

R で学ぶ人口分析  
(Learning demographic analysis with R)

中澤 港

26 June 2017



# 目次

第1章	人口分析の目的と戦略	11
1.1	人口そのものを分析する	11
1.2	人口とそれ以外の現象の関連をみる	12
第2章	人口分析に使うデータ	13
2.1	日本の国勢調査	13
2.2	日本の人口動態統計	16
2.3	死因について	17
2.4	その他のデータソース	20
第3章	人口構造の分析	21
3.1	国勢調査データから人口ピラミッドを描く	22
3.2	手作りのセンサス生データから人口構造を扱う	34
3.3	人口構造の指標値	36
3.4	年齢報告の不正確さについての指標	40
3.5	ふくらみ指数と年少人口構造係数の関係	42
3.6	介護者比	48
第4章	死亡の分析	49
4.1	はじめに	49
4.2	ヒトの死とは？	49
4.3	死亡データの成立	50
4.4	死亡のプロセス	53
4.5	死亡指標	56
4.6	死亡の特性	65
4.7	死因の分解	65
4.8	健康余命	66

---

4.9	Age-Period-Cohort モデル . . . . .	68
4.10	死亡の数学モデル . . . . .	69
4.11	死亡データ分析の例 . . . . .	72
4.12	文献 . . . . .	80
第 5 章	出生の分析 . . . . .	83
5.1	出生というライフイベントの多面性 . . . . .	84
5.2	出生の近成要因 (proximate determinants) . . . . .	113
5.3	出生の数学モデル . . . . .	115
5.4	出生調査の方法論 . . . . .	121
5.5	出生分析の実際 . . . . .	124
5.6	文献 . . . . .	126
第 6 章	結婚と離婚 . . . . .	129
6.1	結婚のモデル . . . . .	129
第 7 章	人口移動と人口分布 . . . . .	131
第 8 章	人口推計とシミュレーション . . . . .	133
8.1	人口推計とは何か? . . . . .	133
8.2	コウホート要因法 . . . . .	134
8.3	システムダイナミクスモデル . . . . .	137
8.4	人口学におけるシミュレーションの意義 . . . . .	138
8.5	人口再生産シミュレーション . . . . .	139
第 9 章	二次資料からマクロな関連をみる . . . . .	149
9.1	R の基本情報とインストールの方法 . . . . .	153
9.2	人口分析に関連したパッケージ . . . . .	153
索引		154

2011年8月14日 第0.1版  
2011年9月1日 第0.2版  
2012年1月8日 第0.3版  
2012年1月25日 第0.4版  
2012年11月30日 第0.5版  
2013年2月12日 第0.6版  
2013年6月29日 第0.7版  
2014年9月22日 第0.8版  
2014年11月28日 第0.8.1版  
2015年8月29日 第0.8.2版  
2016年11月21日 第0.8.3版  
2017年5月29日 第0.8.4版  
2017年5月30日 第0.8.5版  
2017年6月6日 第0.8.6版  
2017年6月26日 第0.8.7版

中澤 港 (著)

神戸大学大学院保健学研究科

国際保健学領域 教授

<minato-nakazawa@umin.net>



## はじめに

2012年1月1日付で発表された、平成23(2011)年人口動態統計の年間推計<sup>\*1</sup>によると、2011年の日本における日本人の人口は20万人以上減少した。人口動態の種類別にみると、出生は最近横這い、あるいはやや上昇傾向だったのが大きく減少し、死亡はそもそも増加傾向だったけれども、その傾向以上に大幅に増加したことがわかる<sup>\*2</sup>。

発表資料を良く読むと、10月までの率に基づいてその後の率を推計しているために年間死亡が過大評価になっている可能性と、一部収集できていない調査票があるためにすべての値が過小評価になっている可能性があることに注意しなければならないと書かれているが、それを割り引いて考えても、死亡が前年確定値より6万人以上も増えたという結果は、東日本大震災がもたらした爪痕の大きさをはっきりと示している。実際、2012年7月26日に公表された、平成23(2011)年の簡易生命表の「1. 主な年齢の平均余命」<sup>\*3</sup>を見ると、平成22(2010)年の簡易生命表では男性79.55年、女性86.30年まで伸びた平均寿命(ゼロ歳平均余命)が、男性79.44年、女性85.90年へと男女とも短くなり、このことへの寄与が最も大きかった死因は(東日本大震災を意味する)地震であった。

このように、大災害や、致命的な感染症の流行は、人口の増減という形で直接、我々の生きる社会に影響を及ぼす。また、現在政府が進めている「税と社会保障の一体改革」は、今日の超高齢社会に適した制度構築ができていないことの帰結であり、1980年代から予想を超えた激減が続いた合計出生率をみれば、超高齢社会が訪れることは明白だったのに、対策を怠ってきたツケが回ってきたといえる。干拓による土地開墾やダム建設による利水をしたけれども利用者が見込みに満たないのは、煎じつめれば人口予測が甘かったためである。人口減少社会で需要の低下を免れることは非常に困難であり、経済活動も停滞して不景気になるのは、海外の市場への依存を大きく伸ばす可能性を除けば、ほぼ自明であった。それなのに、1970年代までに建てられた計画を進めてしまったのは、現在進行形で続いている愚行といえる。

---

<sup>\*1</sup> <http://www.mhlw.go.jp/toukei/saikin/hw/jinkou/suikei11/index.html>

<sup>\*2</sup> もちろん、健康水準が悪化したわけではなく、高齢化が進行して死亡率が高い高齢者が相対的に増えたためである。

<sup>\*3</sup> <http://www.mhlw.go.jp/toukei/saikin/hw/life/life11/dl/life11-02.pdf>

こうした愚行に歯止めを掛けるために、人口の分析方法やその結果の解釈の仕方を知っておくことは、現代の市民社会を生きる誰にとっても、非常に重要なはずである。しかし、上記の説明だけで、分析方法や結果を解釈できる人が、世の中にどれだけいるかと考えると、非常に心許ない。世界の人口学をリードしてきたトップ研究者の一人である Samuel H. Preston ら 3 人の著者による “Demography: Measuring and Modeling Population Processes” という名著が Blackwell Publishing から出ているので、専門家や行政担当者は頑張ってそれを勉強すれば良いが、一般向けに書かれた人口分析のテキストは多くないし、実際にソフトウェアを使って計算したり作図したりできるところまで書かれた日本語の入門書は、現在に至るまで、和田光平 (2006) 『Excel で学ぶ人口統計学』(オーム社) がほぼ唯一のものである\*<sup>4</sup>。同書はコンピュータの画面キャプチャも数多く使われ、非常に分かりやすく丁寧に書かれていて入門書としてはお薦めできるが、その分、説明されている分析方法がそれほど多くないし、人口以外の情報と関連付けて分析することについての説明がほとんどない。

このことは、日本の大学での人口学の扱いが不当に少ないことにも起因している。人口学を専門に扱う大学の附属研究所は日本大学にしかないし、国立研究機関も社会保障・人口問題研究所しか存在しない。人口学の講義は主に経済学部で行われているため、生物人口学などの内容はほとんど扱われていない。医学系の公衆衛生学や保健師養成課程でも、医療法に定められた医療計画の策定などを合理的に行うためには人口学の知識が必要となるにもかかわらず、人口学関連の講義は（あったとしても）1 時間か 2 時間に過ぎないのが普通である\*<sup>5</sup>。米国の大学院で、人類学と人口学の double degree がとれるコースが普通に設置されていたり、セメスター単位で数多くの人口学関連の講義が行われているのに比べると、寂しい状況と言わざるを得ない。

本書は、こうした状況を少しでも改善するために、フリーソフトとして開発され、誰でも自由に使えるソフトウェアである R\*<sup>6</sup>を用い、人口分析を学ぶための入門から中級のテキストとして使われることを目指して執筆するものである。人口分析の方法を学ぶためには適切なサンプルデータを使う必要があり、前出の和田 (2006) では、政府が発表している人口関連のデータをダウンロードして利用する方法が解説されているが、余計な情報を削除するために様々な前処理が必要である。最近では e-Stat\*<sup>7</sup>から多くのデータをダウン

\*<sup>4</sup> 英語ならば、Rowland DT (2003) Demographic methods and concepts. Oxford Univ. Press, ISBN 978-0-19-875263-9 も素晴らしい入門書であり、Excel のワークシートが CD で付属しているのでお薦めできる。

\*<sup>5</sup> 神戸大学大学院保健学研究科において、2013 年度から人口学の 2 単位の講義を開始した。Newell C (1988) Methods and models in demography. Guilford Press, New York. を用いて英語で行っている。

\*<sup>6</sup> <http://www.r-project.org/>

\*<sup>7</sup> <http://www.e-stat.go.jp/SG1/estat/eStatTopPortal.do>

ロードでき、その一部は API 機能<sup>\*8</sup>を使って XML や JSON などの属性付きでダウンロードできるようになったが、さまざまなデータをダウンロードして加工するのはやや面倒である。ある程度のサンプルデータが付属しているソフトが使える方が、初学者には親切だろう。筆者が以前書いた『R による保健医療データ解析演習』のために使った関数群を **fmsb** というパッケージにして CRAN<sup>\*9</sup>にアップロードした際に、いくつかの基本的な日本の人口統計データを R ですぐ使える形で含めたので、本書では、**fmsb** パッケージをインストールしてロードすればすぐに使える、それらのサンプルデータを分析する形で解説した。いちいちデータを入力したりダウンロードするという作業をしなくても分析方法を学べて便利であろう。本書が日本において人口学の講義が1つでも増える一助となれば幸いである<sup>\*10</sup>。なお、R に触れるのが初めてという方のために、基本的な使い方を付録に書いておくので、適宜参照されたい。

---

<sup>\*8</sup> <http://www.e-stat.go.jp/api/> として 2014 年 10 月 31 日から本格運用開始。

<sup>\*9</sup> <http://cran.r-project.org/>

<sup>\*10</sup> <http://minato.sip21c.org/ldaR/> で本書で用いる R コードやデータを公開するので、自習することも不可能ではないと思う。



## 第 1 章

# 人口分析の目的と戦略

人口分析の目的は大別すると2つある。1つは人口そのものの分析で、もう1つは人口とそれ以外の現象の関連を見ることである。前者、すなわち人口そのものの分析をターゲットにする研究を、形式人口学とか人口統計学と呼ぶ。英語では **Demography** となる\*<sup>1</sup>。後者、すなわち人口とそれ以外の現象の関連を見ることを目指す研究は、広く言えば実体人口学であり、英語では **Population Studies** と呼ばれる。数理人口学は、人口現象に法則性を見いだすという意味では前者に含まれるが、感染症の数理モデルによる分析を考えると後者にも含まれる。

### 1.1 人口そのものを分析する

人口には構造がある。一時点での人口構造を把握するタイプの人口統計を、人口静態統計という。国ごとに全住民の居住状況を把握する悉皆調査であるセンサス（国勢調査）は、代表的な人口静態統計であり、すべての人口分析の基本である。国連は西暦末尾が0となる年の前後にセンサスを実施することを加盟国に求めているが、できていない国も少なくない。

センサスで得られる基本構造は、性・年齢別人口であるが、戸別の世帯訪問による調査を原則とすることから、世帯単位の情報も得られるのが普通である。日本では職業や通勤・通学の情報なども収集しているが、プライバシーを盾に取った回答拒否も増えてきたので、郵送による回答を可能にするなど、回収率を上げる工夫がなされている。性・年齢以外にも、就業状態別とか職業別といった属性別の人口構造は、すべて人口静態統計に区分される。

一方、人口学的なイベントの発生を記録するタイプの人口統計を、人口動態統計という。

---

\*<sup>1</sup> 館稔は、かつて「民誌学」という訳語を提唱したが、定着しなかった。

言い換えると、人口動態統計は、出生、死亡、移入、移出といった人口の変化を記録することで行われる。通常、結婚や離婚も、人口動態イベントとして扱われる。人口動態イベントの単純な集計結果から、人口学的基本方程式と呼ばれる、次の関係式が成り立つ。

$$\text{人口増加} = \text{出生} - \text{死亡} + \text{移入} - \text{移出}$$

この関係を応用して、ある時点の性・年齢別人口構造と、それ以降の性・年齢別人口動態の予測値から将来人口を推計する方法を、コウホート要因法と呼び、将来人口推計の方法として広く使われている。詳細は後述するが、人口動態に関わる要因には、生物学的要因、社会経済的要因、文化的要因（宗教的な影響も小さくない）など様々なものがあって、人口動態を予測することは容易ではない。

## 1.2 人口とそれ以外の現象の関連をみる

実体人口学は、形式人口学によって見いだされる人口現象の法則性が、どのようなメカニズムによってもたらされているのかを解釈することに寄与する。例えば、人口指標とマクロな保健医療指標や社会経済指標との関連を見る（散布図、重回帰、ロジスティック回帰分析、ハザード解析などによって）分析は、典型的な実体人口学研究の例といえる。

実体人口学の研究は、人口指標との関連をみる諸指標が属するカテゴリを扱う個別科学の研究ともみなすことができる。例えば、世帯所得や家族構成、居住地、学歴などの社会経済因子が既往出生児数とどのような関連をもっているかという研究は、経済学や社会学の研究として多数行われている。

## 第2章

# 人口分析に使うデータ

人口分析に使うデータは、大別すると、以下4つに分かれる。

- 政府や国際機関、あるいは調査機関が調査して集計後発表している資料。国勢調査や人口動態統計が代表的。
- 上記資料の集計前の個票データ。既存のものでは DHS が有名だが、今後増加が見込まれる。日本の国勢調査の個票はプライバシー保護などの理由で、一般の研究者は利用できないので、総務省統計局に対してカスタム集計を依頼するしかない。人口動態統計の基礎資料である出生小票や死亡小票は、厚生労働省に目的外使用申請をして許可されれば電子化したものを利用できる。
- 研究者が目的を定めてデザインして実施する調査によって得るデータ。サーベイデータとも呼ばれる。
- まだあまり活用例はないが、例えば twitter などの SNS の発信位置情報をトレースすることで移動に関する情報を得るなど、機械的に（自動的に）保存される膨大なデータ、いわゆるビッグデータも人口分析に活用できる可能性がある。

### 2.1 日本の国勢調査

国勢調査は代表的な人口静態統計データである。

- 統計法第4条で定められた義務調査（総務省所管）
- 個人情報保護法の適用除外
- 人口・世帯の実態を把握し、各種行政施策の基本資料を得る目的で実施され、以下のような役に立つ。
  - － 衆議院の選挙区割り決定の元

- 地方交付税，補助金などの算定基礎
- 行政需要（介護需要など）の把握に必須
- 電力需要予測など民間にも役立つ
- すべての標本調査の母集団情報
- 調査員は非常勤の国家公務員で，各調査区で概ね 50 世帯を担当
- 5 年毎に実施。
- 日本に常住している（3 ヶ月以上住んでいる）者全員を対象とする悉皆調査（全数調査）
- 日本の国勢調査人口は 10 月 1 日時点の人口

