

# 年齢別出生率のセミパラメトリックモデル

## Modeling Fertility Age Schedule: A Semiparametric Approach

余田翔平<sup>1</sup>・岩澤美帆<sup>1</sup>・石井太<sup>2</sup>

(1. 国立社会保障・人口問題研究所, 2. 慶応義塾大学)

Shohei Yoda<sup>1</sup>, Miho Iwasawa<sup>1</sup> and Futoshi Ishii<sup>2</sup>

(1. National Institute of Population and Social Security Research, 2. Keio University)

yoda-shouhei@ipss.go.jp

### 1. 目的

年齢別出生率の数理モデルは人口学の中でも最も歴史の古い領域のひとつと言える。特に、比較的少数のパラメータによって出生率の年齢スケジュールを再現する試みは国内外問わず盛んに取り組まれてきた。これらのパラメトリックモデルにはいくつかの利点がある。第1に、あるコホートが出生過程を完結していなくても、ある程度の年齢別出生率が観察されていれば（すなわち、データが利用可能な時点で例えば30代後半に差し掛かっていれば）、パラメータを推定できる。この特徴は、とりわけ出生率の将来推計において大きな意味を持つ。第2に、すでに考案されてきたいくつかのパラメトリックモデルにおいては、パラメータが人口学的な意味を兼ね備えている。例えば、Kaneko (2001) によって考案された一般化コール・マクニールモデルは4つのパラメータを持つが、それぞれ出生率の水準 (level)、出生年齢の最頻値 (mode)、出生タイミングのバラつき (variance)、出生率の年齢スケジュールの (非) 対称性 (shape) を示す。これらのパラメータの時系列的变化を分析することで、出生率の年齢スケジュールについて人口学的理解を深めることができる。

年齢別出生率のパラメトリックモデルにはこうした利点がある一方、実際に観察された年齢別出生率への当てはめという観点からは、とりわけ近年問題が生じつつある。図1は、1970年コホートと1980年のコホートの第1子年齢別出生率を示している。1970年コホートでは典型的な単峰形のベルカーブが描かれている一方、1980年コホートの第1

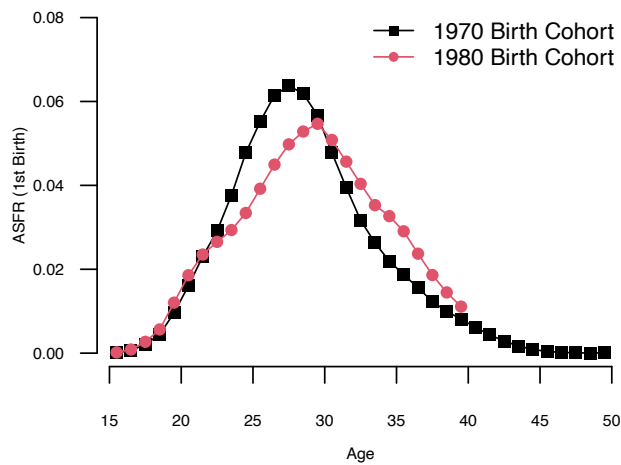


図1 コホート第1子年齢別出生率

子出生率の年齢スケジュールは大きく変化していることがうかがえる。第1に、10代後半から20代前半にかけて出生率カーブに「こぶ」(hump)が見られる。第2に、1970年コホートでは30歳以降は出生率が概ね一定の減少率で減少するのに対して、1980年コホートでは35歳前後において加齢に伴う出生率の低下が一度緩やかになり、その後再度低下していく、という傾向が観察される。

従来のパラメトリックモデルでは捉えることが難しい、こうした近年の年齢別出生率の変化に対して、本報告ではセミパラメトリックなアプローチを提案す

る。その方法とは、出生率の年齢スケジュールのベースラインは観測データからノンパラメトリックに定め、そのベースラインからの乖離をパラメータ化する、というものである。

## 2. 方法

$\mathbf{H} = \{h_{xc}\}$ をコーホート年齢別出生ハザードを示す $A \times C$ 行列とする。 $A$ は年齢階級数、 $C$ は出生コーホート数を表す。

$$\log h_{xc} = \alpha_c + \beta_x + g_{xc}$$

ただし、

$$\alpha_c = \log \sum_x h_{xc}, \quad \beta_x = \frac{1}{C} \sum_c (\log h_{xc} - \alpha_c)$$

特異値分解により、 $\mathbf{G} = \{g_{xc}\}$ は以下のように表現できる

$$\mathbf{G} = \mathbf{USV}^T = \sum_{i=1}^{\rho} s_i \mathbf{u}_i \mathbf{v}_i^T$$

ここで、 $\mathbf{U}$ は左特異ベクトルからなる行列、 $\mathbf{S}$ は特異値を対角成分に持つ対角行列、 $\mathbf{V}$ は右特異ベクトルからなる行列である。 $\mathbf{G}$ から任意の出生コーホートの列ベクトル $\mathbf{g}_c$ を取り出すと、

$$\mathbf{g}_c = \sum_{i=1}^{\rho} s_i v_{ci} \mathbf{u}_i$$

と表現できる。 $v_{ci}$ は $\mathbf{v}_i$ の第 $c$ 成分である。年齢別死亡率の将来推計モデルとして広く活用されている Lee-Carter model では、ピリオドの年齢別死亡率に対して同様の特異値分解を適用し、第 1 特異値および対応する特異ベクトルのみを用いて年齢別死亡率が近似されることが一般的である。これに対し、本報告では第 3 特異値までに対応する項を用いて、コーホート年齢別出生ハザードを近似する（モデリングは出生順位別に行う）。

なお、Lee-Carter model を出生分析に応用した例として Myrskylä et al. (2013) や 2013 年のカナダの公的人口推計などが存在する。ただし、これらの既存研究においてモデリングの対象となっているのは、ピリオドの年齢別出生率であり、さらにその年齢別出生率は分母人口を限定しない率（unconditional rate）である。しかし、晩産化をはじめ、出生行動は年齢パターンの変化が生じやすく、その場合、ピリオドの年齢別出生率はテンポ効果によって歪められることはよく知られている。さらに、分母人口を限定しない年齢別出生率の変動には、当該年次の出生行動の影響のみならず、過去の出生行動の帰結であるリスク人口の影響の両方が反映される（岩澤・金子 2013）。それゆえ、出生の行動要因の変動をモデリングする上では、コーホート年齢別出生ハザードを対象とすることが望ましい。

## 3. データ

各年次に日本人女性から発生した出生数を分子（人口動態統計の調査票情報にもとづく）、日本版死亡データベース（JMD）の生存延べ年数を分母として、各年次の年齢別出生率を算出した。これらピリオドの年齢別出生率をコーホート指標に組み替えたいうえで、リスク人口を未経験者に限定した率、すなわちコーホート年齢別出生ハザードに変換した。

## 4. 結果

試験的分析の結果、第 3 特異値までを用いることで 1980 年コーホートの出生順位別年齢別出生率は概ね再現された。報告当日は既存のパラメトリックモデルとの比較も行う。