

図1 標本モデルのクラス図

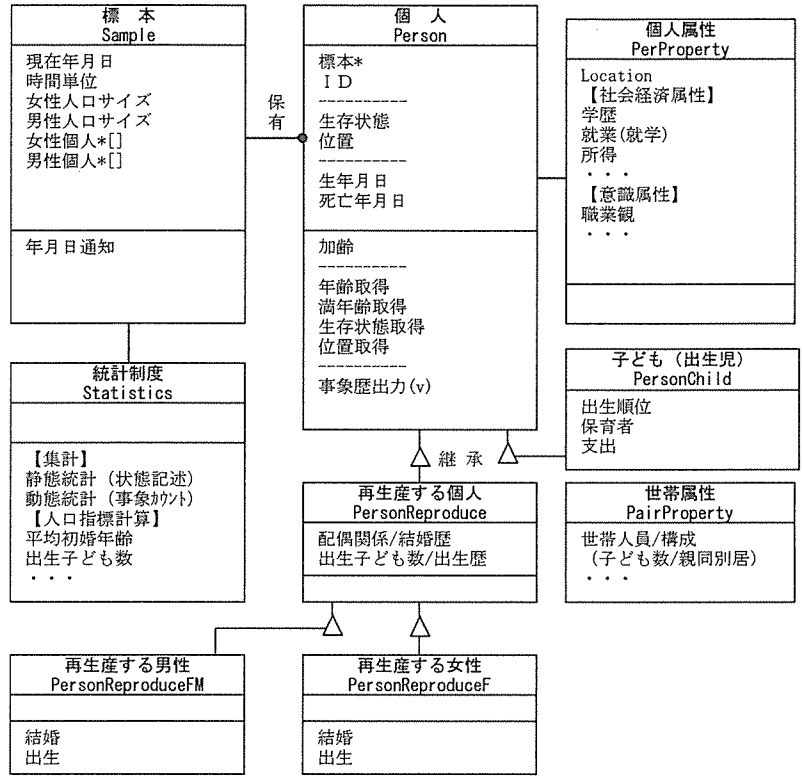
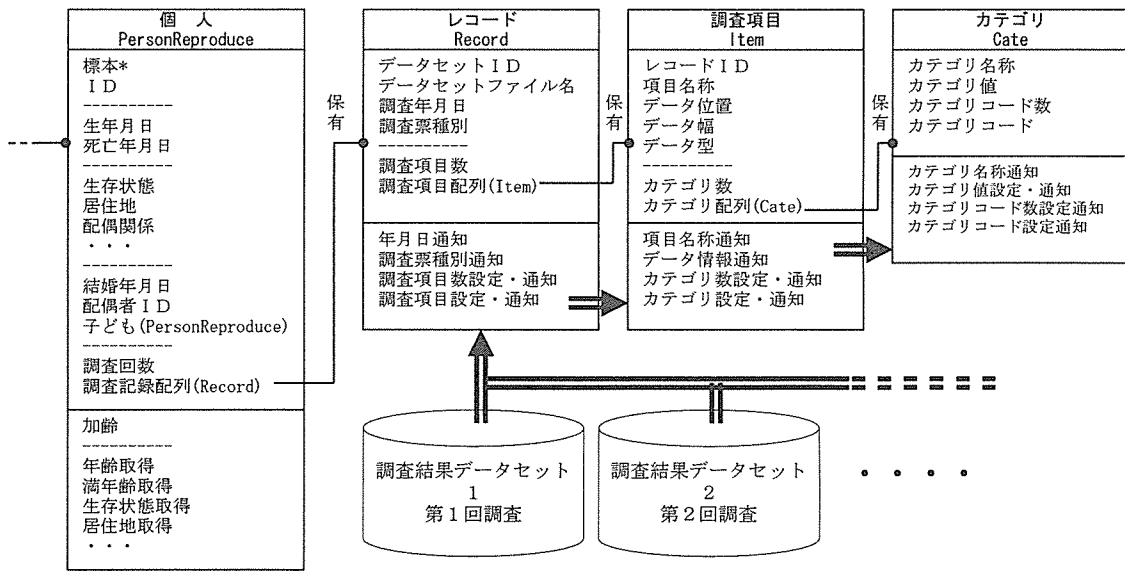


図2 調査結果データのクラス図



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レコード(Record)として保有している。調査レコード(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、調査レコード(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。

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、非線形の関係性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ことである。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。

以下には、図2に示した調査結果データのクラスの実装モデル(C++)を示す。

```
//==[ クラス定義 : Cate ]=====
class Cate { // category

public:

    // コンストラクタ
    Cate(void);
    Cate(char* str);
    Cate(char* name, int cd, double val);

    // コピー-コンストラクタ
    Cate(Cate &);

    // デストラクタ
    virtual ~Cate(void);

    // 演算子(=)
    Cate &operator=(Cate &);

    // メッセージ操作
```

```

char *Name(void);          // 名前取得
char *Name(char *);      // " 設定
double Value(void);      // 値の取得
double Value(double);    // " 設定

int Ncode(void);         // 対応するコードの数を取得
int Code(int i);        // i 番目のコードを取得 i の有効性チェック後)
                        // (i=0,1,...,ncode-1, i が範囲外なら-1 を返す)
int Code(int i, int c);  // i 番目のコードに c を設定(i の有効性チェック後)
                        // (i=0,1,...,ncode-1, c を返すが、i が範囲外なら-1 を返す)

void show(void);        // マナーの表示

private:

char *getName(dscr);
double getValue(dscr);
int getNcode(dscr);
int *getCode(dscr);

private:

char *name;             // category name
double value;          // category value
int ncode;              // # of assigned code to the category
int *code;             // assigned code list
};
//===[ クラス定義 : Item ]=====
class Item { // item (=variable)

public:

// コンストラクタ
Item(char *str);

// デストラクタ
virtual ~Item(void);

// メンバ操作
char* Name(void);      // 名前の取得
char* Name(char *n);  // " 設定
int Loc(void);        // 位置の取得
int Wid(void);       // 幅の取得
int Type(void);      // タイプの取得

int Ncate(void);     // カテゴリ数の取得
char* CateName(int c); // c 番目のカテゴリ名の取得
double CateValue(int c); // c 番目のカテゴリ値の取得
int UnknownCate(int* uk); // 不詳コードを int 配列で取得

void show(int mode=1); // mode... Cate を表示するか否か

private:

int location(char*); // 位置情報文字列から位置情報設定
int category(char*); // カテゴリ情報文字列からカテゴリ情報設定

private:

char* name;          // 名称
char* locstr;       // 位置情報文字列(情報読み取り関数が完成したら削除)
char* catestr;      // カテゴリ情報文字列( " )

int loc;            // location
int wid;           // width
int type;          // type

unsigned ncate;     // # of categories
Cate** cate;       // category list

int nuk;           // # of unknown categories
int* uk;           // unknown category list

