	システム機能名		開発ソフト・開発言語 など	概要
	名称	定義名		
1	パネル調査データ 抽出機能	パネル調査データ抽出.xls	MS Excel VBA マクロ PostgreSQL	<ul style="list-style-type: none"> 指定した抽出条件によりパネル調査データベースから調査票データを抽出する SQL 文を作成する。 作成した抽出 SQL 文を実行する。 抽出されたデータを CSV ファイルに出力する。 指定した抽出項目に関する変換記述表を作成する。 指定した抽出条件をコントロールファイルとして保存する。
2	パネル調査DBデータ作成機能	出生児調査票： パネル調査データ作成 (出生児) .xls 成年者調査票： パネル調査データ作成 (成年者) .xls	MS Excel VBA マクロ MS Access クエリ PostgreSQL	<ul style="list-style-type: none"> パネル調査データベースに格納するDBカタログ (調査票情報・コードマスタ) のテーブルを作成する。 パネル調査データベースに格納するDBカタログ (調査票情報・コードマスタ) の取り込みデータを作成する。 パネル調査データベースに格納する調査票データのテーブルを作成する。 パネル調査データベースに格納する調査票データの取り込みデータを作成する。
3	データベース管理機能		PostgreSQL	<ul style="list-style-type: none"> パネル調査データベースを管理する機能 バックアップ等

2-1. パネル調査データ抽出機能

(1) 処理内容

説明	入出力ファイル一覧	
	ファイル名	ファイルの内容等
① 抽出条件を指定する。 ・抽出したい調査票および調査項目（変数名）を指定する。 ・調査票および調査項目（変数名）とその抽出条件を指定する。 ② 抽出SQL文作成実行 ・指定した抽出条件により抽出SQL文を作成する。 ③ データ抽出の実行 ・作成した抽出SQL文を実行する。 ・抽出されたデータをCSVファイルに出力する。 ④ 変換記述表の作成 ・指定した抽出項目に関する変換記述表を作成する。 ⑤ 抽出条件の保存 ・指定した抽出条件をコントロールファイルとして保存する。	<入力> 抽出項目 ① 調査票番号 ② 調査項目（変数名） 抽出条件 ① 調査票番号 ② 調査項目（変数名） ③ 条件式 <出力> ① データ抽出SQL文 ② 抽出結果 ③ 抽出項目 ④ 抽出条件	<入力> 抽出項目 ① 端末入力 ② 端末入力 抽出条件 ① 端末入力 ② 端末入力 <出力> ① 端末表示・SQLファイル ② 端末表示・抽出データファイル ③ 変換記述表 ④ コントロールファイル

(2) 機能説明

1) 抽出条件を入力する。

- ① 抽出したい調査票および調査項目（変数名）を指定する。

抽出条件入力画面で抽出したい調査票の調査票番号および調査項目（変数名）を入力する。

- ② 調査票および調査項目（変数名）とその抽出条件を指定する。

抽出条件入力画面で抽出条件を入力する。抽出条件は調査票番号および調査項目（変数名）と抽出条件式で指定する。

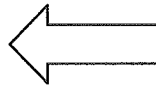
2) 抽出SQL文作成実行

- 1) で指定した抽出条件により抽出SQL文を作成する。

- ① 抽出条件入力画面で抽出SQL文作成を実行する。

SQL文の構成：

```
SELECT (抽出項目)
FROM (調査票テーブル)
WHERE (条件文)
;
```



No.	抽出項目		抽出条件			※
	調査票番号	変数名	調査票番号	変数名	条件	
5	bk03_0	A001	bk03_0	A015	=	3
6		A002		A016	>=	50
7		A003	bk03_1	A039	=	1
8	bk03_1	A068				
9		A069				
10		A070				
11		A071				
12		A072				
13						
14						
15						
16						
17						
18						

抽出項目：[調査票テーブル省略名] + [変数名] ※調査票テーブル省略名：
調査票番号から生成

調査票テーブル：[調査票テーブル] [調査票テーブル省略名]

条件文：[調査票間のキー項目の結合条件]、[調査票テーブル省略名] + [変数名]
+ [条件式 (式・値)]

※各条件間はAND 結合

例) 抽出項目： 調査票番号：bk03_0 変数名：A001、A002、A003

調査票番号：bk03_1 変数名：A068、A069

抽出条件： 調査票番号：bk03_0 変数名および条件式：A015='3'、A016 >='50'