総務省の国勢調査のページ (<http://www.stat.go.jp/data/kokusei/2015/index.htm>) から概要を入手できるほか，データについては，e-Stat の国勢調査一覧ページ (<https://www.e-stat.go.jp/SG1/estat/GL02100104.do?tocd=00200521>) から年度別の詳細な集計値と時系列データを入手できる。また，日本の男女年齢別人口は，R の `fmsb` パッケージにも，`Jpop` 及び `Jpopl` というデータフレームとして入れてある\*1。

### 2.1.1 2005 年国勢調査

**調査事項 (1)**（世帯員に関する事項）(1) 氏名 (2) 男女の別 (3) 出生の年月 (4) 世帯主との続き柄 (5) 配偶の関係 (6) 国籍 (7) 就業状態 (8) 就業時間 (9) 所属の事業所の名称及び事業の種類 (10) 仕事の種類 (11) 従業上の地位 (12) 従業地又は通学地

**調査事項 (2)**（世帯に関する事項）(1) 世帯の種類 (2) 世帯員の数 (3) 住居の種類 (4) 住宅の床面積 (5) 住宅の建て方

**報告形式** 要計表による人口集計，抽出速報集計，第 1 次基本集計，第 2 次基本集計，第 3 次基本集計，抽出詳細集計，従業地・通学地集計，小地域集計

**問題点** 期間内の未回収率が前回の 1.7% から 4.4%（210 万世帯）へ激増

### 2.1.2 2010 年国勢調査

- 日本では，大規模調査となるので，2005 年の項目に加えて，現住居での居住期間・5 年前の住居の所在地・教育・利用交通機関も調査された。

\*1 どちらも 1888 年の国勢調査開始時から現在に至る男女別の年齢各歳別人口を変数として保持しているが，両者の違いは，年齢区分の最上位となるオープンエンドな年齢層が，`Jpop` では 85 歳以上を意味する“85+”であるのに対して，`Jpopl` では 110 歳以上を意味する“110+”となっていることと，`M2015J` のように日本人人口を示す変数を，`Jpop` では 2015 年しか含めていないのに対して，`Jpopl` では 2000 年から含めていることである。

- 600 億円と 70 万人の調査員（非常勤の国家公務員）を投入して実施された。
- 総務省統計局サイト (<http://www.stat.go.jp/data/kokusei/2010/>) での告知はあったが地味であった。
- 米国 Census Bureau での特設サイトは大変工夫されており、「10 分で OK！ 10 問の簡単な質問に答えてポストに投函！」とラッパーに歌わせるプロモーションビデオを作って YouTube で流すことにより、回答をサボりがちな若者に訴求していた。
- 国によってやり方に違いがあり、米国は郵送が基本である。日本は 2010 年から郵送も可となり東京ではネット調査も実施したが、対面集票が基本であった。

### 2.1.3 2015 年国勢調査

- 総務省統計局サイト (<http://www.stat.go.jp/data/kokusei/2015/gaiyou.htm>) に概要がまとめられている。2010 年東京での試行を踏まえ、全国でネット回答が導入されたのが特徴であった。原則は 10 月 1 日現在の情報について、9 月中旬にネット回答してもらい（調査員が各戸ごとに ID とパスワードが書かれた説明書を配布）、回答がなかった世帯は紙の調査票を調査員が配布し記入後集票あるいは市町村に郵送、という手順であった。
- ネット回答に関しては、「先日の大雨の被害の大きかった地域等について、インターネット回答が完了していないため、9 月 26 日から 10 月 20 日までの期間、紙の調査票と並行してインターネット回答を受け付けます」等、地域によって異なるが、期間が延長された\*<sup>2</sup>。なお、9 月中旬のネット回答率は都道府県により異なり、沖縄の 22.7% から滋賀の 48.4% まで幅があった\*<sup>3</sup>。
- 広報にも工夫があったが\*<sup>4</sup>、多数作られて <http://kokusei2015.stat.go.jp/gallery/> から見る事ができた PR 動画は、2017 年 5 月 29 日現在では消滅している。米国センサス局のサイトには <https://www.census.gov/library/video.html> として多数の動画が残っている（例えば、[https://www.census.gov/library/video/lgbt\\_we\\_all\\_count.html](https://www.census.gov/library/video/lgbt_we_all_count.html)）ので、日本の動画も残して欲しかった。
- 結果は 2016 年 2 月 26 日速報で初の総人口減とニュースになった\*<sup>5</sup>。

---

\*<sup>2</sup> <http://kokusei2015.stat.go.jp/internet/>

\*<sup>3</sup> <http://oku.edu.mie-u.ac.jp/~okumura/stat/nippon2.html>

\*<sup>4</sup> YouTube 動画 (<https://www.youtube.com/watch?v=jJfLqT0SjiM>) など。

\*<sup>5</sup> <http://www.stat.go.jp/data/kokusei/2015/kekka.htm>

## 2.2 日本の人口動態統計

- ある期間内（通常は1年）の人数が変わるイベント（＝人口動態）の数を集計する。
- 統計法に基づく指定統計であり，人口動態調査票（出生・死亡・死産・婚姻・離婚）の各項目について分類集計したものである。
- 人口動態調査票は，市区町村長（人口動態調査令の各条文は「市区町村長」だが，第6条で特別区長も「市区町村長」に含めると書かれている）が人口動態調査令および人口動態調査令施行細則に基づき，戸籍法による届書およびその他の関係書類（例えば死亡については，戸籍法第86条，第87条に基いて死亡届が市区町村に提出されるが，医師が作成した死亡診断書または死体検案書が添付される）から作成する（人口動態調査令第3条）。人口動態調査票の用紙は，厚生労働大臣から都道府県知事，保健所長を経て，各市区町村長に交付される（人口動態調査令第4条）
- 作成された人口動態調査票は，市区町村長→保健所長→都道府県知事→厚生労働大臣と提出される（人口動態調査令第5条）。この過程で，保健所長は出生票から出生小票を，死亡票から死亡小票を作成し，3年保管せねばならない。
- 市区町村分類は出生・死亡・死産については住所，婚姻は夫の住所，離婚は別居する前の住所による。死因統計は人口動態調査票の死亡について死因別に分類集計したものである。死因分類はICDによる。
- 厚生労働省が所管する。

### 国際疾病分類 (ICD) について

**定義** ある一定の基準により疾病を分類する体系。

**効用** 疾病の単位を明確にし，異なる調査結果を比較可能にする。

**ICD-10** ICDはWHOの前身である国際会議の協議により1900年に制定され，約10年毎に改定。1995年から第10回修正国際疾病分類（ICD-10）。最大24999種類まで可能。実際の基本分類は約14000項目。2018年5月の世界保健総会でのICD-11採択を目指して作業が行われている。WHO神戸センターの文書 ([http://www.who.int/kobe\\_centre/mediacentre/icd-11/ja/](http://www.who.int/kobe_centre/mediacentre/icd-11/ja/)) が参考になる。

**死因分類** 個々の疾病を約130項目にまとめたもの（ICD-9では「死因簡単分類」）。国連やWHOの統計資料は基本分類でなく死因分類や死因簡単分類でまとめられている。

## 2.3 死因について

日本では人口動態統計がしっかりしているので、年間の死因別死亡数が簡単に集計できるが、その仕組みは決して単純ではない。医師法により、日本では、すべての死亡は届け出られることになっている。医師が看取った場合や、継続診療中の疾患で亡くなった後で死体を検査した（最終診察後 24 時間以内の場合は再度死体を検査しなくてもいい）場合は、医師が死亡診断書を発行する（医師法 20 条）。それ以外の死はほぼすべて「異状死」として扱われ、医師が死体を検査した後で死体検案書を発行し、24 時間以内に所轄警察署に届け出なければならない（医師法 21 条）。なお、東京都など監察医制度があるところでは死体検案は監察医によって行われる。

日本では死後の解剖が義務付けられていないので、死後に解剖が実施される割合は 2% に過ぎず、死因の判定は主に医師が臨床診断として解剖せずに行うことになるが、臨床診断は誤っている場合が 12–30% に及ぶという指摘が、作家の海堂尊（本職は病理医）によってなされ（海堂尊『死因不明社会』講談社ブルーバックス）、死後に CT 等の画像診断をして真の死因を明らかにすること（画像診断だけでは確定できない場合には解剖を行う根拠ともなりうる）が提唱されている。

通常、死亡診断書または死体検案書は、A3 横の死亡届書の右半分で作成される<sup>\*6</sup>。死因については、直接死因から遡って原死因（WHO の定義では、「直接に死亡を引き起こした一連の事象の起因となった疾病もしくは損傷」または「致命傷を負わせた事故もしくは暴力の状況」）まで記載することになっている。例えば、10 年前から糖尿病を患っていた人が 1 ヶ月前に肺炎に罹り、回復することなく 5 日前に膿胸となり、2 時間 30 分前に敗血症性ショックを起こして死に至った場合、直接死因は敗血症性ショックだが、原死因は肺炎となるため、死亡の原因を記載する I 欄には、(ア) 敗血症性ショック、(イ) 膿胸、(ウ) 肺炎と記載する。糖尿病は「直接には死因に関係しないが I 欄の傷病経過に影響を及ぼした傷病名等」を記載する II 欄に記載する。

届出人（優先順位は同居の親族、同居していない親族、同居者、家主、地主、家屋管理人、土地管理人、公設所の長の順）は、死亡届書の左半分に記入押印し、市町村役場（優先順位は、死亡者の本籍地、死亡地、届出人の現住所地の順）の戸籍課に提出する<sup>\*7</sup>（ただし、通常、死亡届書の記入と提出は葬儀社により代行されている）。戸籍課では死亡者本人の戸籍の抹消と火埋葬許可証の交付を行い、死亡届書を 1 ヶ月保管した後、地方法務局へ送付する。

<sup>\*6</sup> <http://www.mhlw.go.jp/toukei/manual/index.html> として厚生労働省の web サイトで記入マニュアルが公開されている。

<sup>\*7</sup> このことを規定しているのは、上述の通り戸籍法である。

一方、人口動態調査令（厚生労働省令の1つ）により、厚生労働大臣は、人口動態調査票の用紙を保健所長経由で各市町村長に交付することになっている。市町村長は、上記死亡届書が提出されたらすぐに、人口動態調査票を作成し、人口動態調査票市町村送付票を添えて管轄保健所長に提出する（人口動態調査令施行細則）。保健所長はこれに基づき死亡小票を作成し3年間保存するとともに、1ヶ月ごとに取りまとめ（前月15日から当月14日まで）、人口動態調査票保健所送付票を添えて、当月25日までに都道府県知事（ただし政令指定都市や中核市に設置された保健所では市長）に提出する。都道府県知事（あるいは市長）は、これを取りまとめ、人口動態調査票都道府県送付票を添えて翌月5日までに厚生労働大臣に提出する。人口動態調査票を厚生労働大臣が保管する期間は翌年1年間だが、電子化されたデータは永久に保管することになっている。

上記のプロセスを経て、日本人すべての死亡データは厚生労働大臣のもとに集まり、人口動態統計として集計され報告されることになる<sup>\*8</sup>。そのおかげで、年齢別死因別死亡数といった数字を得ることができる。しかし、疫学研究などで個人ベースの詳細な死亡データが欲しいときは、死亡小票の目的外使用を申請し、総務大臣の許可を得てから、保健所で閲覧する必要がある、大変面倒である。平成19年の統計法改正を受けて学術会議から提言（次の枠内に一部引用）が出たが、いまだに改善されていない模様である。

---

<sup>\*8</sup> 人口動態統計では、原死因を死因として集計する。原死因は上記Ⅰ欄のもっとも下に書かれているものとなるのが原則だが、それが呼吸不全、心不全など死因としてふさわしくないとICD-10でWHOが定めたものになっている場合はⅡ欄の傷病を原死因とする。

## 学術会議分科会提言

**(5) 日本版 National Death Index の創設と運用**

疫学研究においては、一定集団について長期間観察を行うことによって始めて結果が得られることが多く、人の生死情報がしばしば必須であり、人口動態統計との連結なしには科学的な結論が見出せない。これまで、人口動態統計使用に当たっては、厚生労働省を經由し、総務省に目的外使用に関する申請を行い、認可を得て情報を入手することとなっていた。申請から認可までに要した期間は、平均 11.8 か月（最短で 3 か月、最長は 23 か月）であり、申請書等の修正回数は平均 10.2 回（最少で 0 回～最多は 60 回）であった（日本疫学会将来構想検討委員会「政府統計の利活用に関する調査報告書」平成 19 年 11 月 30 日、[<http://jeaweb.jp/activities/reports/pdf/20071206seifu.pdf>]）。これに比し、欧米諸国では健康情報登録（疾病登録、患者登録等）が整備され、死亡データとのリンケージも容易である。例えば、米国では、National Death Index システムを用いて、氏名、性、生年月日、住所、社会保障番号などを指標とする死亡ファイルとの記録照合による追跡調査が容易に可能である。このシステムを利用する場合、利用申請書のフォーマットや、それを作成するための詳細なマニュアル等が整備されており、申請処理期間は約 2 か月であり、多くの医学研究がこのシステムを用いて行われ、大きな成果をあげている（National Health Statistics. National Death Index. [<http://www.cdc.gov/nchs/ndi.htm>]）。これまで我が国の死亡統計ファイルには氏名が入力されていなかったため、このような死亡ファイルとの記録照合は不可能であった。しかし、2003 年 1 月より厚生労働省が進めている人口動態オンライン報告システムでは、死亡小票氏名が入力されるようになり、このシステムを充実させて全国的な運用を図れば、日本版 National Death Index の基盤が整うこととなる。統計情報の高度利用の一環として、死亡統計ファイルとの記録照合の技術的および制度的な検討を早急に開始し、日本版 National Death Index システムの構築とそれを活用するための手順の標準化・マニュアル化を進めるべきである。これらのシステムの円滑な運用によって、我が国の保健医療分野における疫学研究が飛躍的に進展し、国民の健康と安全のための有益な種々の施策提案に結びつくことが期待される。

出典：提言「保健医療分野における政府統計・行政資料データの利活用について—国民の健康と安全確保のための基盤整備として」, pp.10-11, 平成 20 年（2008 年）8 月 28 日, 日本学術会議基礎医学委員会・健康・生活科学委員会合同パブリックヘルス科学分科会

<http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-20-t62-6.pdf>

欧米先進諸国では概ね日本よりも信頼性の高い死亡データが得られる。地域がん登録の登録率も高いし、電子化も進んでいる。死因確定のための剖検割合も日本より高い国が多い。途上国では事情は一変する。パプアニューギニアやソロモン諸島では、そもそも死亡事象を行政が把握するまで年単位の遅れが存在する場合さえあるし、死因も必ずしも医師によって診断されない。個人ベースの死因を把握しようと思ったら、綿密な聞き取り（ただしクロスチェックは必須）によって再構成する方が、公式統計よりもむしろ信頼性が高い場合さえある。

年齢別死亡率のデータ（生命表データ）については、世界 37 ヶ国について Human Mortality Database (HMD; <http://www.mortality.org/>) で公開されていて、日本につ

いては国立社会保障・人口問題研究所の石井太さんが JMD (<http://www.ipss.go.jp/p-toukei/JMD/>) を公開されている。これらから入手できるデータはテキストファイルだが、HMD については、`demography` パッケージの中に R で直接読む関数がある。