```

```

};
//==[ クラス定義 : Record ]=====
class Record { // record

public:

    // コンストラクタ
    Record(const char *rdf);           // rdf からすべて読む
    Record(const char *rdf, char** list); // rdf から list にある item を選択して読む

    // デストラクタ
    virtual ~Record(void);

    // メッセージ操作
    char* Name(void);                 // 名前取得
    char* Name(char *n);              // " 設定
    char* DataFile(void);             // ファイル名取得
    char* DataFile(char *n);         // " 設定
    int Nitem(void);                  // item の数を取得
    Item* ItemAd(int i);              // i 番目の item を取得(ポインタ)
    char* ItemName(int i);            // i 番目の item の名前取得
    int ItemLoc(int i);              // i 番目の item の位置取得
    int ItemWid(int i);              // i 番目の item の幅取得
    int ItemType(int i);             // i 番目の item のタイプ取得

    int ItemNmOnlist(int i);         // 呼び出し側から与えられたリスト上での、
    // item[i] の位置を取得
    Item* ItemOnList(int j);        // 呼び出し側から与えられたリスト上 j 番目の
    // item を取得

    // メンバ表示
    void show(int mode=1);          // mode... Gate を表示するか否か

private:

    int readRdf(const char *rdf, const char** list);
    // レコード記述ファイルの読み込み
    // list にある item のみ選択して読み込む
    // list=0 ならすべて読み込む

    int isitemIn(const char* istr, const char** list);
    // item 記述の中の item 名が list にあるかどうか check
    // 有れば位置を、無ければ-1 を返す
    // : list の最後の要素は空文字列

private:

    char *name;                     // record name
    char *datafile;                 // data file name(full path)
    int nitem;                      // # of variables
    Item *item[MAXITEM];           // item list
    int itnm[MAXITEM];             // 各 item が、呼び出し側から与えられたリスト上の
    // 何番目(0,1...)を記録
};

```

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。次年度はここで設計したシステムを実装し、実際のシミュレーション分析を行う。

5 諸外国のパネル調査の概要の閲覧システムの開発

釜野 さおり

本研究では、平成 16～17 年度の「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補助金（統計情報高度利用総合研究事業））において実施した、諸外国のパネル調査に関して収集した情報（詳細は、同研究総合報告書の「諸外国におけるパネル調査レビュー」を参照のこと）を、ウェブサイトで閲覧できるシステムを開発した。これらのパネル調査は、国際的に見ても充実しているミシガン大学の大規模なデータバンクと、小規模であるものの、社会科学系の縦断調査のデータをまとめて所有している、ハーバード大学のデータバンク(Murray Center)を中心にサーベイを行って収集したものである。ミシガン大学の ICPSR から抜き出した調査は、2005 年 8 月時点で”panel”ならびに”longitudinal”のキーワードで検索された調査の内容を、それぞれ確認した結果に基づいている。Murray Center のデータバンクから抜き出したものは、2005 年 5 月時点で登録されていたもののデータ形式を確認し、longitudinal と記されたものである。これらのうち、調査対象者が 100 人に満たない規模の小さなものや、データ自体がパネル形式であっても個人に対して行った調査に基づいたものではないもの（たとえば州の統計に基づいたデータ）は除外している。

現在、本閲覧システムを、国立社会保障・人口問題研究所のホームページで実験的に公開する準備をしている。今後も引き続き情報を収集し、将来は一般にも公開できる閲覧システムを目指している。

パネル調査閲覧用のウェブサイトは、以下のような構成になっている。

- 1) まず、content にアクセスすると、以下の【画面 1】のように、本一覧をなす 3 分類が List として表示される。
- 2) それぞれの List をクリックすると、その分類に入る調査の一覧（各分類における通し番号、英文調査名と和文調査名、調査のテーマ）が表示される。たとえば List1 をクリックすると、【画面 2】のように表示される。List 1 には、149 の調査、List 2 には 40 の調査、List 3 には 18 の調査が含まれている。
- 3) 目次上で、さらに詳細を知りたい調査名をクリックすると、その調査の概要の書かれた画面が表示される。たとえば、List 1 の、8. Americans' Changing Lives: Waves, I, II, and III, 1986, 1989, and 1994（「アメリカ人の生活の変容」）をクリックすると、【画面 3】のような表が表示される。
- 4) 可能な限り、オリジナルの情報源へのリンクが張ってある。たとえば、上記の調査であれば、表上の「ウェブサイトのリンク先」掲載されている、<http://webapp.icpsr.umich.edu/cocoon/ICPSR-STUDY/06438.xml> をクリックすると、この情報の元となる ICPSR (Inter-University Consortium For Political and Social Research) の中、当該調査やデータ使用の方法に関する元の情報にアクセスすることができる。

*****【画面1】*****

パネル調査 タイトル一覧

- ・ LIST 1. 健康・医療・家族生活・家計・労働・教育に関する調査
- ・ LIST 2. 政治・選挙に関する調査
- ・ LIST 3. 犯罪・防犯・警察・更生プログラム・依存治療等のプログラム効果の調査

*****【画面2】*****

パネル調査 タイトル一覧

1. 健康・医療・家族生活・家計・労働・教育に関する調査

#	調査タイトル	分類
1	<u>Abortion Study, 1969-1974</u> <u>「人工妊娠中絶に関する調査」</u>	妊娠出産・リプロ
2	<u>Adolescent Health Care Evaluation Study, 1984-91</u> <u>「思春期ヘルスケア評価研究」</u>	H I V・青少年健康
3	<u>Adolescent Mental Health Study, 1979-1980</u> <u>「青少年の精神的健康に関する研究」</u>	青少年・精神健康
4	<u>Adolescent and Family Development Study, 1978-1982</u> <u>「青少年と家族の発達に関する研究」</u>	青少年・家族関係
5	<u>Aging, Status, and Sense of Control (ASOC), 1995 and 1998</u> <u>「加齢、状況、コントロールの実態：1995年と1998年」</u>	高齢化・健康・生活
6	<u>American Couples, 1978-79</u> <u>「アメリカのカップルに関する研究」</u>	家族関係・仕事
7	<u>American Management Association's Survey of Women Business Owners, 1978-1979</u> <u>「アメリカ経営者協会による女性のビジネスオーナーの調査」</u>	仕事・女性経営者
8	<u>Americans' Changing Lives: Waves, I, II, and III, 1986, 1989, and 1994</u> <u>「アメリカ人の生活の変容」</u>	生活・意識・健康
9	<u>Australian [Adelaide] Longitudinal Study of Aging, Waves 1-5, 1992-1997; 1999-2000</u> <u>「オーストラリア高齢化長期研究 第1～5波、第6波」</u>	高齢者・健康・生活
10	<u>Beginning School Study, 1982-2002</u> <u>「就学開始に関する調査」</u>	子どもの発達・成長・成人
11	<u>Berkeley Social Contexts of Learning Study, 1981-1991</u> <u>「バークレーにおける学習の社会的背景研究」</u>	小学校・バイリンガル
12	<u>British Cohort Study, 1970</u> <u>「イギリス1970年生まれの調査」</u>	出生児・生活・発達
	(省略)	
147	<u>Youth-Parent Socialization Panel Study, 1965-1982: Three Waves Combined; Youth Wave IV, 1997</u> <u>「若者と親の社会化に関するパネル調査 1965-1982 (3波統合)、1997年」</u>	若者・政治意識