調査票番号：bk03_1 変数名および条件式：A039='1'

以下の抽出SQL文が作成される。

```
SELECT
    tbl1.A001, tbl1.A002, tbl1.A003,
    tbl2.A068, tbl2.A069
FROM
    bdt_bk03_0 tbl1, bdt_bk03_1 tbl2
WHERE
    tbl1.A001 = tbl2.A001 AND tbl1.A002 = tbl2.A002 AND
    tbl1.A003 = tbl2.A003
    AND tbl1.A015 = '3'
    AND tbl1.A016 >= '50'
    AND tbl2.A039 = '1'
;
```

② 作成された抽出SQL文は表示画面に出力される。

3) データ抽出の実行および結果の出力

① 作成した抽出SQL文を実行する。

本機能ではSQL文の厳密な構文チェックは行なっていない。

SQL文エラーなどで検索に失敗する場合、PostgreSQL クライアントツールで抽出SQL文を実行しエラー原因を確認する。

(後述「※PostgreSQL クライアントツール “pgAdmin” でのデータ抽出方法」および「データ抽出マニュアル」を参照)

② データベースの検索が正常に実行された場合、抽出結果をCSVファイルで出力する。

出力ファイル名：「out_ (日付 YYYYMMDD) .csv」 例) 2005年12月16日の場合「out_20051216.csv」

出力先フォルダ：データ抽出ツール実行フォルダ

※PostgreSQL クライアントツール “pgAdmin” でのデータ抽出方法：

① 作成した抽出SQL文はPostgreSQL クライアントツール “pgAdmin” でも実行できる。

出力された抽出SQL文をコピーし、PostgreSQL クライアントツール “pgAdmin” のSQL実行画面に貼り付け実行する。

② 抽出されたデータを PostgreSQL クライアントツール “pgAdmin” でファイルに出力する。

PostgreSQL クライアントツール “pgAdmin” に抽出結果が表示される。

表示された結果をクライアントツール “pgAdmin” のエクスポート機能でファイルに出力する。

4) 変換記述表の作成

1) で指定した抽出項目をもとに変換記述表を作成する。

出力する内容は以下の通り。

[ヘッダー部分]

No.	項目名	表示行	列	内容	備考
1	表題	1行目	A	「パネル調査データ抽出」	※出力内容固定
2	データセット名	2行目	A	out_ (日付 YYYYMMDD)	例) 2005年12月16日の場合「out_20051216」
3	データファイル名	2行目	B	出力ファイル名「out_ (日付 YYYYMMDD) .csv」	例) 2005年12月16日の場合「out_20051216.csv」
4	データフォルダ名	2行目	F	出力先フォルダ名	(実行フォルダ)

[調査票情報部分] 調査票情報テーブルから指定した調査票番号(survey_id)・変数名(item_id)で取得する情報

No.	項目名	表示行	列	内容	備考
1	変数名	4行目以降	A	指定された調査票番号・変数名 [調査票情報テーブル・調査票番号] - [調査票情報テーブル・変数名] で記述	例) 調査票番号 : bk03_0 変数名 : A015 の場合
2	項目名	4行目以降	B	調査票情報テーブル・項目名(item_name)	
3	位置	4行目以降	C	調査票情報テーブル・位置(pt_position)	
4	幅	4行目以降	D	調査票情報テーブル・幅(length)	
5	コード範囲	4行目以降	E	調査票情報テーブル・コード範囲(range)	
6	内容	4行目以降	F	調査票情報テーブル・内容(content)	
7	変換情報	4行目以降	G	調査票情報テーブル・変換情報(content)	c : 文字、n : 整数

[コードマスタ部分] コードマスタテーブルから指定した調査票番号(survey_id)・変数名(item_id)で取得する情報

No.	項目名	表示行	列	内容	備考
1	コード	4行目以降	E	コードマスタテーブル・コード(item_code)	
2	カテゴリ (コードの内容)	4行目以降	F	コードマスタテーブル・カテゴリ (category)	