## 2.4 その他のデータソース

HMD に相当する世界各国の年齢別出生率データは、Human Fertility Database (HFD; <http://www.humanfertility.org/cgi-bin/main.php>) として公開されている。ユーザ登録をしてログインすれば利用可能である。

多くの発展途上国で、人口と健康に関する調査 (Demographic and Health Surveys; DHS) が行われており、個人情報を含まないように匿名化された個票レベルのデータを利用することができる (<http://dhsprogram.com/data/available-datasets.cfm> からダウンロードできる)。

世界各国で行われている国勢調査から抽出された匿名化個票を集積しているプロジェクトが IPUMS INTERNATIONAL (<https://international.ipums.org/international/>) である。登録すれば 85 ヶ国 301 センサスのデータをダウンロードして利用することができる (ただし日本のデータはない)。IPUMS INTERNATIONAL は IPUMS (<https://www.ipums.org/>) の一部で、他には IPUMS USA として米国のセンサスマイクロデータを 1850 年から得られるプロジェクトや IPUMS NHGIS として 1790 年以降の米国センサスデータを GIS に関連づけて提供しているプロジェクト、IPUMS DHS として 1980 年以降の DHS データを時空間分析するために統合したものもある。

## 第3章

# 人口構造の分析

人口構造は、人口の特性の、ある時点における断面を示すものである。その意味で、人口静態 (static of population, state of population) の指標といえる。ただし、断面といっても、調査時点にたまたま家の外にいたというような場合や、たまたま客として訪問中だった人をどうするか、が常に問題となる。つまり、断面には、一定の時間幅がある。この時間幅をほとんど許さずに、調査期間中にいたかないかによって人口静態を切る方法が現在人口 (*de facto* population) 主義であり、常住 (すなわちある一定の期間の在・不在) によって切るのが常住人口 (*de jure* population) 主義である。

日本ではセンサス (国勢調査、つまり全国規模の人口静態調査) が5年ごとに行われているが (正確に言えば10年ごとに本調査でその間5年で補完調査として実施されており、補完調査の方が若干調べる項目数が少ない)、1947年までのセンサスが現在人口主義で行われていたのに対して、1948年常住人口調査を経て、1950年国勢調査から常住人口主義となった。変更の理由は、人口移動が激しくなったためとされている。なお、国勢調査の場合、在・不在を決めるための「一定の期間」は3ヶ月である。また、国勢調査では一時不在人口と一時現在人口も併せて調査されているので、常住人口 = 現在人口 + {(一時不在人口) - (一時現在人口)} という関係によって、一方から他方を求めることは可能である。

出生、死亡、結婚、移動といった人口動態 (population dynamics, dynamics of population, movement of population)\*<sup>1</sup>によって、人口の規模や構造は安定または変動する。言い換えると、人口静態は人口動態によって生み出される結果である。人口動態が一定のままある程度の時間が経過すれば、人口構造も安定する (総人口に占める年齢別人口の割合が全ての年齢において一定値に収束する) ことがわかっている。簡単な数学モデルによってこの

---

\*<sup>1</sup> 人口規模を決定する要因である出生、死亡、移動と、人口構造の変動を決定する要因のうち、結婚、離婚、疾病および傷害の発生及び治癒、職業又は所属産業の転換など、これに胎児死亡と死産を加えて、人口動態の要因といい、これらのできごとを人口動態上の事件 (vital events) という (館, 1968)。

ことを示したのはロトカ (Lotka, Alfred) であった。この理論は安定人口モデルと呼ばれ、人口学の中でも最も有名な数理モデルの1つである。

人口構造の分析は形式人口学の基本であり、他の分析に先立って必ず行うべきである。もっとも基本的な人口構造は、性・年齢別人口である。これを図示する方法としては、人口ピラミッドがもっとも良く知られており、小学校の社会科でも登場する。けれども、Excelなどで作成しようとするといかに面倒である。Rでは、`plotrix`パッケージの `pyramid.plot()` 関数や、`Pyramid` パッケージでも作成できるが、おそらくもっとも多機能なのは著者が開発した `pyramid` パッケージであろう。本章では、まず `pyramid` パッケージを使って人口ピラミッドを描く方法を具体的に紹介する。

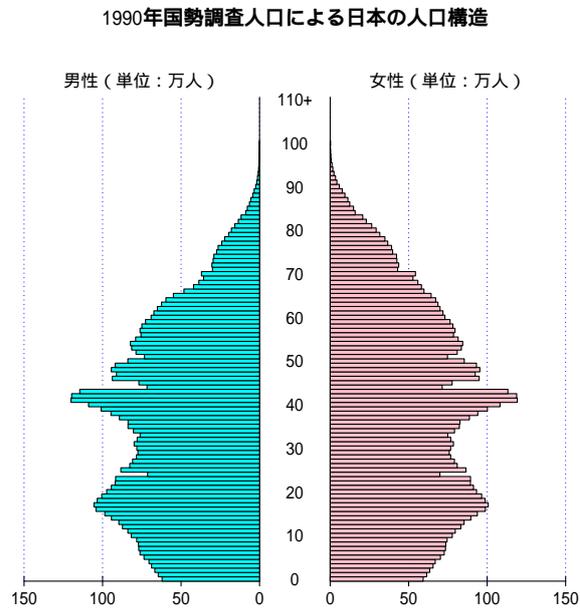
人口構造の分析としては、もちろん図示だけではなく、年少人口指数、65歳以上人口構造係数、老年化指数、ふくらみ指数など、さまざまな指標が知られている。本章の後半では、さまざまな人口構造指標を紹介し、計算方法とともにその意味を説明する。

### 3.1 国勢調査データから人口ピラミッドを描く

まずは、日本全国を対象とした人口ピラミッドを描いてみよう。これは、`fmsb` パッケージと `pyramid` パッケージを使えば簡単である。例えば、1990年の国勢調査人口で男女年齢1歳階級別の人口ピラミッドを描画するには下枠内のコード (<http://minato.sip21c.org/ldaR/pp1990J.R>) を打てば良い。

pp1990J.R

```
if (!require(fmsb)) { install.packages("fmsb"); library(fmsb) }
if (!require(pyramid)) { install.packages("pyramid"); library(pyramid) }
pyramid(data.frame(M=Jpopl$M1990/100000, F=Jpopl$F1990/100000, A=Jpopl$Age),
  Laxis=0:3*50, Llab="男性 (単位: 万人)", Rlab="女性 (単位: 万人)",
  Clab="", Cstep=10, main="1990年国勢調査人口による日本の人口構造")
```



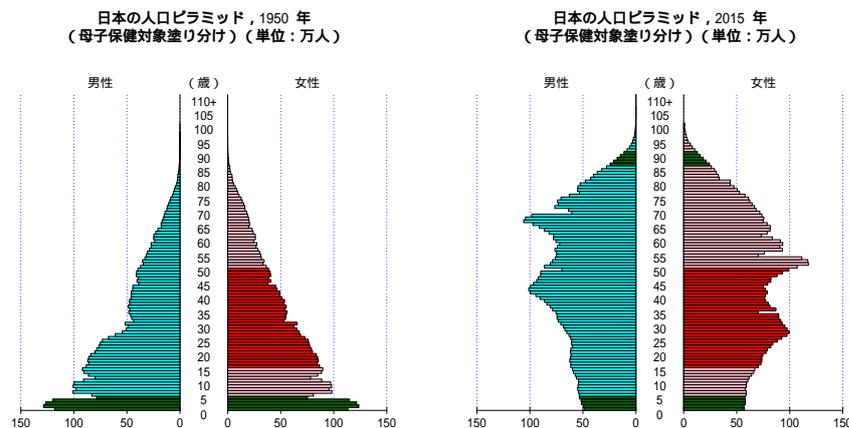
### 3.1.1 特定年齢層を強調する

`pyramid()` 関数には塗りつぶし色を指定する `Lcol` と `Rcol` というオプションがあり、柔軟な色指定が可能である。特定年齢層だけを別の色で塗ることもできる。母子保健対象年齢層を強調するために色を変え、戦後すぐと最新の2015年のピラミッドを並べて描くコードは以下の通りである。

```

if (!require(pyramid)) { install.packages("pyramid"); library(pyramid) }
if (!require(fmsb)) { install.packages("fmsb"); library(fmsb) }
layout(t(1:2))
d2015 <- data.frame(M=Jpopl$M2015/10000, F=Jpopl$F2000/10000, A=Jpopl$Age)
d1950 <- data.frame(M=Jpopl$M1950/10000, F=Jpopl$F1950/10000, A=Jpopl$Age)
pyramid(d1950,
  Laxis=0:3*50, Llab="男性",Rlab="女性",Clab="(歳)", Cstep=5, Cadj=-0.05,
  Lcol=c(rep("darkgreen",5), rep("cyan",81)),
  Rcol=c(rep("darkgreen",5), rep("pink",10), rep("red",35), rep("pink",36)),
  main="日本の人口ピラミッド, 1950年\n(母子保健対象塗り分け)(単位:万人)")
pyramid(d2015,
  Laxis=0:3*50, Llab="男性",Rlab="女性",Clab="(歳)", Cstep=5, Cadj=-0.05,
  Lcol=c(rep("darkgreen",5), rep("cyan",81)),
  Rcol=c(rep("darkgreen",5), rep("pink",10), rep("red",35), rep("pink",36)),
  main="日本の人口ピラミッド, 2015年\n(母子保健対象塗り分け)(単位:万人)")

```



### 3.1.2 pyramid() 関数詳細

pyramid() 関数には、次の枠内に示す通り、さまざまなオプションがある。また、データフレームを与える代わりに `Left=`、`Right=`、`Center=` というオプションで、左右に描くための数値ベクトルと中央のラベルを与える関数 `pyramids()` と、既に何か描かれているグラフィックウィンドウ上に、`frame=` オプションで指定した枠（省略時は `c(-1.15, 1.15, -0.05, 1.1)`）となる、`pyramid()` 関数のデフォルトで想定される座標系の左端、右端、下端、上端）の中に人口ピラミッドを重ね描きする関数 `pyramidf()` も用意してある。これを使えば、2つ以上の人口ピラミッドを重ね描きしたり、地図上に自治体別の人口ピラミッドを描くことが可能になる。

## pyramid() 関数

```
pyramid(data, Laxis=NULL, Raxis=NULL,
  AxisFM="g", AxisBM="", AxisBI=3, Cgap=0.3, Cstep=1, Csize=1,
  Llab="Males", Rlab="Females", Clab="Ages", GL=TRUE, Cadj=-0.03,
  Lcol="Cyan", Rcol="Pink", Ldens=-1, Rdens=-1, main="", ...)
```

data	1 列目が左側に表示させたい数値, 2 列目が右側に表示させたい数値, 3 列目が中央に表示させたいラベルからなるデータフレーム。3 列目が存在しない場合は, <code>row.names(data)</code> が中央ラベルとなる。
Laxis	左側軸目盛。省略すると, データ 1 列目の最大値までを適当に区切る
Raxis	右側軸目盛。省略すると Laxis と同じ値になる
AxisFM	軸目盛表示のため <code>formatC</code> に与える書式指定文字列。省略すると "g" が与えられる
AxisBM	軸目盛をコンマで桁区切り表示させたいときは, このオプションに "," を指定する。省略時は無し
AxisBI	軸目盛を区切る桁を指定する。省略時は 3 桁区切り
Cgap	中央ラベルのためにとる幅 (左描画領域に対する比として与える)。省略時は 0.3
Cstep	中央ラベルの表示間隔。省略時は 1 なので, すべてのラベルを表示する。10 と指定すれば年齢ラベルが 10 階級ごとに表示される
Csize	中央ラベルのフォントサイズ。省略時は 1 なので, 他の文字と同じサイズ
Cadj	中央ラベルの表示位置調整。フォントサイズ等により微妙にずれることがあるので, その調節のためのパラメータ。省略時は -0.03
Llab	左側の図のラベル。省略時は "Males"
Rlab	右側の図のラベル。省略時は "Females"
Clab	中央ラベルの上に表示するラベル。省略時は "Ages"。消したければ, <code>Clab=""</code> と指定すれば良い
GL	グリッド線を表示するか否かを論理値で与える。省略時は TRUE で, グリッド線が描かれる
Lcol	左側の図の描画色。省略時は "Cyan"。もう少し濃い色にしたいときは, "deepskyblue" をお勧めする
Ldens	左側の図をハッチングするときの斜線の密度。省略時は -1 で, 塗りつぶしになる。塗りつぶし時は枠線が黒色だが, ハッチング時は枠線も描画色になる
Rcol	右側の図の描画色。省略時は "Pink"。もう少し濃い色にしたいときは, "deependpink" をお勧めする
Rdens	右側の図をハッチングするときの斜線の密度。省略時は -1 で, 塗りつぶしになる。塗りつぶし時は枠線が黒色だが, ハッチング時は枠線も描画色になる
main	人口ピラミッド上部に表示するタイトル文字列。省略すると "" なので, タイトル無しになる
...	他のオプション文字列

### 3.1.3 pyramidf() を使った重ね描き

同じく `fmsb` パッケージ内の国勢調査人口と `pyramidf()` 関数を使って、1950年と2000年の人口ピラミッドを重ね描きするコード<sup>\*2</sup>は次の枠内のようなになる。ここでは `pyramidf()` でオプション `frame=` を指定していないため、既に描かれたピラミッドの上に重ね描きされる。その場合は、`Ldens` と `Rdens` を適切に指定することにより、元のピラミッドが透過で見えるようにする必要がある。

alt-jpopl-pyramid.R

```
if (!require(fmsb)) { install.packages("fmsb"); library(fmsb) }
if (!require(pyramid)) { install.packages("pyramid"); library(pyramid) }
d2000 <- data.frame(M=Jpopl$M2000/1000,F=Jpopl$F2000/1000, A=Jpopl$Age)
d1950 <- data.frame(M=Jpopl$M1950/1000,F=Jpopl$F1950/1000, A=Jpopl$Age)
pyramid(d2000, Cstep=10, Laxis=0:6*200, Llab="", Rlab="", Clab="",
  main="日本の人口 (1950年と2000年, 単位:千人)")
pyramidf(d1950, Cstep=10, Laxis=0:6*200, Llab="", Rlab="", Clab="",
  Lcol="navy", Rcol="salmon", Ldens=10, Rdens=10)
legend("topleft",
  legend=c("2000年男性", "2000年女性", "1950年男性", "1950年女性"),
  fill=c("cyan", "pink", "navy", "salmon"),
  density=c(NA, NA, 30, 30))
```

### 3.1.4 都道府県間比較

次に都道府県別の人口ピラミッドを並べて表示し、相互に比較してみよう。この場合は、年齢各歳では細かすぎる（一つの階級の人数が少なくなりすぎる）ので、5歳階級で十分である。総務省統計局の国勢調査の結果の概要ページ<sup>\*3</sup>で公開されている、「抽出速報集計結果」のデータを使うと<sup>\*4</sup>、2010年10月の国勢調査人口の速報集計値による、全国と都道府県別の年齢5歳階級別人口ピラミッドを描くことができる。