▲このページのトップへ
▲ パネル調査メニューへ

LIST 1. 健康・医療・家族生活・家計・労働・教育に関する調査

8. Americans' Changing Lives: Waves, I, II, and III, 1986, 1989, and 1994
「アメリカ人の生活の変容」

分類：生活・意識・健康

<ul style="list-style-type: none"> ● 調査情報出典 (※1) ● ウェブ上のリンク先 ● 調査対象国・地域／調査対象年齢 (※2)／調査継続期間・調査頻度 (※3) 	<p>[ICPSR 6438]</p> <p>http://webapp.icpsr.umich.edu/cocoon/ICPSR-STUDY/06438.xml</p> <p>全米／全年齢・高齢者／3年後／8年後</p>
<ul style="list-style-type: none"> ● 調査の目的・サマリー (調査主体(※4)) 	<p>生産領域での人間関係と参加している活動、健康や通常の活動、生産活動を妨げる予期しない出来事や慢性的ストレスにどう対処しているか、生産活動と、そこでの対人関係・意味づけ・その決定要因・その影響の社会文化的違いを調べる。</p>
<ul style="list-style-type: none"> ● 具体的な調査内容・調査項目 	<p>対人関係 (配偶者・パートナー、子ども、親、友人)、つきあい、余暇、トラウマ的な出来事の実験、退職の捉え方、飲酒喫煙肥満休養の実態、医療サービスの利用、身体的健康、心理的健康、認知的機能、世帯の消費、子どもや孫の数、就業状態、職歴、収入、家族の経済状況、信仰、エスニシティ、人種、教育、性別、居住地。</p>
<ul style="list-style-type: none"> ● 母集団・標本抽出法・サンプル数・調査方法・その他関連情報 	<p>第2波で60歳以上をオーバーサンプル。</p>
<ul style="list-style-type: none"> ● 回収数、脱落等に関する情報 	<p>1波 3917人、2波 2867人、3波 2562人 (1波回答者全員にコンタクト+追加サンプル 164人)</p>

※1: データ所有先における整理番号。ICPSRはInter-university Consortium for Political and Social Researchにおける整理番号。

表内のウェブリンクから、オリジナルの情報へのアクセスが可能。Log#はHarvard MIT Data CenterのHenry Murray Research Archiveにおける整理番号。<http://vdc.hmdc.harvard.edu/VDC/index.jsp>へリンクし、サイト内のSearchのボックスに整理番号 (たとえば00715) を入力することによって、オリジナルの情報へのアクセスが可能 (2006年10月現在)。

※2: 情報がある場合は、オーバーサンプル等によって重点の置かれた年齢も記載。

※3: 情報がある場合のみ。「X年後」は、調査開始年からの年数。対象者の年齢で表示した場合もある。

※4: 個人の研究者の場合は省略。政府機関、大学研究機関等の場合は記載

[▲このページのトップへ](#)
[← LIST 1のメニューへ](#)
[<< パネル調査メニューへ](#)

II. 個別研究報告（分析事例）

6 1990年代以降における第1子出生の傾向と要因： 『第2回 21世紀成年者縦断調査』を用いた分析

福田 節也

1. はじめに

近年の人口学研究においては、夫婦出生力の低下が指摘されている（佐々井 1998, 岩澤 2002, 金子 2004）。1990年代以降、夫婦が結婚してから第1子をもつまでの期間が長期化しており、第2子以降をもたない夫婦が増加していることがその一因である。第1子出生タイミングの遅れならびに第2子出生確率の低下は、晩婚化や女性の高学歴化といった従来の要因では説明できないことが明らかとなっている（佐々井 1998, 金子 2004）。これまでわが国においては、結婚すれば夫婦は平均して2人以上の子どもを生むという状態が持続してきた。しかし、近年における夫婦出生力の低下は、このような2子規範の崩壊を示唆するものである。したがって、このような夫婦の産み方の変化が、どのような構造の下に生起しているのかを明らかにすることは、わが国の今後の出生力変動を見極める上で重要な課題である。

本稿では、厚生労働省大臣官房統計情報部によって実施されている「21世紀成年者縦断調査」の第2回調査のデータを用いて、わが国の1990年代以降における第1子出生の傾向と要因に関する分析を行う。第1子出生がどの年齢で、あるいは結婚からどのくらい経過した時点で生起しているのかは、個人の出生力を規定する上で重要な要素である。また、第1子出生のタイミングは、婚前妊娠や晩婚化といった結婚行動における変化をもっとも敏感に反映しているため、近年における出生パターンの変化を理解するうえで多くの情報を与えるものと思われる。本稿では、晩婚化、高学歴化、雇用の非正規化、若者の失業、そして有配偶女性の雇用労働力化といった変化が、1990年代以降における第1子出生のタイミングにどのような影響を与えているのかについて知見を得る。

2. データ

本稿において用いるデータは、前出の「21世紀成年者縦断調査」の第2回調査（2003年11月5日実施）の個票データである。同調査は、2002年10月末時点で20-34歳であった全国の男女（及びその配偶者）を対象とし、そのうち、第1回調査において協力を得られた者等（及びその配偶者）を調査客体としている（厚生労働省大臣官房統計情報部 2005）。同調査は、同一個人を対象とした追跡調査であり、パネル調査と呼ばれる。パネル調査を利用することの利点として、同一個人の属性の時間的変化を特定できること、また横断調査に比べてより直接的に標本のもつ異質性を統制し、説明要因と非説明変数との因果関係