5) 抽出条件の保存

① 指定した抽出条件をコントロールファイルとして保存する。

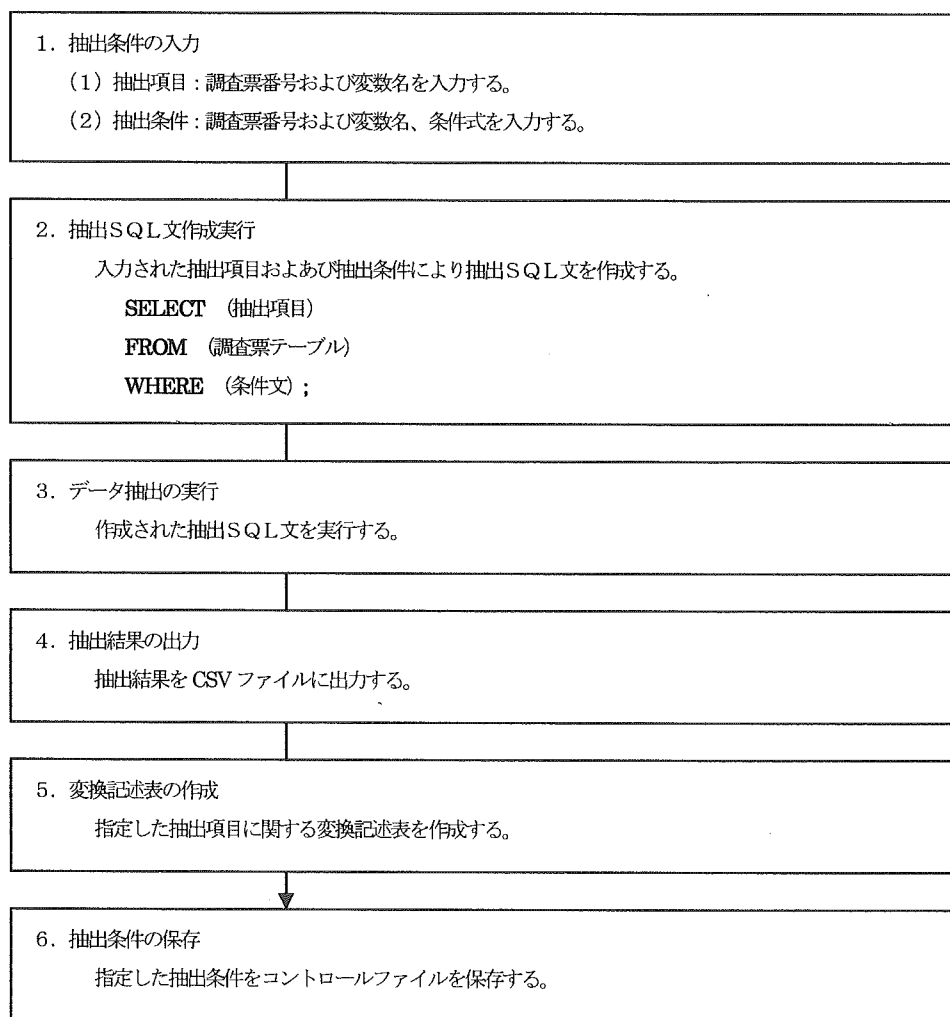
抽出項目の情報として調査票番号および調査項目(変数名)、抽出条件として調査票番号および調査項目(変数名)とその条件式をコントロールファイルに出力する。

② ファイル名および保存先

ファイル名 : panel_ctl.txt 保存先 : データ抽出ツール実行フォルダ

※抽出条件コントロールファイルは常に上書きされるので注意

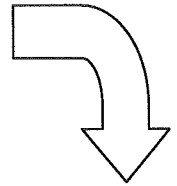
【処理フロー】



(3) 処理イメージ

抽出条件入力画面

抽出項目	抽出条件	条件
調査年度	調査番号	条件
bk03 0	A001	A015 = 3
	A002	A016 >= 50
	A003	
bk03 1	A068	
	A069	
	A070	
	A071	
	A072	



出生児調査票1回、2回から
父の出生年が昭和50年以上の
A001(調査年)・A002(調査番号)・A003(認識番号)
A069~A072の調査項目を抽出

```

1 抽出SQL文:
2 SELECT
3   a.A001, a.A002,
4   b.A068, b.A069,
5 FROM bk03_0 a, bk03_1 b
6 WHERE a.A001 = b.A001
7 AND a.A002 = b.A002
8 AND a.A003 = b.A003
9 AND a.A015 = 3
10 AND a.A016 >= 50
    
```