次の R コード<sup>\*5</sup>を実行すれば6行8列で全国と各都道府県の人口ピラミッドが描かれ

<sup>\*2</sup> <http://minato.sip21c.org/ldaR/alt-jpopl-pyramid.R>

<sup>\*3</sup> <http://www.stat.go.jp/data/kokusei/2010/kekkagai.htm>

<sup>\*4</sup> 最近の総務省統計局のデータ提供方法はだいぶ改善されており、例えば、政府統計の総合窓口 (e-Stat) というサイトの中の、都道府県別の男女年齢5歳階級別人口の年次推移データ <http://www.e-stat.go.jp/SG1/estat/List.do?bid=000001007702&cycode=0> から Excel 形式のファイルをダウンロードすれば、男女年齢5歳階級別人口が隙間なく連続したセルに入っているため、以前より遥かに加工が容易になった。

<sup>\*5</sup> <http://minato.sip21c.org/ldaR/compare-pyramids-among-pref.R>

る。なお、Excel 形式の主要統計表の第 2 表から、縦長形式に組み替えつつ年齢不詳（意外に多いことに注意されたい）を除いてテキスト化したファイル\*6をいったんダウンロードして作業ディレクトリ\*7に置けば、1 行目のデータ読み込みはインターネット URL でなくローカルファイルからでも良い。

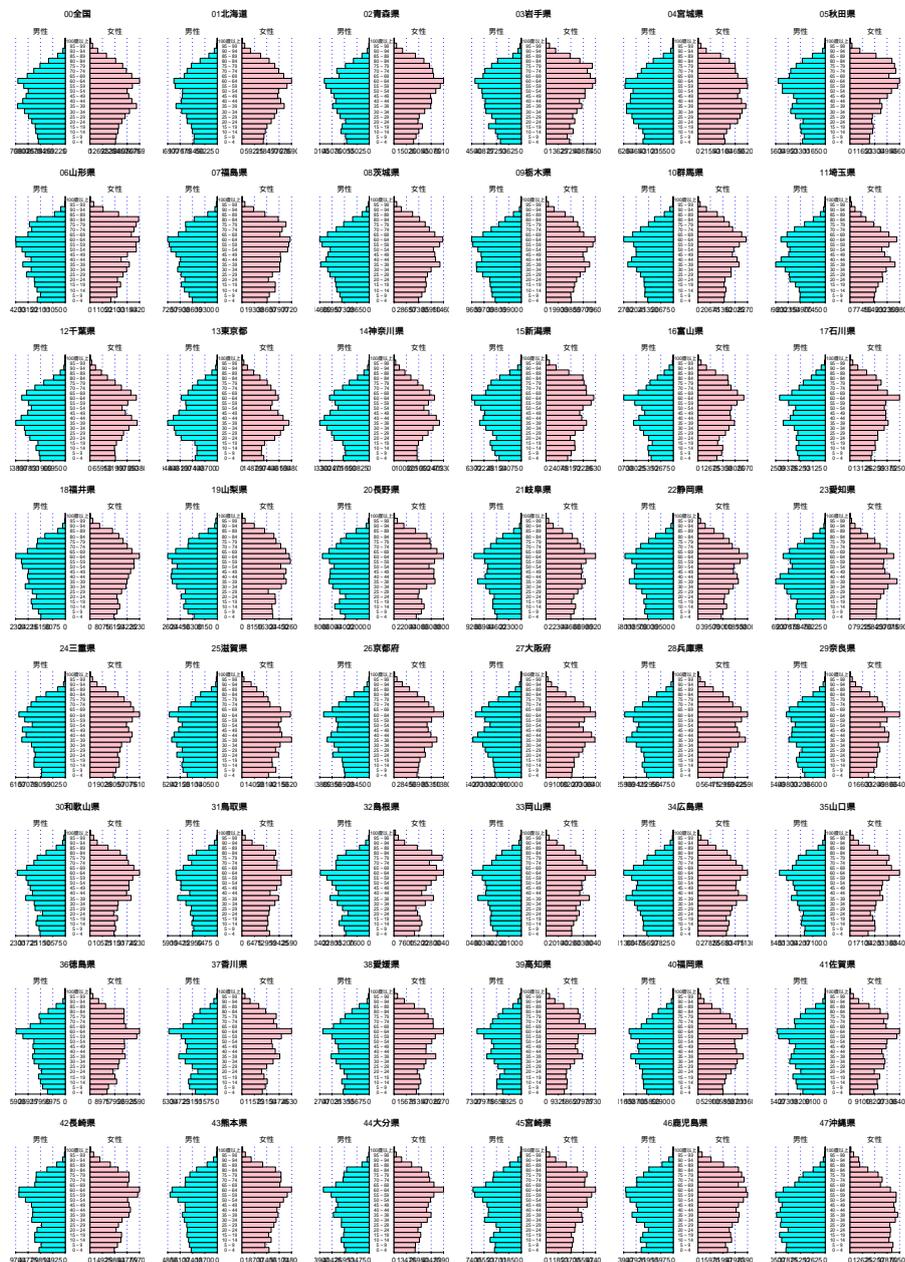
compare-pyramids-among-pref.R

```
x <- read.delim("http://minato.sip21c.org/ldaR/japancensus2010tp.txt",
  encoding="CP932")
if (!require(pyramid)) { install.packages("pyramid"); library(pyramid) }
layout(matrix(1:48, 6, 8, byrow=TRUE))
par(mar=c(1, 1, 2, 1), cex=0.6)
AreaList <- names(table(x$Area))
for (i in AreaList) {
  y <- subset(x, Area==i)
  pyramid(data.frame(Males=y$Males, Females=y$Females, Ages=y$Ages),
    Cadj=-0.01, Cgap=0.5, Csize=0.6, Llab="M", Rlab="F", Clab="",
    main=paste(i), AxisFM="fg")
}
```

軸目盛を手抜きしているの見にくいですが、ピラミッドの形を見るには十分だろう。明らかに東京都は 19 歳未満が少ないのに 20 歳以上で急増しているし、長野県と香川県は 20-24 歳だけ少なくなっていることがわかる。これは大学生や専門学校生の就学人口移動を意味していると思う。それにしても沖縄県だけがかろうじてピラミッドと呼べる形になっている。出生の章で示すように、TFR の年次変化を都道府県別にプロットすると沖縄県だけが他の都道府県の塊から離れて高いことが一目瞭然だが、人口ピラミッドの形からも沖縄県の出生率が高いことはわかる。

\*6 <http://minato.sip21c.org/ldaR/japancensus2010tp.txt>

\*7 Windows 版の R では、起動アイコンのプロパティで作業ディレクトリを指定することができる。また、Linux でも MacOS X でも、R を起動後、`setwd()` 関数で作業ディレクトリを指定できる。RStudio をフロントエンドにすれば、プロジェクトごとに作業ディレクトリを使い分けることが容易にできる。



### 3.1.5 地図上に人口ピラミッドを重ね描きする

日本全体だと小さくなりすぎるので適さないが、地方レベルでなら地図の上に人口ピラミッドを重ね描きすることもできる。

まず、関西6府県の人口ピラミッドを地図上に描いてみよう。都道府県レベルの地図ならば、わざわざシェイプファイルをダウンロードしなくても、Nippon パッケージや mapdata パッケージに日本地図情報が入っている。人口ピラミッドを重ね描きするためには各府県の中心座標が必要である。Nippon パッケージの JapanPrefecturesMap() 関

数は各都道府県にラベルを表示するための座標を返してくれる点は便利だが、日本全国を描いてしまい、関西だけの描画ができない。そこで、ここでは `mapdata` パッケージを用いることにする。基本的な使い方は以下のようにする。

```
library(mapdata)
map("japan", "hyogo") # 兵庫県の白地図を描く
map("japan") # 日本の白地図を描く
map("japan", "Hokkaido", fill=TRUE, col="red", add=TRUE) # 北海道を赤く塗る
```

地理情報は `mapdata` パッケージに入っているのですが、関西6府県の人口ピラミッドを描くために必要なのは、データだけである。e-Stat の国勢調査から都道府県及び大都市の男女年齢各歳別人口が入っているファイル [http://www.e-stat.go.jp/SG1/estat/GL08020103.do?\\_csvDownload\\_&fileId=000007809775&releaseCount=3](http://www.e-stat.go.jp/SG1/estat/GL08020103.do?_csvDownload_&fileId=000007809775&releaseCount=3) をダウンロードし、そこから関西6府県のデータをコピーして貼り付けることで、`kansai2015` というデータフレームに `SHIGAM`, `SHIGAF` のような変数として人口データを参照できるようにする。それができたら、下記のコード<sup>\*8</sup>を実行すると、その下の画像が描かれる。ポイントは `Ldens=1` と `Rdens=1` という指定である。これはハッチングさせるためというよりは、各年齢ごとの人口を示すバーの枠を塗りつぶし色で描かせるためである。そうでないと画像が黒く潰れてしまう。

---

<sup>\*8</sup> データ定義部分からの全てを <http://minato.sip21c.org/ldaR/kansai-pyramids.R> からダウンロードできる。

## kansai-pyramids.R

```

# まず関西全体を描画
map("japan", c("shiga", "kyoto", "osaka", "hyogo", "nara", "wakayama"))
gpa <- par()$usr # 描画全体の枠を得る
par(cex=0.5) # 描画文字サイズを0.5倍にする

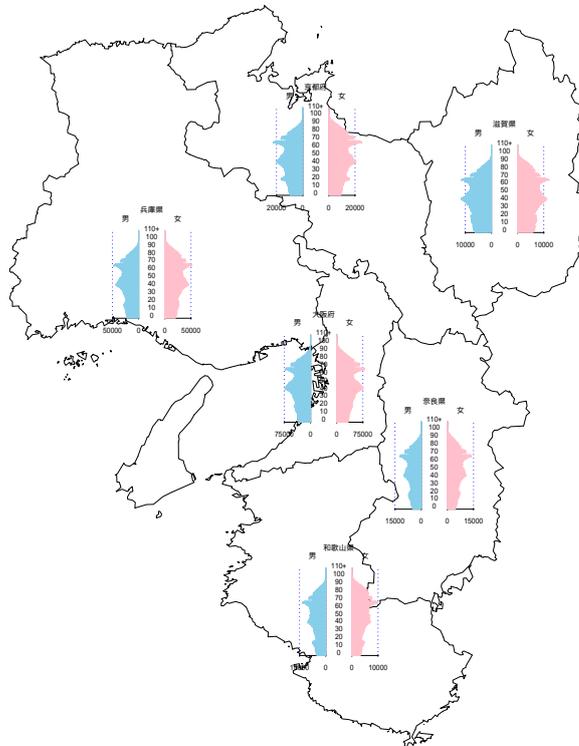
# 重心座標をフレームに換算する関数を定義
cent2frame <- function(x, y,
  sizex=(gpa[2]-gpa[1])/15, sizey=(gpa[4]-gpa[3])/15)
  { return(c(x-sizex, x+sizex, y-sizey, y+sizey)) }

# 各府県の人口ピラミッドを描画
# 滋賀県
pyramidf(data.frame(M=kansai2015$SHIGAM, F=kansai2015$SHIGAF, A=kansai2015$AGE),
  frame=cent2frame(geo.shiga.x, geo.shiga.y), Cstep=10,
  Lcol="skyblue", Ldens=1, Rcol="pink", Rdens=1, Clab="", Cgap=1,
  Llab="男", Rlab="女", Laxis=c(0, 10000),
  main="滋賀県", AxisFM="d")
# 京都府
pyramidf(data.frame(M=kansai2015$KYOTOM, F=kansai2015$KYOTOF, A=kansai2015$AGE),
  frame=cent2frame(geo.kyoto.x, geo.kyoto.y), Cstep=10,
  Lcol="skyblue", Ldens=1, Rcol="pink", Rdens=1, Clab="", Cgap=1,
  Llab="男", Rlab="女", Laxis=c(0, 20000),
  main="京都府", AxisFM="d")
# 大阪府
pyramidf(data.frame(M=kansai2015$OSAKAM, F=kansai2015$OSAKAF, A=kansai2015$AGE),
  frame=cent2frame(geo.osaka.x, geo.osaka.y), Cstep=10,
  Lcol="skyblue", Ldens=1, Rcol="pink", Rdens=1, Clab="", Cgap=1,
  Llab="男", Rlab="女", Laxis=c(0, 75000),
  main="大阪府", AxisFM="d")
# 兵庫県
pyramidf(data.frame(M=kansai2015$HYOGOM, F=kansai2015$HYOGOF, A=kansai2015$AGE),
  frame=cent2frame(geo.hyogo.x, geo.hyogo.y), Cstep=10,
  Lcol="skyblue", Ldens=1, Rcol="pink", Rdens=1, Clab="", Cgap=1,
  Llab="男", Rlab="女", Laxis=c(0, 50000),
  main="兵庫県", AxisFM="d")
# 奈良県
pyramidf(data.frame(M=kansai2015$NARAM, F=kansai2015$NARAF, A=kansai2015$AGE),
  frame=cent2frame(geo.nara.x, geo.nara.y), Cstep=10,
  Lcol="skyblue", Ldens=1, Rcol="pink", Rdens=1, Clab="", Cgap=1,
  Llab="男", Rlab="女", Laxis=c(0, 15000),
  main="奈良県", AxisFM="d")
# 和歌山県
pyramidf(data.frame(M=kansai2015$WAKAYAMAM, F=kansai2015$WAKAYAMAF, A=kansai2015$AGE),
  frame=cent2frame(geo.wakayama.x, geo.wakayama.y), Cstep=10,
  Lcol="skyblue", Ldens=1, Rcol="pink", Rdens=1, Clab="", Cgap=1,
  Llab="男", Rlab="女", Laxis=c(0, 10000),
  main="和歌山県", AxisFM="d")

title(main="関西地区の2015年国勢調査人口", cex.main=2)

```

関西地区の2015年国勢調査人口



次に、神戸市の地図上に区別に人口ピラミッドを描いてみる。必要なファイルは、(1) 神戸市の地図情報データ、(2) 神戸市各区の年齢5歳階級別人口データ、の2つである。

- (1) 神戸市地図情報 国土地理院の国土数値情報行政区画データ ([http://nlftp.mlit.go.jp/ksj/gml/datalist/KsjTmplt-N03-v2\\_3.html](http://nlftp.mlit.go.jp/ksj/gml/datalist/KsjTmplt-N03-v2_3.html)) から兵庫県のシェイプファイルをダウンロードし展開する。欲しいのは神戸市のデータだけなので、CDCが無償公開しているソフトである EpiInfo (<https://www.cdc.gov/epiinfo/pc.html>) の EpiMap を使って兵庫県のデータから神戸市のデータだけを抜き出して保存する。具体的には AddLayerPartial 機能を使って神戸市だけの地図情報を kobe-city.shp として保存すれば良い<sup>\*9</sup>。
- (2) 神戸市各区の年齢5歳階級別人口 神戸市が独自に国勢調査結果をまとめているページ (<http://www.city.kobe.lg.jp/information/data/statistics/toukei/jinkou/tyoubetsujinkou.html>) から、平成22年国勢調査結果の「全市」をダ