を考察できることなどが挙げられる（樋口他 2004）。しかし、このようなパネル調査の利点を享受するには、データの十分な蓄積が必須であり、長い時間と調査継続のための莫大な費用を要する。それでは、データが蓄積するまでの間、パネルデータを用いて有効な分析を行うことはできないのであろうか。答えは否である。本稿においては、パネル調査において得ている事歴情報（event history data）を用いた分析事例を示したい。

同調査においては、調査対象者のライフコースに関する様々な事歴情報を得ている。事歴情報とは、結婚、出生、就業などのライフイベントについて、それぞれのイベントを対象者が「いつ」経験したのかを表す情報である。こうした事歴情報は、イベントを経験した年齢や年月によって得られる。このように対象者のイベント経験履歴について情報を集めている調査を遡及的調査（retrospective survey）という。「21世紀成年者縦断調査」の第2回調査においては、対象者の現在の配偶者との同居開始年月、すべての子どもについての出生年月、最終学校卒業あるいは中退の年月、そして過去13回までの就業の開始および終了の年月を得ている。これらの時点に関する情報を生かすことにより、「時間」の概念を反映したモデルを構築し、結婚や出生の規定要因に関する縦断的分析（longitudinal analysis）を行うことが可能である。以下に今回の分析において用いた分析手法であるイベントヒストリー分析（event history analysis）の概略と分析モデルについて述べる¹。

なお、今回の分析では同調査のうち第1回から第3回調査までの個票データ使用することが可能であったが、1) 卒業・中退及び就業に関する履歴情報は第2回調査以降において得られること、そして2) 脱落の発生により、第2回調査のほうが第3回調査よりもサンプル数が多いことの2つの理由により、第2回調査のデータのみを用いた。

3. 分析手法とモデル

3-1. 分析手法

本稿における分析は、イベントヒストリー分析によって行う。イベントヒストリー分析とは、あるイベントの発生パターンとその要因に関する分析手法の総称である。別名、生存分析（survival analysis）ともいわれる。

イベントヒストリー分析では、リスク人口（population at risk）におけるイベント発生確率である「ハザード率（hazard rate）」を分析の対象とする。リスク人口とは、イベントを経験する可能性がある人口を指す。例えば、離婚をイベントとして分析を行う場合、離婚のリスク人口は有配偶の男女であり、未婚者や死別者、離別者はリスク人口に含まれない。ハザード率は、より正確には「時間 t に至るまでの期間に、当該イベントが起こらなかったという条件のもとでの、時間 t におけるイベント発生の瞬間確率（instantaneous rate）」（津谷

¹ なお、Blossfeld and Rowher（2002）は、このような事歴情報を用いた因果分析における注意点として、1) 価値観や態度などの変数を遡及的に扱うことができない、2) 記憶違いによる誤った情報の収集の可能性、3) 生存者のみを対象とすることによるサンプリング・バイアスなどを挙げている。

2002, p. 429 右段, ll. 49-52) を指し、以下のように表わされる。

$$h(t) = \lim_{\Delta t \rightarrow 0} [P(t + \Delta t > T \geq t | T \geq t) / \Delta t]$$

ハザードとは、英語で「危険」を意味する言葉であるが、これはハザード率の概念が死亡を分析対象とすることの多い生物統計において発展したことに由来している。通常、リスク人口におけるイベント発生確率は、イベント発生のリスク開始時点からの「時間」によって異なる。また、イベント発生確率が時間の経過とともにどのようなパターンを示すのかも、対象となる集団・人口によって異なる。イベントヒストリー分析は、このハザード率を時間の関数として特定し、それが単数あるいは複数の説明要因によってどのように変化するのかを明らかにする多変量回帰分析である。モデルのパラメータは、最尤法 (maximum likelihood method) もしくは部分尤度法 (partial likelihood method) によって推定される。時間の関数として表わされるハザード率は、ベースライン・ハザード (baseline hazard) と呼ばれ、モデルの他の要因を統制した場合におけるイベント発生確率の基本的なパターンを表わす。

また、モデルにおける説明変数は共変量 (covariate) と呼ばれる。共変量には、時間によって値が変化する変数と、そうでないものがある。前者を「時間依存性共変量 (time-varying covariate)」といい、年齢や配偶関係、職業、あるいは学歴といった変数がこれにあたる。一方、後者を「時間独立共変量 (time constant covariate)」と呼ぶ。性別や生年月日、出身地などがこれにあたる。時間依存性共変量を用いることができるのは、時間の概念をもつイベントヒストリー分析ならではの利点である。

イベントヒストリー分析において重要な概念にセンサリング (censoring) がある。観察対象となるイベントのリスク期間について、開始時点と終了時点のいずれかが明らかではない場合をセンサリングという。このうち、観察期間中にイベントが生起しないケースを右センサリング (right-censoring) という。イベントヒストリー分析には、右センサリングのケースを偏りなく分析に反映できるという利点がある。例えば、離婚までの期間を従属変数とした重回帰分析を行った場合、離婚経験者しか分析の対象とはならない。しかし、イベントヒストリー分析では、イベントが生起しなかった時点までの情報を分析に反映できるため、リスク人口全体を対象とした分析を行うことができる。

イベントヒストリー分析にはいくつかのモデルがある。本論文において用いるのは、イベントヒストリー分析のうち、もっとも汎用的なモデルとされる比例ハザード・モデル (proportional hazard model) である。イベントヒストリー分析を行うためには、イベント生起の有無のみならず、イベント生起の時点に関する情報が必要となる。「21世紀成年者縦断調査」では、イベントの生起時点について月レベルで回答を得ており、リスク時間の測定単位は連続的と考えてよいほどに細かいといえる。比例ハザード・モデルは、このように

時間の測定単位が連続的である場合に用いられるモデルであり、以下の式によって表わされる。

$$\text{Ln}[h(t)/h_0(t)] = a + b_1X_1 + b_2X_2 + \dots + b_kX_k$$

上式においては、 $h_0(t)$ がベースライン・ハザードであり、 a は定数、 X_k が共変量にあたる。共変量は時間依存性共変量と時間独立性共変量のどちらを含むこともできる。比例ハザード・モデルにおいては、対数化したハザード率の比を従属変数としているため、共変量の影響力はパラメータ b を指数化したハザード比として表わされる。ハザード比とは、共変量の当該カテゴリーにおけるハザード率が基準カテゴリーにおけるハザード率の何倍高いのか、あるいは低いのかを表わす数値である。共変量のハザード比により、人口内のサブ・グループでイベント生起のリスクがどの程度異なるのかについて容易に知ることができる。共変量の影響はリスク期間を通じて一定であり、ベースライン・ハザードを上下に平行的にシフトさせるものと仮定される。そのため、リスク期間のどの時点においても、常にベースライン・ハザードと当該カテゴリーのハザード率の比は一定である。これを比例ハザードの仮定という。現実には共変量 X の影響力が、リスク期間を通じて一定であるというのは強い仮定であるが、モデルではリスク期間と共変量との間に交互作用をとることにより、共変量の影響力に時間依存性を許容するモデルを構築することができる。このことはリスク人口のサブ・グループごとに異なるベースライン・ハザード関数を設定し、それを同一のモデルで推定することに等しい。

なお、比例ハザード・モデルの名称は、ベースライン・ハザードをどのような関数によって表わすのかによってさらに細分化される。本稿における分析では、時間区分線形スプライン・モデル (piecewise liner spline model) を用いる。モデルでは、リスク期間を1年ごとや5年ごとなどの任意の長さによって区切り、各区分におけるハザード率を推定する。ただし、時間区分線形スプライン・モデルでは、区分内におけるハザード率 ($\text{Ln}[h(t)/h_0(t)]$) が一定の傾きで線形に増加あるいは減少するスプライン関数であり、その傾きが区分間で変化すると仮定する。時間区分線形スプライン・モデルでは、同一区分内でも常にハザード率が変化していると仮定するため、より少ないパラメータでハザード率の形状を表わすことができるという特徴がある。ただし、リスク期間はスプライン関数で表わされる「傾き」であるため、共変量との間に交互作用項をとり、共変量の影響力がリスク期間を通じて変化することを仮定するモデルを構築することができない。そのため、本稿における分析では、リスク期間によってサンプルを分けて、リスク期間別の分析を行うことで比例ハザードの仮定を緩めることとする。

以上のように、イベントヒストリー分析では、ハザード率の概念を用いることにより、イベントの生起確率と生起タイミングの両方を考慮した分析を行うことが可能となる。また、多変量回帰分析の手法をもちいることにより、個人の多様なライフコースをモデル化

する柔軟性を備えており、ライフコースにおける移行過程を分析するのに最適な手法であるといえる (Wu 2003)。

3-2. モデル

本稿では、第1子出生の分析を行うのにあたり、2つの多変量モデルを構築した。ひとつは、年齢を基底時間とするモデルであり、もうひとつは結婚期間を基底時間とするモデルである。イベントヒストリー分析においては、ハザード率の推移（イベント発生リスクの変化）を最もよく反映する時間軸（time axis）を選択する必要がある（Allison 1995）。ほぼすべての出生が婚姻内で発生していることを考慮すれば、第1子出生のリスクは結婚によって開始し、結婚期間を通じて変化すると仮定することができよう。このモデルでは、夫婦出生力を直接の対象として、その規定要因を分析することができる。また、有配偶者を分析の対象としているため、夫妻双方の社会経済的属性を説明要因として含めることができるという利点もある。しかしこのモデルにおいては、サンプリング・セレクションの問題が生じる。つまり、結婚したサンプルのみが分析対象となるため、結婚をしない／できないために第1子を生まない女性が分析から排除されてしまうこととなる。未婚化が進行している場合や結婚と出産の意思決定が近接している場合には、夫婦のみを分析対象とすることによって、このようなサンプリング・セレクションの問題が大きく作用するものと思われる。また、近年においては婚前妊娠が増加傾向にあり、結婚を出生のリスク開始イベントとすることの意義が薄れていることも指摘されている（金子 2004）。さらに、「21世紀成年者縦断調査」においては、調査時点の結婚が初婚であるのか再婚であるのかを区別することができない。そのため、結婚期間を時間軸としたモデルでは、現在の結婚が再婚である場合に、第1子出生タイミングにバイアスが生じることとなる²。

一方、年齢を基底時間としたモデルにおいては、夫婦に限定せずに、個人が未婚の状態から第1子を生むまでを対象として分析を行うことができる。モデルでは結婚を所与としないため、単純に第1子の出生が年齢的に遅い（早い）のはどのような属性を持つ人々かを明らかにすることとなる。そのため分析の解釈が容易である。また、初婚・再婚を区別できないことによるバイアスも生じない。しかし、一方でリスク開始時点（本稿では女性が16歳、男性が18歳と定義した）において対象者は未婚であるために、配偶者の情報を用いることができない。また、第1子出生タイミングの遅れが晩婚化によるものなのか、結婚後の出生タイミングの遅れによるものなのかを区別できないために、各要因がどちらに作用しているのかが明確ではない。

このようにそれぞれのモデルには長所と短所がある。そこで、本稿でははじめに年齢を基底時間としたモデル（以後、年齢モデル）によって分析を行い、第1子出生タイミン

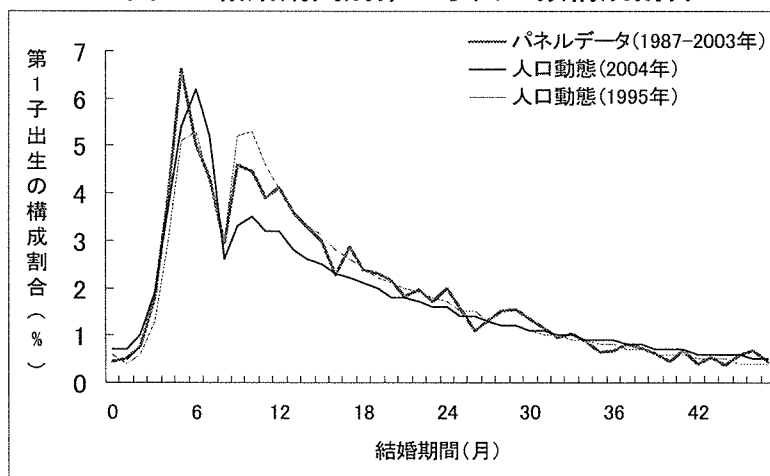
² 第1子出生年月から現在の配偶者との同居開始年月を引くと、出生間隔がマイナスとなるサンプルがあったが、このようなサンプルは分析から除外した。

グが年齢的に遅い（早い）のはどのような要因によるものかを明らかにする。その後結婚期間を基底時間としたモデル（以後、結婚期間モデル）によって、結婚から第1子出生までの期間を対象とした分析を行う。両者の結果を比較することにより、1990年代以降におけるわが国の第1子出生タイミングを年齢と結婚期間の2つの側面から考察し、その変動要因について明らかにする。