<sup>\*9</sup> この加工済ファイルをダウンロード可能にするわけにはいかないのですが、面倒でも上記プロセスを踏んで加工していただきたい。

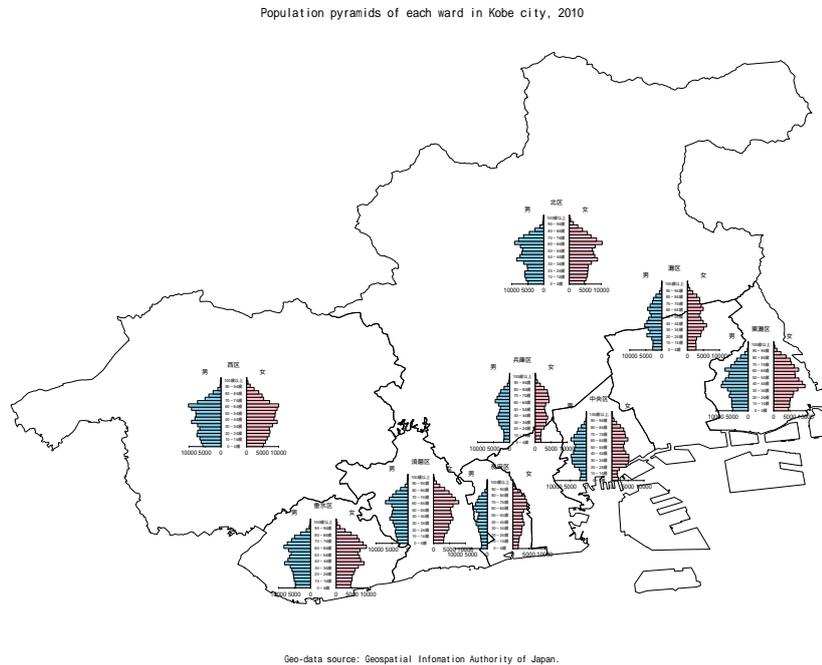
ダウンロードしてタブ区切りテキスト形式に加工したファイルで、<http://minato.sip21c.org/ldaR/2010kobepop-sex-ageby5.txt> からダウンロードできる。

Rの作業ディレクトリ（Windowsの場合、起動用ショートカットのプロパティの「作業フォルダ」で設定するか、起動後にメニューの「ファイル」の「ディレクトリ変更」で指定するか、`setwd()` 関数で指定することができる。また、RStudioでは作業ディレクトリを明示的に指定するのが普通）にシェイプファイル（`kobe-city.shp`）を置き、以下のコード<sup>\*10</sup>を実行すると、その下の画像が描画される。

kobe-plot.R

```
library(maptools)
library(pyramid)
# Kobe extracted by EpiMap's AddLayerPartial from Hyogo of Kokudo-chiriin
kobe <- readShapeSpatial("./kobe-city.shp")
kobedata <- kobe @ data
cents <- coordinates(kobe)
geo <- data.frame(x=cents[,1], y=cents[,2])
geo$areas <- sapply(slot(kobe,"polygons"),slot,"area")
geo$JCODE <- kobedata$N03_007
pop <- read.delim("http://minato.sip21c.org/ldaR/2010kobepop-sex-ageby5.txt",
  fileEncoding="CP932")
par(family="Japan1GothicBBB", mar=c(6,1,5,1), cex=0.5)
plot(kobe, col=NA, xlab="", ylab="", axes=FALSE)
gpa <- par()$usr
cent2frame <- function(x, y, sizex=(gpa[2]-gpa[1])/20, sizey=(gpa[4]-gpa[3])/15)
  { return(c(x-sizex, x+sizex, y-sizey, y+sizey)) }
for (i in names(table(kobedata$N03_007))) {
  ku <- subset(geo, JCODE==i)
  j <- which.max(ku$areas)
  kupop <- subset(pop, JCODE==i)
  pyramidf(kupop[,1:3],
    frame=cent2frame(ku[j,1]-ifelse(i==28105, 0.01, 0),
      ku[j,2]+ifelse(i==28105, 0.04, ifelse(i==28102, 0.03, 0))),
    Cstep=2, Csize=0.6, Cadj=-0.07, Cgap=0.8, Lcol="skyblue", Rcol="pink", GL=FALSE,
    Llab="男", Rlab="女", Clab="", Laxis=0:2*5000, main=kupop$Ward[1], AxisFM="d")
}
title(main="Population pyramids of each ward in Kobe city, 2010",
  sub="Geo-data source: Geospatial Infomation Authority of Japan.",
  cex.main=2, cex.sub=1.5)
```

\*10 <http://minato.sip21c.org/ldaR/kobe-plot.R>



### 3.1.6 アニメーションを作る

`animation` パッケージを使うと, `gif`, `swf` など, さまざまなファイル形式で, 人口ピラミッドが経時的に変化していく様子をアニメーションにすることができる。例えば `fmsb` パッケージの `Jpopl` に入っている日本の国勢調査人口のピラミッドを `gif` アニメーションにするコード\*<sup>11</sup>は以下の通りである (<http://minato.sip21c.org/ldaR/> 参照)。

\*<sup>11</sup> <http://minato.sip21c.org/ldaR/JapanCensusAnimation.R>

```

JapanCensusAnimation.R
if (!require(fmsb)) { install.packages("fmsb"); library(fmsb) }
if (!require(pyramid)) { install.packages("pyramid"); library(pyramid) }
if (!require(animation))
  { install.packages("animation"); library(animation) }
ani.options(interval=1, ani.type="png")
drawanim <- function() {
  for (i in c(1:24*2, 52, 56, 60)) {
    pyramid(
      data.frame(M=Jpopl[, i]/100000, F=Jpopl[, i+1]/100000, A=Jpopl[, 1]),
      Laxis=0:4*50, AxisFM="d", Llab="男性 (万人)", Rlab="女性 (万人)",
      Cstep=10, main=gsub("M([0-9]+)", "Year \\1", names(Jpopl)[i]))
    ani.pause()
  }
}
saveGIF(drawanim(), "./JapanCensusAnimation.gif")

```

## 3.2 手作りのセンサス生データから人口構造を扱う

国レベルや都道府県レベルなどでは、既に集計済のデータを扱えば良いが、もっと小さい区分のデータを扱う目的で、自分で調査をするか、あるいは住民基本台帳や Demographic and Health Surveys (DHS) のようなソースから生データを入手する場合がある。このとき存在するデータは個人ベースの性別や年齢の情報なので、性・年齢別に集計する必要が生じる。小集団の場合、年齢各歳別では0人の階級が多く生じる可能性があるため、5歳階級等の集計も必要になる。

### 3.2.1 データファイルの構造

性別は基本的に変化しないが（もちろん性同一性障害で性別再判定手術を受けたりすれば別である）、年齢は毎年増えるため、データファイルに保持する情報としては、年齢よりも生年月日の方が良い。途上国の小集団では自分の年齢を知らない人も多く、その場合はカレンダー法などさまざまな年齢推定のテクニックを駆使する必要が生じる<sup>\*12</sup>。

ここでは、既に生年月日はわかっているとして、そのデータファイルの構造をどうすれば良いか、その後の集計をどうすれば良いかについて説明する。

現在では研究倫理や個人情報保護の観点から、解析対象のデータファイルには、個人を特定できる名前や住所などの情報を含むべきではないとされている。しかし、縦断的な研

<sup>\*12</sup> 例えば、中澤 港 (2007) 小集団人口学. In: 稲葉 寿 (編) 現代人口学の射程. ミネルヴァ書房を参照。

究をしたい場合は、異なる時点で得られた同一人に対する情報を連結できなくては困る。そこで使われるデータベースの技法が、連結可能匿名化である。簡単にいってしまえば、個人情報が入ったファイルとデータを保持するファイルを分け、個人には ID 番号を振って、両方のファイルを同じ個人 ID 番号で照合できるようにするのである。

個人情報が入ったファイルは、ID が重複しないようにすることが肝要である。そのため、死亡や脱落があっても ID は再利用しない方が良い。セキュリティのためには暗号化すべきである。

データファイルは表形式にする。原則として 1 人が 1 行になるように入力する。個人のさまざまな属性を列にしていく。一番左の列を ID 番号、次の列を性別にするのが普通である。生年月日の入力にはさまざまな方法があるが、例えば西暦年を 4 桁、月を 01 から 12 の 2 桁、日を 01 から 31 の 2 桁で表した 8 桁の数字で入力することもできるし、年と月と日を別々の変数として入力することもできる。どちらが良いかは一長一短である。

### 3.2.2 架空データを使った集計の例

<http://minato.sip21c.org/ldaR/hyppop.txt> は、乱数を使って生成した架空の人口である。変数は ID (ID 番号)、SEX (1 が男性、2 が女性)、AGE (年齢)、MRD (配偶者の ID。配偶者がいない場合は 0)、NC (既往子供数)、LCB (一番下の子供が生まれた年齢) の 6 つからなる。年齢のところは生年月日にして、読み込んだ後で年齢を計算させる場合もあるが、実際の小集団調査で得られるデータもこのような形で入力することができる。

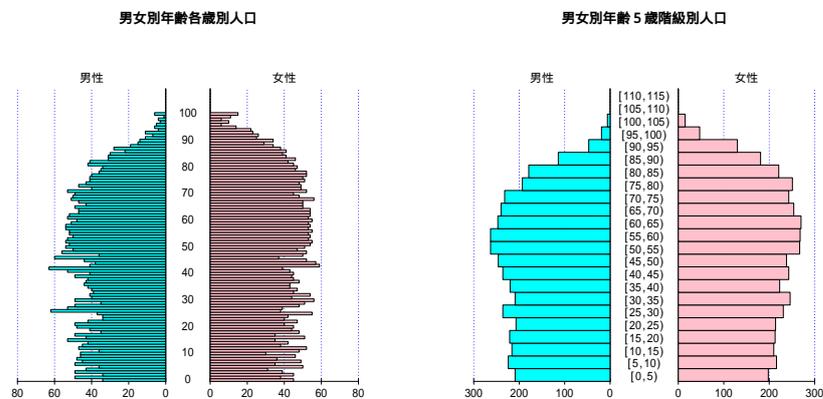
このデータをインターネット上から読み込み、`cut()` 関数を使って AGE を 1 歳階級と 5 歳階級の階級データに変換し (1 歳階級であっても AGE そのままではなく階級データに変換するのは、偶然その年齢の人がいない階級もスキップされないようにするためである)、`xtabs()` 関数を使って性別と年齢階級のクロス集計結果を行列に保存し、それを使って人口ピラミッドを描画するコードと結果のグラフは次の通り。

```

makepyramid-from-raw.R
if (!require(pyramid)) { install.packages("pyramid"); library(pyramid) }
raw <- read.delim("http://minato.sip21c.org/ldaR/hyppop.txt")
MAXAGE <- ifelse(max(raw$AGE)>=110, max(raw$AGE)+1, 110)
raw$AGEC1 <- cut(raw$AGE, 0:MAXAGE, right=FALSE)
MAC <- (MAXAGE %% 5) + 1
raw$AGEC5 <- cut(raw$AGE, (0:MAC)*5, right=FALSE)
TAB1 <- xtabs(~ AGEC1 + SEX, data=raw)
TAB5 <- xtabs(~ AGEC5 + SEX, data=raw)

layout(t(1:2))
pyramid(data.frame(M=TAB1[, 1], F=TAB1[, 2], A=0:(MAXAGE-1)),
  Laxis=0:4*20, Llab="男性", Rlab="女性", Clab="", Cstep=10,
  main="男女別年齢各歳別人口")
pyramid(data.frame(M=TAB5[, 1], F=TAB5[, 2], A=rownames(TAB5)),
  Laxis=0:3*100, Llab="男性", Rlab="女性", Clab="", Cgap=0.5,
  Cadj=-0.02, main="男女別年齢5歳階級別人口")

```



### 3.3 人口構造の指標値

ピラミッドを見ればわかるとはいえ、長期的な変化の傾向を見るなどの目的では、人口構造の情報を集約する指標値が役に立つ。良く用いられる指標が、年少人口指数、老年人口指数、老年化指数、従属人口指数の4つである<sup>\*13</sup>。

これらは、年齢三区分別人口の比に基づいている。社会において生産に従事する年齢を生産年齢と呼ぶが、通常は義務教育が終わってから定年退職するまでなので、15歳から

\*13 <http://www.un.org/esa/population/publications/worldageing19502050/pdf/95annexi.pdf>

59歳までか、15歳から64歳までの人口になる（2017年現在の日本では後者を用いている）。生産年齢人口以外の人口は従属人口と呼ばれる。従属人口のうち若い方はこれから生産年齢になり、定年退職後は二度と生産年齢にはならないため、区別される。若い方の0歳から14歳までを年少人口と呼び、高齢側の60歳以上または65歳以上の人口（2017年現在の日本では後者）を老年人口と呼ぶ。

一方、分母を総人口として、そのうちの各区分人口が占める割合を示す指標が年齢構造係数である。x歳年齢構造係数とは、総人口に占めるx歳人口の割合である。上記年齢三区分別人口を使えば、年少人口構造係数、生産年齢人口構造係数、老年人口構造係数を計算することができるが、老年人口構造係数は、65歳以上人口構造係数と呼ぶことが多い。

### 3.3.1 年少人口構造係数

社会生活、ことに経済活動の見地から、通常、15歳未満は年少人口として区分される。細分すれば、0歳が乳児期、1～4歳が幼児期、5～14歳が学齢期とされる。年少人口の総人口の中での割合が年少人口構造係数（0～14歳人口構造係数）である。

### 3.3.2 生産年齢人口構造係数

同様の見地から、15歳以上で、仕事をリタイヤする以前の、60歳未満あるいは65歳未満は、生産年齢人口として区分される。日本の人口統計では、通常65歳を区切りとする。生産年齢人口が総人口に占める割合が、生産年齢人口構造係数（15～64歳人口構造係数）である。