4. 第1子出生の傾向

はじめに、「第2回 21世紀成年者縦断調査」における第1子出生の全体的な傾向について把握する。図1は、同調査における結婚期間別第1子出生数の構成割合を1995年と2004年の人口動態統計と比較したものである。「第2回 21世紀成年者縦断調査」では、1987年から2003年までの期間において第1子出生が確認されている。その構成割合の推移は、結婚から5ヶ月目において出生数のピークを迎えている。次いで、第1子出生の生起割合は結婚後9-12ヶ月において第2のピークを迎え、以降においては若干の変動を経ながら低下する傾向にある。全体として同調査における第1子出生のパターンは、2004年の人口動態と1995年の人口動態との中間的な推移傾向を示している。また、同調査では婚前妊娠による出生を表す第1のピークが、人口動態統計よりもやや早くかつ高めにでている傾向があるが、全体として全国レベルの出生傾向をよく反映しているといえよう。

図1 結婚期間別第1子出生数構成割合



資料：厚生労働省統計情報部 2006. 『平成17年度 出生に関する統計の概況：人口動態特殊報告』

図2 性別第1子出生の生存確率

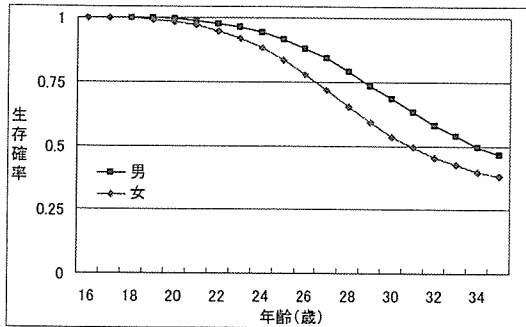
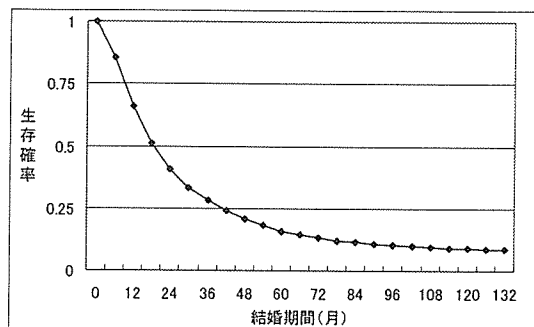


図3 第1子出生の生存確率の推移



次に、 Kaplan-Meier法 (Kaplan-Meier Estimation Method) を用いて、第1子出生の生存確率 (survival probability) の推移を表わしたのが図2ならびに図3である。生存確率とは、時点 $t-1$ から t までの間に、リスク人口においてイベントが生起しなかった確率 $(1-p)$ を累積した値である。図2では、年齢を軸として、第1子出生の経験パターンが性別によってどのように異なるのかを表わしている。第1子出生の生存確率は男女ともに20歳代後半において急激に低下している。35歳時点での第1子出生の経験割合は、男性で53.6%、女性で62%となっている。未婚化の進展を反映して、1990年代以降では男性の半数、女性の4割が35歳までに出生を経験していない。

また、図3では結婚している夫婦を対象に、結婚期間を軸として第1子出生の生起パターンを表わした。その結果、結婚から6ヶ月までに14.3%、1年目までに34.2%、そして2年目までに60%の夫婦が第1子の出生を経験している。結婚後8年6ヶ月の時点で9割以上の夫婦に出生経験があることがみてとれる。したがって、離婚や死別が生起せずに、結婚が持続すると仮定できるのであれば、夫婦の9割以上が少なくとも1子をもつことが期待される。

次節以降では、男女あるいは夫妻のもつ社会経済的属性を説明要因として、第1子出生の生起パターンがどのように異なるのかを検証する。

5. 年齢による第1子出生タイミングの分析

はじめに、年齢を基底時間としたモデルにおける第1子出生タイミングについて考察する。年齢モデルにおける分析では、法定結婚年齢を基準として男性は18歳、女性は16歳より出生のリスクが始まると仮定し、年齢別の第1子出生ハザード率の分析を行う。

5-1. 記述統計

表1は、モデルにおける説明変数の分布を表わしている。分析は男女別に行われ、説明変数として、出生コーホート、学歴、そしてリスク期間における就業状態を用いた。なお、就業状態は就業履歴より作成した時間依存性共変量である。

表1 年齢モデルに用いた共変量の分布

	男			女		
	サンプル数 N	割合 %	第1子出生数 N	サンプル数 N	割合 %	第1子出生数 N
出生年						
1965-69年	1544	15.5	821	1653	15.2	1133
1970-74年	3593	36.0	1293	3954	36.5	2068
1975-79年	3068	30.7	338	3301	30.4	582
1980-84年	1778	17.8	36	1941	17.9	58
合計	9983	100.0	2488	10849	100.0	3841
学歴						
中学・高校	4081	40.9	1333	3733	34.4	1888
専門・高専・短大	2156	21.6	508	4830	44.5	1555
大学・大学院	3746	37.5	647	2286	21.1	398
合計	9983	100.0	2488	10849	100.0	3841
第1子出生時年齢 ^{*1}						
16-19歳	29280	21.5	39	55077	33.1	151
20-24歳	63127	46.3	671	70709	42.5	1341
25-29歳	33944	24.9	1315	32435	19.5	1948
30-34歳	10134	7.4	463	8138	4.9	401
合計	136485	100.0	2488	166359	100.0	3841
就業状態 ^{*1}						
無職	11967	8.8	48	24039	14.5	2388
自営業主・役員・ 家族従業員	3306	2.4	3306	1502	0.9	1502
正社員	28375	20.8	860	35310	21.2	386
非正社員・内職	7829	5.7	54	16386	9.9	111
在学中	35093	25.7	37	56446	33.9	62
不明	49915	36.6	1348	32676	19.6	818
合計	136485	100.0	2488	166359	100.0	3841
配偶関係(第1子出生時点) ^{*2}						
独身	6771	67.8	129	6390	58.9	362
有配偶	2783	27.9	2127	3768	34.7	2949
不明	429	4.3	232	691	6.4	530
合計	9983	100.0	2488	10849	100.0	3841

*1: リスク期間中に値が変化する時間依存共変量。サンプル数はパーソン-ピリオド数をもとに算出。

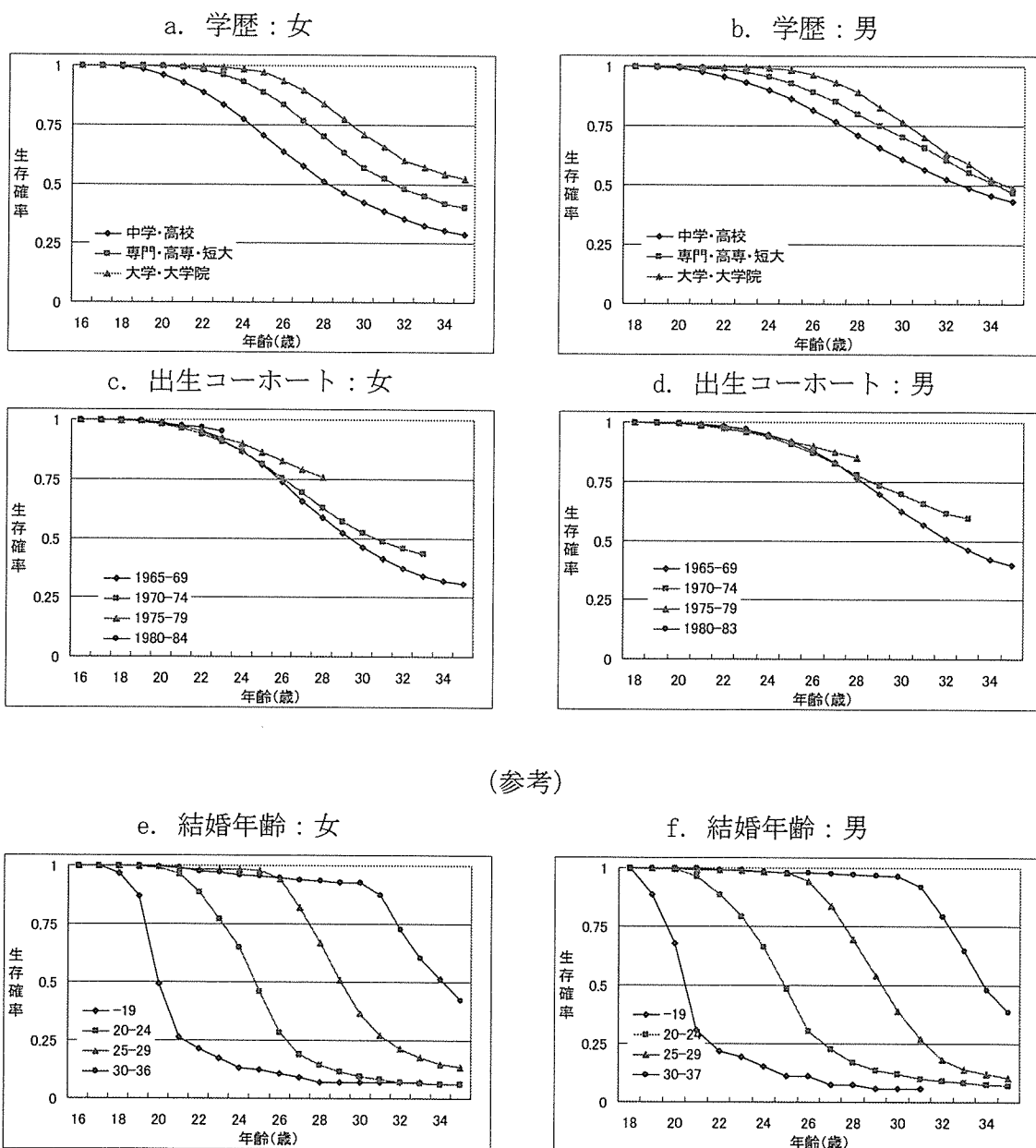
*2: 調査時点の配偶関係を表す。そのため、「独身」であっても前の結婚において有配偶である可能性もある。

男性は 9983 サンプル、女性は 10849 サンプルが分析の対象となった。1965 年から 1984 年に生まれた男女が分析対象となっているが、出生経験がある男女はおもに 1970 年代後半までの男女であることがわかる。学歴は、中退を含む最終学歴として測定されている。女性は男性よりも高卒以上の学歴構成割合が高いが、大学・大学院への進学割合は男性のほうが高い。年齢についてみると、男性においては 20 歳代後半、女性では 20 歳代前半より、第 1 子出生数が多い傾向がある。就業状態については、過去の就業の開始・終了年月および就業形態のいずれかに欠損値があると、就業状態が特定できないため、「不明」の割合が、特に男性において高い傾向がある。これを考慮に入れずに比較すると、女性は男性よりも在学中や非正規就業に就いている割合が高い傾向にある。在学中であるか否かは、最後に行った学校を卒業もしくは中退した年月より特定した。なお、無職のカテゴリーには、学校卒業後に無職である男女が分類されている。

参考までに、現在の配偶者との同居開始年月を元に、配偶関係の分布も示した。独身

で第1子出生を経験している割合は、男性が2.0%、女性が5.7%であった。したがって、ほぼすべての出生が婚姻内で生起していることが確認された。ただし、ここで分類した配偶関係は、調査時点の配偶関係に基づくため、「独身」と分類されていても実際には前の配偶者と婚姻関係にある可能性もある。そのため、非嫡出出生の割合は、ここで示された値よりも小さいものと思われる。

図4 カプラン-マイヤー法による第1子出生の推移：年齢モデル



(参考)

5-2. 2 変量解析

次に、前出のカプラン-マイヤー法によって、時間独立性共変量と第1子出生との関係についてみてみよう。図4は、性別に学歴や出生コホートによって第1子出生の年齢別生起パターンがどのように異なるのかを表わしている。また、ここでは参考として、結婚年齢による出生パターンについても表わした。これについては、調査時点において結婚している男女に対象が限定されていることを付記しておく。

図4によると、女性は学歴が高いほど、第1子の出生タイミングが遅く、また34歳時点における出生割合も低い傾向にあることが明らかである。一方、男性においては、30歳代前半において、大学・大学院進学者によるキャッチ・アップがみられる。出生コホートについては、男女ともに近年のコホートほど20歳代後半における出生タイミングが遅れる傾向がある。

結婚年齢別のグラフ(図4-e, 図4-f)をみると、結婚年齢が遅くなるにしたがって、第1子出生のタイミングも遅れることが明らかである。結婚年齢が20歳以降における出生パターンは、リスク開始時点が後方にずれることを除けば、生存確率の低下傾向にはそれほど大きな差はみられない。しかし、結婚年齢20歳未満の層においては、第1子出生が若年齢において急激に経験されていることが明らかである。このことは婚前妊娠が、早婚において生起しやすいという人口動態統計における知見とも一致する(厚生労働省2006)。また、結婚年齢が30歳以上の男女においては、結婚年齢以前より第1子出生が経験される傾向がある。これについては、先にも述べたように、非嫡出出生の発生というよりは、前回婚における出生と理解することができる。

5-3. 多変量解析

最後に、各共変量を用いて、イベントヒストリー分析の一種である時間区分線形スプラインモデルによる多変量解析を行った。分析では、年齢を基底時間として、出生コホート、学歴、そして就業状態を統制した。なお、各共変量は、年齢によって異なる影響をもつことが予測されたため、リスク期間をいくつかの区間に分けて分析を行った。