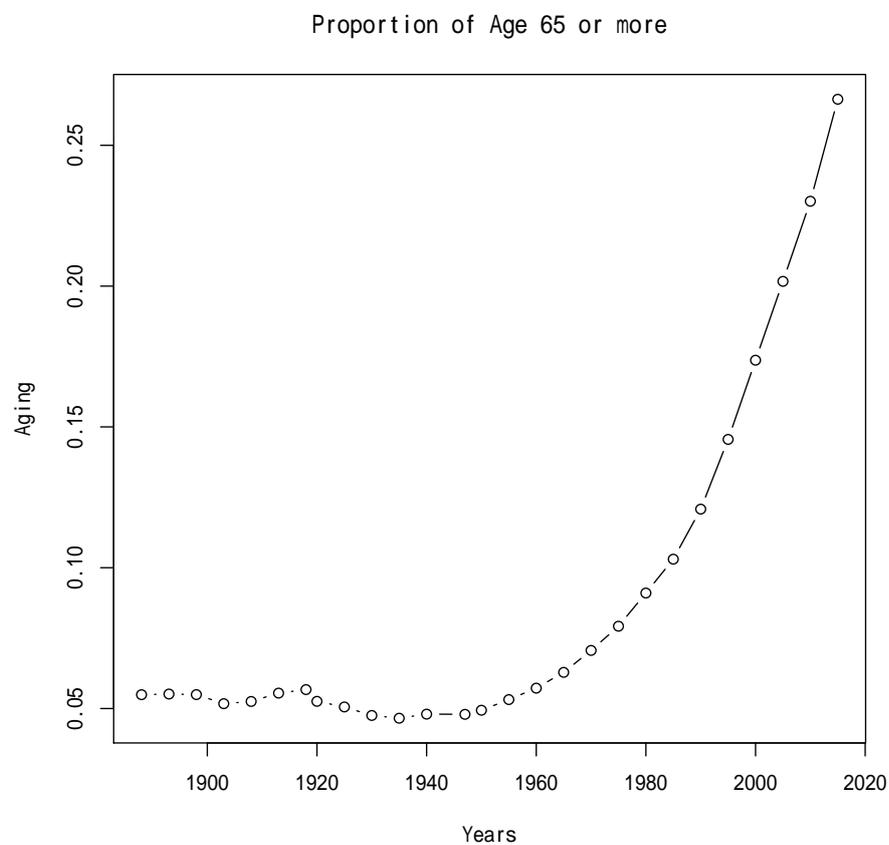
### 3.3.3 老年人口構造係数（65歳以上人口構造係数）

年少人口と生産年齢人口を除いた、生産に従事しない高齢者の人口を老年人口という。日本の人口統計では、65歳以上人口構造係数と呼ぶのが普通である。国勢調査人口を使って日本人口の高齢化の様子を折れ線グラフにすると、1960年代以降急速に高齢化が進行してきたことがわかる。コード\*<sup>14</sup>とグラフを以下に示す。

\*<sup>14</sup> <http://minato.sip21c.org/ldaR/Aging-Trend-Japan.R>

## Aging-Trend-Japan.R

```
if (!require(fmsb)) { install.packages("fmsb"); library(fmsb) }
Years <- c(1888+0:6*5, 1920+0:4*5, 1947, 1950+0:13*5)
Aging <- numeric(27)
for (i in 1:27) {
  Aging[i] <- sum(Jpop[66:86, c(i*2, i*2+1)])/sum(Jpop[, c(i*2, i*2+1)])
}
plot(Years, Aging, main="Proportion of Age 65 or more", type="b")
```



### 3.3.4 年少人口指数 (youth dependency ratio)

年少人口の生産年齢人口に対する比である。2015年の日本の国勢調査データによると、約0.208である。

```
if (!require(fmsb)) { install.packages("fmsb"); library(fmsb) }
#日本の 2015 年の年少人口指数
sum(Jpop$M2015[1:15] + Jpop$F2015[1:15]) /
  sum(Jpop$M2015[16:65] + Jpop$F2015[16:65])
```

### 3.3.5 老年人口指数 (aged dependency ratio / old-age dependency ratio)

老年人口の生産年齢人口に対する比。2015 年の日本の国勢調査データによると、約 0.439 である。

```
if (!require(fmsb)) { install.packages("fmsb"); library(fmsb) }
#日本の 2015 年の老年人口指数
maxage <- length(Jpop$M2015)
sum(Jpop$M2015[66:maxage] + Jpop$F2015[66:maxage]) /
  sum(Jpop$M2015[16:65] + Jpop$F2015[16:65])
```

### 3.3.6 従属人口指数 (dependency ratio / total dependency ratio)

従属人口の生産年齢人口に対する比。年少人口と老年人口は、現在生産に従事していないという点では同じなので、これらを併せて従属人口と呼ぶ。従属人口の生産年齢人口に対する比を従属人口指数と呼ぶ。なお、人口循環の視点からみれば、年少人口が今後成長して生産年齢人口となり、労働市場に入ってくるのに対して、老年人口は再び労働市場に現れることがないという意味で大きく異なる点に注意すべきである。2015 年の日本の国勢調査データでは約 0.647 である。

```
if (!require(fmsb)) { install.packages("fmsb"); library(fmsb) }
#日本の 2015 年の従属人口指数
maxage <- length(Jpop$M2015)
sum(Jpop$M2015[c(1:15, 66:maxage)] + Jpop$F2015[c(1:15, 66:maxage)]) /
  sum(Jpop$M2015[16:65] + Jpop$F2015[16:65]) # ただコンマで区切ると 2 次元行列
  になるので c() で括る
```

### 3.3.7 老年化指数 (ageing index)

老年人口の年少人口に対する比。2015 年の日本の国勢調査データでは約 2.11 である。

```
if (!require(fmsb)) { install.packages("fmsb"); library(fmsb) }
#日本の2015年の老年化指数
maxage <- length(Jpop$M2015)
sum(Jpop$M2015[66:maxage] + Jpop$F2015[66:maxage]) /
  sum(Jpop$M2015[1:15] + Jpop$F2015[1:15])
```

### 3.4 年齢報告の不正確さについての指標

先進国ではあまり問題にならないが、国連の推奨に基づいて国勢調査が行われていても、途上国では年齢の報告が不正確な場合がある。そもそも自分の正確な年齢を知らない人が珍しくないので、国勢調査担当として選ばれた人や村の有力者が外見などから勝手に年齢を決めてしまうこともある。このような理由で年齢報告が不正確になると、(1)1桁目が5や0といった切りの良い数字で報告される傾向と、(2)若い方が良いとされる文化であれば若い方へ、年をとっている方が尊いとされている文化での高齢者なら高齢側にサバを読む傾向が出現する。

報告される年齢の1桁目が5や0に集中すると、ピラミッドを描いたときに、そこだけスパイク状に人口が増える「年齢の盛り上がり (age heaping)」が見られる。この age heaping の程度を示す指標としては、Whipple's Index が有名である。

Whipple's Index は、25歳から60歳まで5歳刻みの「切りが良い」年齢の人口を合計して5倍し、23歳から62歳人口の合計で割って100を掛けた値である。age heaping がなければ100に近い値になるはずである。目安としては、105未満なら「極めて正確 (highly accurate)」, 105以上110未満なら「かなり正確 (fairly accurate)」, 110以上125未満で「近似的 (approximate)」, 125以上175未満で「粗い (rough)」, 175以上で「極めて粗い (very rough)」と判定する。日本の国勢調査データは一貫して「極めて正確」である。

インドの2011年国勢調査結果を <http://www.censusindia.gov.in/2011census/C-series/C-13.html> から Excel 形式でダウンロードし、全国の男女別人口だけを抜き出したテキストファイルを <http://minato.sip21c.org/ldaR/India2011census.txt> に置いてある。これを読み込んで人口ピラミッドを描くコード<sup>\*15</sup>は以下の通り。

<sup>\*15</sup> <http://minato.sip21c.org/ldaR/India2011.R>

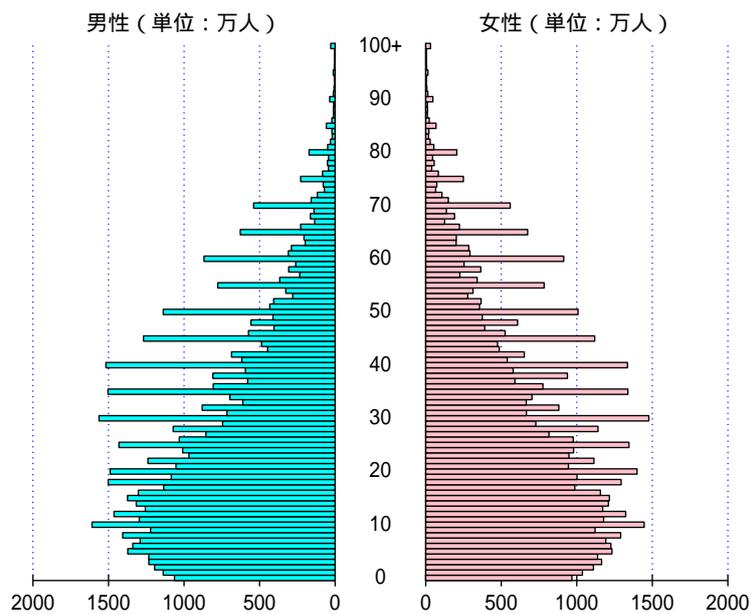
India2011.R

```

if (!require(pyramid)) { install.packages("pyramid"); library(pyramid) }
India <- read.delim("http://minato.sip21c.org/ldaR/India2011census.txt")
pdf("India2011census.pdf") # pdf ファイルへ出力
par(family="Japan1GothicBBB") # 日本語が文字化けしないようにするため
pyramid(
  data.frame(M=India$Males/10000, F=India$Females/10000, A=India$Age),
  Laxis=0:4*500, Llab="男性 (単位: 万人)", Rlab="女性 (単位: 万人)",
  Clab="", Cstep=10, main="2011年国勢調査人口によるインドの人口構造")
dev.off()
if (!require(fmsb)) { install.packages("fmsb"); library(fmsb) }
print(WhipplesIndex(India$Males+India$Females))

```

2011年国勢調査人口によるインドの人口構造



明らかに多くの年齢の盛り上がりが見られる。Whipple's Index は、171.0422 で「粗い」という判定結果である。

## 3.5 ふくらみ指数と年少人口構造係数の関係

### 3.5.1 ふくらみ指数とは

ふくらみ指数とは、既に故人となられた黒田俊夫先生が1976年に書かれた『日本人口の転換構造』（古今書院）の中で提唱されている人口流入の指標で、15歳から34歳までの人口が比較的移動しやすいのに対して（都会への就職，地元に戻るUターン，地方にマイホームを買うIターンなど），比較的移動しないと考えられる5歳から14歳までと，35歳から44歳までの比をとって，%単位で示した値である。人口構造データだけを使って移動という人口動態を推定しようとしている点がユニークである。式で書けば，

$$\text{ふくらみ指数} = \frac{(15-34 \text{ 歳人口})}{\{(5-14 \text{ 歳人口}) + (35-44 \text{ 歳人口})\}} \times 100$$

となる。100より大きければ人口が流入していて，100より小さければ人口が流出していると考ええる。

しかし21世紀の現在では中卒で就職する人がほとんどいないため，高校卒業まではあまり移動しないと考えられるし，平均初婚年齢やマイホーム取得年齢も上がっていると考えられるので，これを5歳上方シフトさせ，

$$\text{ふくらみ指数改} = \frac{(20-39 \text{ 歳人口})}{\{(10-19 \text{ 歳人口}) + (40-49 \text{ 歳人口})\}} \times 100$$

と定義してみよう。英語にすれば Population Expansion Index となるため，PEI(C) という関数名でコード<sup>\*16</sup>を書いた。

<sup>\*16</sup> <http://minato.sip21c.org/ldaR/PEI.R>

PEI.R

```
PEI <- function(X, CLS=1, MODE=1) {  
  # X はゼロ歳からの年齢階級別人口ベクトル。  
  # CLS は年齢階級幅。デフォルトは年齢各歳。1 か 5 か 10 のみ取れる。  
  # MODE=1 でふくらみ指数改, 他の MODE でオリジナルのふくらみ指数  
  N1 <- 20/CLS + 1 - ifelse(MODE==1, 0, 5/CLS)  
  N2 <- N1 + 20/CLS - 1  
  D1 <- 10/CLS + 1 - ifelse(MODE==1, 0, 5/CLS)  
  D2 <- D1 + 10/CLS - 1  
  D3 <- D1 + 30/CLS  
  D4 <- D2 + 30/CLS  
  return(sum(X[N1:N2]) / (sum(X[D1:D2]) + sum(X[D3:D4])) * 100)  
}
```

### 3.5.2 ふくらみ指数と年少人口との関係

この数字との関係をみたら面白いと思われるのは、移動できる可能性のある人口を提供する能力という意味で、既に述べた年少人口構造係数が良からう。これも % 単位であらわす関数を以下のようにコードする\*17。

Y.P.R

```
YP <- function(X, CLS=1) {  
  # X はゼロ歳からの年齢階級別人口ベクトル。  
  # CLS は年齢階級幅。デフォルトは年齢各歳。1 か 5 のみ取れる。  
  NE <- 15/CLS  
  return(sum(X[1:NE])/sum(X)*100)  
}
```

その上で、既に使った 2010 年 10 月の国勢調査人口の速報集計値により、都道府県別に提供されている男女年齢 5 歳階級別人口のデータを使って、ふくらみ指数改と年少人口構造係数を男女別に都道府県ごとに計算させ、横軸に年少人口構造係数、縦軸にふくらみ指数改をとった散布図を描かせるコード\*18は以下のように書ける。

\*17 <http://minato.sip21c.org/ldaR/YP.R>

\*18 <http://minato.sip21c.org/ldaR/reasons-PEI-YP.R>

```
relations-PEI-YP.R

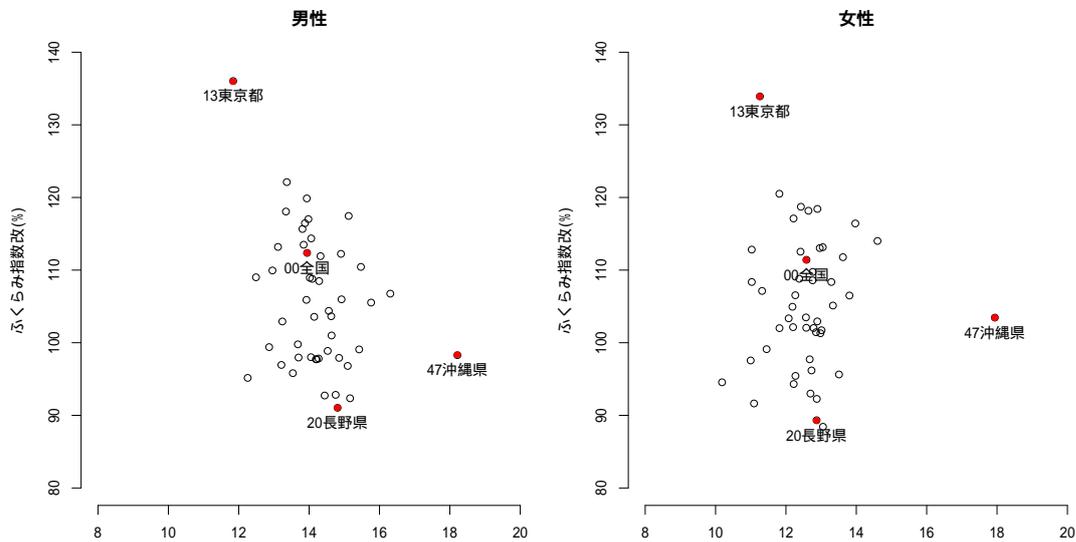
x <- read.delim("http://minato.sip21c.org/ldaR/japancensus2010tp.txt",
  encoding="CP932")

source("http://minato.sip21c.org/ldaR/PEI.R")
source("http://minato.sip21c.org/ldaR/YP.R")

areanames <- names(table(x$Area))
malePEI <- as.vector(by(x$Males, x$Area, PEI, CLS=5))
maleYP <- as.vector(by(x$Males, x$Area, YP, CLS=5))
femalePEI <- as.vector(by(x$Females, x$Area, PEI, CLS=5))
femaleYP <- as.vector(by(x$Females, x$Area, YP, CLS=5))
textdisp <- c(1,14,21,48)

layout(t(1:2))
plot(maleYP, malePEI, main="男性", frame.plot=FALSE,
  xlim=c(8,20), ylim=c(80,140), xlab="年少人口構造係数 (%)",
  ylab="ふくらみ指数改 (%)")
points(maleYP[textdisp], malePEI[textdisp], pch=16, col="red")
text(maleYP[textdisp], malePEI[textdisp], paste(areanames[textdisp]),
  pos=1)
plot(femaleYP, femalePEI, main="女性", frame.plot=FALSE,
  xlim=c(8,20), ylim=c(80,140), xlab="年少人口構造係数 (%)",
  ylab="ふくらみ指数改 (%)")
points(femaleYP[textdisp], femalePEI[textdisp], pch=16, col="red")
text(femaleYP[textdisp], femalePEI[textdisp], paste(areanames[textdisp]),
  pos=1)
```