表2は、女性サンプルを対象とした時間区分線形スプライン・モデルの分析結果を表わしている。分析結果は、すべてハザード比(hazard ratio)によって表わした。ハザード比の値は、当該カテゴリーにおける第1子出生のハザード率が、基準カテゴリーに比べて何倍高いのか(あるいは低いのか)を表わしている。モデル1を例にとると、1965-69年に生まれた女性は、1970-74年に生まれた女性に比べて1.1倍、第1子出生のリスクが高いと解釈される。また、ハザード比は掛け算のモデルであるため、1975-79年に生まれた大卒女性のハザード比は、 $0.35 (=0.66 \times 0.53)$ となる。この場合の基準カテゴリーは1970-74年に生まれた中学・高校進学女性となる。なお、年齢スプラインのハザード比の解釈には若干の注意を要する。なぜならば、他の共変量と異なり、スプ

ライン関数は、当該リスク期間におけるハザード率の上昇比率を表わしているためである。例えば、モデル1では16歳から24歳の期間において、ハザード率（定数の値）が1歳ごとに1.3倍ずつ上昇することを意味する。同様に24歳から30歳においては、24歳時点におけるハザード率（ $4.40E-06 \times 1.38$ ）から1歳毎に1.03倍ずつ上昇し、30歳から34歳においては、30歳時点におけるハザード率（ $4.40E-06 \times 1.39 \times 1.034$ ）から1歳ごとに0.74倍ずつ低下する。こうして描かれるハザード率の推移がベースライン・ハザードであり、各共変量のハザード比はベースライン・ハザードを平行に上下させるものと解釈される（Lillard and Panis 2003）。

表2 時間区分線形スプライン・モデルによる
第1子出生の年齢別ハザード率の規定要因に関する分析結果：女

女		Model 1	Model 2	Model 3	Model 4
		16-34歳 Hazard Ratio exp(b)	16-24歳 Hazard Ratio exp(b)	25-29歳 Hazard Ratio exp(b)	30-34歳 Hazard Ratio exp(b)
年齢スプライン	16-24歳	1.30 ***	1.31 ***	-	-
	24-30歳	1.03 ***	-	1.02	-
	30-34歳	0.74 ***	-	-	0.78 ***
出生コーホート	1965-69年	1.10 **	0.91	1.17 ***	1.40 ***
	1970-74年	1	1	1	1
	1975-79年	0.66 ***	0.76 ***	0.53 ***	-
	1980-84年	0.44 ***	0.45 ***	-	-
学歴	中学・高校	1	1	1	1
	専門・高専・短大	0.70 ***	0.44 ***	0.94	1.01
	大学・大学院	0.53 ***	0.18 ***	0.70 ***	1.07
就業状態	無職	15.84 ***	23.26 ***	11.56 ***	10.55 ***
	自営業主・役員・家従	4.05 ***	5.82 ***	3.17 ***	2.07 *
	正社員	1	1	1	1
	非正社員・内職	0.71 ***	0.81	0.59 ***	0.58 *
	在学中	0.35 ***	0.55 ***	0.82	0.00
	不明	2.32 ***	3.11 ***	1.81 ***	1.80 ***
定数	定数	4.40E-06 ***	3.42E-06 ***	2.78E-03 ***	2.26E-03 ***
	Person-period数	166359	125786	32435	8138
	サンプル数	10849	10849	6845	2303
	イベント数	3841	1492	1948	401
	log-likelihood	5659.66	1627.87	3536.52	639.60
	chi-square	10027.26	4827.57	2563.55	488.70
	d.f.	14	12	11	10

* p<.10; ** p<.05; *** p<.01

表3 時間区分線形スプライン・モデルによる
第1子出生の年齢別ハザード率の規定要因に関する分析結果：男

男		Model 1	Model 2	Model 3	Model 4
		18-34歳 Hazard Ratio exp(b)	18-26歳 Hazard Ratio exp(b)	26-29歳 Hazard Ratio exp(b)	30-34歳 Hazard Ratio exp(b)
年齢スプライン					
	18-26歳	1.31 ***	1.34 ***	-	-
	26-30歳	1.07 ***	-	1.01	-
	30-34歳	0.87 ***	-	-	0.87 ***
出生コーホート					
	1965-69年	1.15 ***	0.90	1.37 ***	1.37 ***
	1970-74年	1	1	1	1
	1975-79年	0.79 ***	0.82 ***	0.46 ***	-
	1980-83年	0.77	0.74 *	-	-
学歴					
	中学・高校	1	1	1	1
	専門・高専・短大	0.75 ***	0.60 ***	0.89	1.16
	大学・大学院	0.63 ***	0.32 ***	0.83 **	1.27 **
就業状態					
	無職	0.29 ***	0.38 ***	0.26 ***	0.13 ***
	自営業主・役員・家従	1.11	1.54 ***	0.86	0.80
	正社員	1	1	1	1
	非正社員・内職	0.29 ***	0.30 ***	0.20 ***	0.39 ***
	在学中	0.12 ***	0.16 ***	0.19 ***	0.17 *
	不明	0.87 ***	0.80 ***	0.90	0.96
定数		6.05E-06 ***	4.50E-06 ***	6.47E-03 ***	5.32E-03 ***
	Person-period数	136485	109558	16793	10134
	サンプル数	9983	9983	5257	2910
	イベント数	2488	1233	792	463
	log-likelihood	1196.99	107.42	750.66	436.13
	chi-square	2669.96	1553.78	140.94	77.31
	d.f.	14	12	11	10

* p<.10; ** p<.05; *** p<.01

はじめに、全リスク期間を対象としたモデル1から解釈を行うと、近年の出生コーホートほど、また高学歴女性ほど第1子出生のリスクが低くなる傾向がみられる。また、就業状態についてみると、正社員の女性と比べて、無職や自営の女性の出生リスクが高く、非正社員や学生において低い傾向がある。無職の女性の出生リスクが高いのは、女性が出産前に退職しているためと思われる。また、自営や役員、家族従業の女性の出生リスクが高いのも、就業を中断することなく出産できるためであろう。したがって、ここでは出産時における就業継続の容易さが、第1子出生リスクの差となって表れているといえよう。しかし、出産の意思決定が就業状態によってどのような影響を受けているのかを明らかにしようとするのであれば、妊娠が判明する前である7ヶ月前の就業状態を説明変数として用いるべきである。これについては次回以降の課題とする。

16歳から24歳までの若年齢における第1子出生は、1970年代前半のコーホートにおいて最も顕著にみられる。また、高卒以上の学歴におけるハザード比が著しく低いことから、