実行すると以下の散布図が描かれる。



他の道府県に比べ、東京都と沖縄県が特異点であることが一目瞭然であろう。また、長野県が男女とも年少人口構造係数が中程度なのに、ふくらみ指数改がきわめて低い値になっているのが目立つ。移動統計によると東北各県の方が長野県よりも人口流出は大きいので、この結果が意味することは、長野県には大学が少ないために20代と30代人口が流出するが、その後で帰ってきたりIターン就農したりする人口が一定の水準で存在するということと思われる。

### 3.5.3 ふくらみ指数によるコロプレス図

この「ふくらみ指数改」を何段階かに区分して地図上で塗り分け表示した、いわゆるコロプレス図の作成は、地理的な分布を考察する上で有益であろう。

### シェイプファイルを使う王道な方法

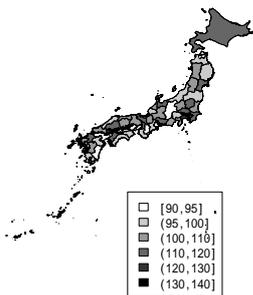
かつては、ESRI 社や DIVA-GIS などからシェイプファイルとして地図情報をダウンロードして `mapproj` パッケージで描画する必要があった。世界中のどこでも、シェイプファイルさえあれば地図にできるし、ない場合も、米国 CDC が無償で公開している EpiInfo に入っている EpiMap（に限らず多くの GIS ソフト）を使えば自作できる。

DIVA-GIS の場合は、Country:を Japan, Subject:を Administrative Areas (GADM) と指定すると、都道府県が黒、市町村が青線で縁どりされた図が表示され、Download というリンクをクリックすると、約 17 MB の zip 圧縮されたファイル (JPN\_adm.zip) を入手することができる。展開すると、JPN\_adm0.\*, JPN\_adm1.\*, JPN\_adm2.\*と、3つのレベルのシェイプファイル群が出てくるので、都道府県境界なら JPN\_adm1.\*, 市町村境界なら JPN\_adm2.\*を使えば良い。mapproj パッケージの `readShapePoly()` 関数で JPN\_adm1.shp などを読み込み、data スロットにふくらみ指数改を（先頭の全国データを除いて）マージし、塗り分ければ良い。

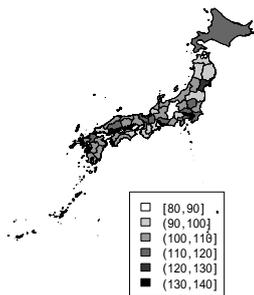
この方法を使っていた当時、DIVA-GIS の都道府県データはローマ字のアルファベット順に並んでいたもので、単純な `cbind()` などではマージできなかった。ふくらみ指数改を計算した後、ローマ字の都道府県名を変数名 PN として `data.frame()` を定義してから `merge()` 関数でマージしていた (`order(PN)` を [] の行の添字に使うってソートしておく必要がある)。DIVA-GIS に含まれる都道府県名変数 NAME\_1 の長崎にスペルミスがあって、“Naoasaki”となっていたため、そのままではうまくマージされず、PN の方を Naoasaki にすることで問題回避できた。

できあがったコードは <http://minato.sip21c.org/ldaR/fukuramikmap.R> である。DIVA-GIS から得た JPN\_adm1.\*と、<http://minato.sip21c.org/ldaR/japancensus2010tp.txt> を作業ディレクトリに置いて実行すると次の地図が得られる。

都道府県別ふくらみ指数(改)によるコロプレス図  
(男性, 2010年国勢調査)



都道府県別ふくらみ指数(改)によるコロプレス図  
(女性, 2010年国勢調査)



既に示した通り、日本の都道府県塗り分けならば、`mapdata` パッケージか `Nippon` パッケージを使うのが簡単である。2010年男性の「ふくらみ指数改」を使ったコロプレス図を描くには、次のようにする。

```
http://minato.sip21c.org/ldaR/PEImap.R

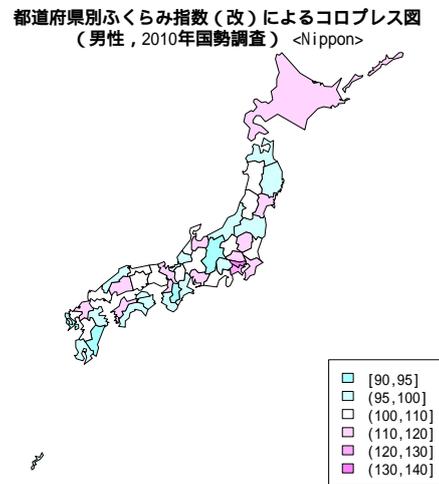
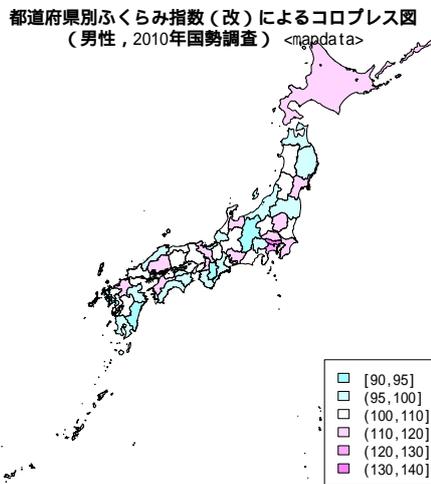
# データの読み込み
x <- read.delim("http://minato.sip21c.org/ldaR/japancensus2010tp.txt",
  encoding="CP932")

# ふくらみ指数改を計算する関数
source("http://minato.sip21c.org/ldaR/PEI.R")

# データから都道府県別の男性のふくらみ指数改を計算
malePEI <- as.vector(by(x$Males, x$Area, PEI, CLS=5))[2:48]

# ふくらみ指数改から6段階の塗り分け色
classes <- cut(malePEI, c(9, 9.5, 10:14)*10, include.lowest=TRUE)
cols <- cm.colors(7)[-1]

layout(t(1:2))
# using mapdata
library(mapdata)
# ふくらみ指数改と参照都道府県名からなるデータフレーム定義
PN <- c("hokkaido","aomori","iwate","miyagi","akita","yamagata","fukushima",
  "ibaraki","tochigi","gunma","saitama","chiba","tokyo","kanagawa",
  "niigata","toyama","ishikawa","fukui","yamanashi","nagano","gifu",
  "shizuoka","aichi","mie","shiga","kyoto","osaka","hyogo","nara",
  "wakayama","tottori","shimane","okayama","hiroshima","yamaguchi",
  "tokushima","kagawa","ehime","kochi","fukuoka","saga","nagasaki",
  "kumamoto","oita","miyazaki","kagoshima","okinawa")
map("japan", type="n")
for (i in PN) {
  map("japan", region=i, fill=TRUE, add=TRUE, col=cols[classes[PN==i]])
}
legend("bottomright", legend=names(table(classes)), cex=1, fill=cols)
title("都道府県別ふくらみ指数(改)によるコロナプレスマップ  
(男性, 2010年国勢調査) <mapdata>")
# using Nippon (definition of PN is unnecessary)
library(Nippon)
JapanPrefecturesMap(cols[classes], inset=FALSE)
legend("bottomright", legend=names(table(classes)), cex=1, fill=cols)
title("都道府県別ふくらみ指数(改)によるコロナプレスマップ  
(男性, 2010年国勢調査) <Nippon>")
```



### 3.6 介護者比

Rowland (2003) は、人口構造の指標として、介護者比 (Caretaker Ratio) を提案している。これは、介護を要する可能性が高い 80 歳以上の超高齢者人口の、主に介護行為を行う 50 歳以上 65 歳未満の女性人口に対する比である（100 を掛けた値で表示することになっている）。1990 年の UK では 46 である。

日本人口のデータについては、`fmsb` パッケージに実装した `CaretakerRatio()` 関数を使えば簡単に計算できる。男女別の 0 歳から始まる年齢各歳人口のデータが必要なので、例えば 1990 年と 2015 年の日本の介護者比の計算は以下の R コードで可能である。それぞれ、25.6 と 81.7 となり、日本の介護者比が急激に上昇したことがわかる<sup>\*19</sup>。

```
if (!require(fmsb)) { install.packages("fmsb"); library(fmsb) }
CaretakerRatio(Jpop$M1990, Jpop$F1990)
CaretakerRatio(Jpop$M2015, Jpop$F2015)
```

\*19 ちなみに、2005 年で 45.3 と 1990 年 UK と同程度であり、2010 年に 61.0 であった。

## 第4章

# 死亡の分析

### 4.1 はじめに

死亡は、さまざまな死因の集積による多面的現象である。このことは出生よりも広く認識されているが、同時に、年齢別死亡に何らかの規則性が存在することも古くから指摘されてきた。形式人口学における死亡の分析は、集団レベルでのそれを明らかにしようとする試みである。なかでも生命表による死亡の分析は、形式人口学のバックボーンとも言われる。

一方、実体人口学、例えば公衆衛生学や疫学の分野では、死亡は健康水準の最終的な指標である。乳児死亡率は対象集団の一般的な衛生状態や医療水準を反映するし、死因別死亡率を求めればその集団の死因構造において何が大きなリスク因子となっているのかを評価することができ、公衆衛生政策の立案の根拠となる情報を提供することができる。

人口転換にともなう死亡の高齢期への集中という（生存曲線の矩形化とも呼ばれる）現象を個人差の減少とみなせば、個人レベルの死亡現象と集団レベルの死亡現象がはっきりとつながっていることがわかるが、発生までの期間が長いために、個人レベルの現象としての死亡（発生した個体が死に至るまでの生物学的なプロセス）は、なかなか捉えにくい。

### 4.2 ヒトの死とは？

個々のヒトの死についてわかっていることは、誰でもいつかは必ず死ぬことや、個人個人で時期や原因が異なること（先天異常などで生まれつき死のリスクが高い人、直前までピンピンしていながら、突然、偶発的に交通事故や戦争で死ぬ人、百歳以上まで生き延びた後に老衰で死ぬ人、など）、以上2点から明らかのように、遺伝と環境の両方の影響を受けること、くらいであろう。

集団として考えるとどうだろうか。ある集団に属している人の死に方を表現する手段は

あるだろうか。例えば、一般に、途上国の人々は全体として先進国の人より若くして死んでしまう人が多いとはいえる。このことを、途上国は先進国より「死亡水準」が高いというためには、どうやって（即ち、どんな指標を使って）示せばいいのだろうか？

本稿では死亡データを分析して得られる代表的な指標について解説し、実際の計算例を示す。

### 4.3 死亡データの成立

日本では人口動態統計がしっかりしているので、年間の死因別死亡数が簡単に集計できるが、その仕組みは決して単純ではない。すべての死亡が死因とともに漏れなく届け出られ、登録されなくては、これは実現できないことなのである。医師法により、日本では、すべての死亡は届け出られることになっている。医師が看取った場合や、継続診療中の疾患で亡くなった後で死体を検査した（最終診察後 24 時間以内の場合は再度死体を検査しなくてもいい）場合は、医師が死亡診断書を発行する（医師法 20 条）。それ以外の死はほぼすべて「異状死」として扱われ、医師が死体を検査した後で死体検案書を発行し、24 時間以内に所轄警察署に届け出なければならない（医師法 21 条）。なお、東京都など監察医制度があるところでは死体検案は監察医によって行われる。

#### 実は死因は怪しいという話

日本では死後の解剖が義務付けられていないので、死後に解剖が実施される割合は 2% に過ぎず、死因の判定は主に医師が臨床診断として解剖せずにを行うことになるが、臨床診断は誤っている場合が 12-30% に及ぶという指摘が、作家の海堂尊（本職は病理医）によってなされている（海堂尊『死因不明社会』講談社ブルーバックス）。海堂は、この事実を踏まえ、死後に CT 等の画像診断をして真の死因を明らかにすること（画像診断だけでは確定できない場合には解剖を行う根拠ともなりうる）を提唱している。これはオートプシー・イメージング (Ai) と呼ばれ、全国各地で徐々に採用する施設が増えてきている。

通常、死亡診断書または死体検案書は、A3 横の死亡届書の右半分で作成される<sup>\*1</sup>。死因については、直接死因から遡って原死因（WHO の定義では、「直接に死亡を引き起こした一連の事象の起因となった疾病もしくは損傷」または「致命傷を負わせた事故もしくは暴力の状況」）まで記載することになっている。例えば、10 年前から糖尿病を患っていた人が 1 ヶ月前に肺炎に罹り、回復することなく 5 日前に膿胸となり、2 時間 30 分前に敗血症性ショックを起こして死に至った場合、直接死因は敗血症性ショックだが、原死因は肺炎となるため、死亡の原因を記載する I 欄には、(ア) 敗血症性ショック、(イ) 膿胸、(ウ) 肺炎と記載する。糖尿病は「直接には死因に関係しないが I 欄の傷病経過に影響を

<sup>\*1</sup> <http://www.mhlw.go.jp/toukei/manual/index.html> として厚生労働省の web サイトで記入マニュアルが公開されている。

及ぼした傷病名等」を記載する II 欄に記載する。

届出人（優先順位は同居の親族，同居していない親族，同居者，家主，地主，家屋管理人，土地管理人，公設所の長の順）は，死亡届書の左半分に記入押印し，市町村役場（優先順位は，死亡者の本籍地，死亡地，届出人の現住所地の順）の戸籍課に提出する（ただし，通常，死亡届書の記入と提出は葬儀社により代行されている）。戸籍課では死亡者本人の戸籍の抹消と火埋葬許可証の交付を行い，死亡届書を 1 ヶ月保管した後，地方法務局へ送付する。

一方，人口動態調査令（厚生労働省令の 1 つ）により，厚生労働大臣は，人口動態調査票の用紙を保健所長経由で各市町村長に交付することになっている。市町村長は，上記死亡届書が提出されたらすぐに，人口動態調査票を作成し，人口動態調査票市町村送付票を添えて管轄保健所長に提出する（人口動態調査令施行細則）。保健所長はこれに基づき死亡小票を作成し 3 年間保存するとともに，1 ヶ月ごとに取りまとめ（前月 15 日から当月 14 日まで），人口動態調査票保健所送付票を添えて，当月 25 日までに都道府県知事（ただし政令指定都市や中核市に設置された保健所では市長）に提出する。都道府県知事（あるいは市長）は，これを取りまとめ，人口動態調査票都道府県送付票を添えて翌月 5 日までに厚生労働大臣に提出する。人口動態調査票を厚生労働大臣が保管する期間は翌年 1 年間だが，電子化されたデータは永久に保管することになっている。

上記のプロセスを経て，日本人すべての死亡データは厚生労働大臣のもとに集まり，人口動態統計として集計され報告されることになる<sup>\*2</sup>。そのおかげで，年齢別死因別死亡数といった数字を得ることができる。しかし，疫学研究などで個人ベースの詳細な死亡データが欲しいときは，死亡小票の目的外使用を申請し，総務大臣の許可を得てから，保健所で閲覧する必要がある，大変面倒である。平成 19 年の統計法改正を受けて学術会議から提言（次の枠内に一部引用）が出たが，いまだに改善されていない模様である。

---

<sup>\*2</sup> 人口動態統計では，原死因を死因として集計する。原死因は上記 I 欄のもっとも下に書かれているものとなるのが原則だが，それが呼吸不全，心不全など死因としてふさわしくないと ICD-10 で WHO が定めたものになっている場合は II 欄の傷病を原死因とする。

## 学術会議分科会提言

(5) 日本版 **National Death Index** の創設と運用

疫学研究においては、一定集団について長期間観察を行うことによってはじめて結果が得られることが多く、人の生死情報がしばしば必須であり、人口動態統計との連結なしには科学的な結論が見出せない。これまで、人口動態統計使用に当たっては、厚生労働省を経由し、総務省に目的外使用に関する申請を行い、認可を得て情報を入手することとなっていた。申請から認可までに要した期間は、平均 11.8 か月（最短で 3 か月、最長は 23 か月）であり、申請書等の修正回数は平均 10.2 回（最少で 0 回～最多は 60 回）であった（日本疫学会将来構想検討委員会「政府統計の利活用に関する調査報告書」平成 19 年 11 月 30 日、[<http://www.soc.nii.ac.jp/jes/news/pdf/20071206sei fu.pdf>]）。これに比し、欧米諸国では健康情報登録（疾病登録、患者登録等）が整備され、死亡データとのリンケージも容易である。例えば、米国では、National Death Index システムを用いて、氏名、性、生年月日、住所、社会保障番号などを指標とする死亡ファイルとの記録照合による追跡調査が容易に可能である。このシステムを利用する場合、利用申請書のフォーマットや、それを作成するための詳細なマニュアル等が整備されており、申請処理期間は約 2 か月であり、多くの医学研究がこのシステムを用いて行われ、大きな成果をあげている（National Health Statistics. National Death Index. [<http://www.cdc.gov/nchs/ndi.htm>]）。これまで我が国の死亡統計ファイルには氏名が入力されていなかったため、このような死亡ファイルとの記録照合は不可能であった。しかし、2003 年 1 月より厚生労働省が進めている人口動態オンライン報告システムでは、死亡小票氏名が入力されるようになり、このシステムを充実させて全国的な運用を図れば、日本版 National Death Index の基盤が整うこととなる。統計情報の高度利用の一環として、死亡統計ファイルとの記録照合の技術的および制度的な検討を早急に開始し、日本版 National Death Index システムの構築とそれを活用するための手順の標準化・マニュアル化を進めるべきである。これらのシステムの円滑な運用によって、我が国の保健医療分野における疫学研究が飛躍的に進展し、国民の健康と安全のための有益な種々の施策提案に結びつくことが期待される。

出典：提言「保健医療分野における政府統計・行政資料データの利活用について—国民の健康と安全確保のための基盤整備として」、pp.10-11, 平成 20 年（2008 年）8 月 28 日, 日本学術会議基礎医学委員会・健康・生活科学委員会合同パブリックヘルス科学分科会

<http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-20-t62-6.pdf>

欧米先進諸国では概ね日本よりも信頼性の高い死亡データが得られる。地域がん登録の登録率も高いし、電子化も進んでいる。死因確定のための剖検割合も日本より高い国が多い。

途上国では事情は一変する。パプアニューギニアやソロモン諸島では、そもそも死亡事象を行政が把握するまで年単位の遅れが存在する場合さえあるし、死因も必ずしも医師によって診断されない。個人ベースの死因を把握しようと思ったら、綿密な聞き取り（ただしクロスチェックは必須）によって再構成する方が、公式統計よりもむしろ信頼性が高い場合さえある。

## 4.4 死亡のプロセス

個体に現実に起こる死亡のプロセスは、死亡をもたらす要因を最小限にとどめたとしても生物学的必然として起こる場合（一般に言う、天寿を全うする場合）の他に、物理化学的要因や生物学的要因などさまざまな要因からなる病気によって死がもたらされる場合や、あるいは交通事故や天災によって突然訪れる死もある。同じ病気に罹っても保健医療水準が高い先進諸国と、高度医療がほとんど受けられない途上国とでは死に至る可能性は異なることが普通である。

社会的な要因は、治療だけではなく、物理化学的要因や生物学的要因との接触、高脂肪食を中心とする食生活の変化まで含めて、病気をもたらすリスクにも影響する。近年では、社会疫学という研究分野がクローズアップされ、隣人がどれくらい信頼できるか、社会組織にどれくらい参加しているか、といった、ソーシャル・キャピタルと呼ばれる要因が健康状態に及ぼす影響が注目されている（このことは、2009年12月22日の厚生政策セミナー「長寿革命」でも、複数の演者が指摘していたが、例えば『社会格差と健康』（川上ら編、2006）などが参考になる）。

多くの社会では相対的に若齢では事故死が、高齢では病死が多いが、日本では中年層の自殺が多いことが目立つ。自殺も、かつてのブラジルで砂金掘りのガリンペイロたちが精錬のための水銀蒸気曝露によって中枢神経系に異常をきたし、結果として高い自殺率を示したような物理化学的要因もあるかもしれないが、社会的な要因の影響が強い死因である。

### 4.4.1 寿命とは何か？

一般に寿命といえば、事故がなかった場合に、そのものが機能停止するまでの時間をさすように思う。何らかのメカニズムで老化が進行すると、やがて機能停止に至ると考えると、寿命を考えることは老化を考えることと表裏一体である。老化がなぜ起こるかということについては、体細胞廃棄説など諸説紛々であり、結論は出ていない（杉本、古市、1998）。アロメトリーの成果によれば（たとえば Charnov, 1992; 本川, 1992）、生まれてから死ぬまでに心臓がドクンと鼓動する回数や、呼吸する回数といったものは、哺乳動物にほぼ共通していることが知られており、代謝速度を基準に考えれば、寿命はどの哺乳動物でも一定と考えられることが指摘されている。テロメアの長さという点から考えると、体細胞の分裂回数で寿命が決まってくるとも考えられる。

いっぽう、人口学で寿命といえば、普通は平均寿命をさす。人口学の専門用語ではゼロ歳平均余命と呼ぶ。「平均」といっても死亡年齢の算術平均をとるわけではなく、同時に

生まれたゼロ歳児が10万人いたとして、その人たちがある時点の年齢別死亡率に従って死んでいった場合に、平均何年生存するかという値である。言い換えると、平均寿命は年齢別死亡率を静止人口モデルで計算した生存年数の期待値であり、計算には生命表を作ることが普通である（連続量として生存時間解析を使うこともできる）。死因が何であるかには関係なく年齢別死亡率だけで計算されるため、戦争があったり大災害があったりすれば、平均寿命は短くなる。その意味で平均寿命は、遺伝と環境の両方の影響を受ける指標である。

#### 4.4.2 ヒトはなぜ長寿なのか？

生物の世界を見渡すと、ヒトは少数の子どもを産んで大事に育てる、いわゆるK戦略者である。K戦略をCharnov (1992) の生活史戦略の観点に立って見直すと、寿命が長くなるような戦略と考えることもできる。

先にもあげた体細胞廃棄説とは、何十億もの体細胞すべてのエラー修復を長い間続けることは非効率なため、生殖細胞だけのエラー修復をして、そこから再び個体発生をして次世代を構築することで全体としてのエラー蓄積を避けることができるわけであるが、その代償として廃棄すべき体細胞にはエラーがたまっていくのは避けがたく、それが老化なのだ、という説である。この説で考えると、K戦略は、生殖細胞よりも体細胞にエネルギーを振り分ける戦略と説明できる。必然的に寿命は長くなるはずである。

しかし、次世代の生産を終えた閉経後の生存が次世代の生存に寄与しないなら、閉経後の生存は淘汰を受けないことになり、遺伝的な影響はないことになる。したがって、同じように50歳で閉経するならば、寿命が60歳でも100歳でも子ども数とは関係しないわけで、閉経ぎりぎりまで体細胞がもつように振り分けをコントロールするのが最も適応的と思われる。しかし、現在の先進諸国では、どういうわけか閉経後もかなりの期間、生存するのが普通であり、日本を始めとして平均寿命が延びつづけている国がいくつもある。なぜ包括適応度に寄与しないはずの生存期間が延長しつづけるのかは大きな謎であり、いくつかの仮説が提唱されている。

#### 4.4.3 おばあちゃんの孫育て仮説

ヒトの場合は子どもが再生産年齢に達するまでの養育コストが大きいので、孫や曾孫を育てるという形による閉経後の家系の生存への寄与は無視できない、ということに着目したユタ大学のホークスは、チンパンジーと再生産期間はそれほど変わらないのにヒトの再生産完了後の生存年数が長いのはおばあちゃんが忙しい娘の子育てを手伝うためである、という説を唱えている (Hawkes et al., 1998; Alvarez, 2000)。彼らの説を簡単に紹介する

と、以下のようにまとめられる。

長い閉経後の生存期間は、他のどんな霊長類にすぎない、ヒト独自の特徴である。このパターンは、母子間の食物の共有という、年をとった女性が娘の出生力を強め、それによって老化に対抗する淘汰圧を大きくする行為を進化させてきたかもしれない。哺乳類の生活史についての Charnov (1992) の無次元の集合法則と組み合わせれば、この仮説は我々の成熟が遅いこと、離乳時の身体が小さいこと、そして（霊長類としては例外的といえるほど）高い出生力を説明する（次の表を参照； $\alpha M$  がどの哺乳動物でも一定というのが Charnov の法則）。このことは、過去のヒトの生存環境選択と社会組織に対して、またヒトの進化における学習能力の拡大と父親による食料調達の重要性を示唆するものである。

種	成体期間 <sup>(1)</sup>	成熟年齢	離乳年齢	$\alpha$ <sup>(2)</sup>	$\alpha M$	体重比 <sup>(3)</sup>	娘出産率 ( $b$ )	$ab$
オランウータン	17.9	14.3	6.0	8.3	0.46	0.28	0.063	0.52
ゴリラ	13.9	9.3	3.0	6.3	0.45	0.21	0.126	0.79
チンパンジー	17.9	13.0	4.8	8.2	0.46	0.27	0.087	0.70
ヒト	32.9	17.3	2.8	14.5	0.44	0.21	0.142	2.05

(1)  $1/M = 0.4 * \omega - 0.1$  ただし、 $M$  は平均成人死亡率、 $\omega$  は理論上の最高年齢 (Charnov, 1992)

(2) 離乳から成熟するまでの期間

(3) 離乳時の体重の成人の体重に対する比

#### 4.4.4 繁殖成功を犠牲にした長寿仮説

老化の原因としての体細胞廃棄説によれば、長寿に寄与する体細胞のエラーを嚴重にすることは生殖細胞が利用可能なリソースを減らすので、代償として子ども数が減ることは想像に難くない。実際ショウジョウバエなどでは以前よりよく知られていた。

Westendorp ら (1998) が、17 世紀から 19 世紀のイギリス貴族の家系データにより、死亡年齢別の子ども数を集計したところ、60 歳までは長生きするほど子ども数が多かったが、70 歳、80 歳と長生きになるほど子ども数が減っていた。このことはヒトでも体細胞のエラーを減らす傾向が強いほど（つまり DNA 修復などの体細胞維持作用が強いほど）、妊孕力が低いことを示唆する。これは、ヒトでも体細胞廃棄説が成り立っていることを間接的に支持する証拠である。

長生きする方がより蓄財できるとすれば、それが子どもの生存に有利に働くことによって結果的に包括適応度を上げるならば、長寿に寄与する遺伝子が自然淘汰によって残っていく可能性はある。もしそうならば、少子化社会は長寿の所産として必然なのかもしれない。

#### 4.4.5 「課題と挑戦」仮説

上記 2 つのような進化的な説明ではないが、Wilmoth J が提唱しているのが（注：2009 年 12 月 22 日、厚生政策セミナーでの基調講演による），“Why mortality falls over time?”

という問いに対して、あらゆる種類の Death に対しての human population の対応が、Recognition（認識）、Reaction（対策）を経て、Reduction（減少）へと至り、やがて別の原因による Death が主要死因になると、改めてこの“3R”プロセスが起こり、それが繰り返されることによって、死亡水準は永遠に低下し続けるという仮説である。Wilmoth 自身は“3R-theory of mortality decline”と呼び、堀内四郎は「課題と挑戦」仮説と訳している。

なるほどと思わされるが、この仮説の弱点は、地球上のリソースの有限性を考えていないところである。リソースが枯渇したらそれ以上の対策はとれないので、やはり寿命延長には限界があると考えるのが自然であろう。

## 4.5 死亡指標

いよいよ、さまざまな死亡データを分析して得られる、集団の死亡水準を示す指標値について説明する。どんな死亡指標を計算できるかは、研究方法に依存する。すべての死亡を記録しておき、1年ごとにその数を層別に集計するのが基本であるが、歴史的コホートとして出生年ごとに集計する方法もあるし、前向きコホートとして集計することもある。

また、高齢者の方が若齢者より死にやすいのが普通なので、年齢構造の異なる集団間では、単純に死亡数だけで比較することには問題があり、年齢構造で補整<sup>\*3</sup>する方法がいくつか提案されている。

### 4.5.1 年間死亡数から推計できる指標

まず用語説明を含め、年間死亡数から推計できる指標についてざっと説明する。

**CDR** Crude Death Rate の略。普通死亡率または粗死亡率という。率であるから、本来は延べ生存人年あたりの死亡数を求めるべきであり、それによって次元が（/年）となる。けれども、通常は単純化して1年間の人口変化が線型であると仮定し、分母は年央人口（7月1日現在の人口）で代替する。さらに、日本の人口動態統計では、

$$\frac{\text{年間総死亡数}}{\text{その年の10月1日人口}} \times 1000$$

と定義されている。10月1日人口を用いるのは、国勢調査人口が10月1日現在の人口だからである。一般に1年ごとに求めるが、小規模な集団の場合は、複数年のデータから平滑化して補整することもある。

**ADR** Age-specific Death Rate の略（Chamberlain, 2006）。Age-Specific Mortality Rate の略で ASMR と表記する場合もあるが同じものである。年齢別死亡率を指す。分母

<sup>\*3</sup> 補正と書くこともあるが、人口学でこの意味で使う場合は、補整と書くのが普通である。

が年齢別の延べ生存人年（と考えると、実はいつの時点での年齢かという難しい問題があることがわかるが、ここではとりあえずそこまでは考えずにおく）であるが、その近似値として年齢別年央人口をとり、ADR の定義としては

$$\frac{\text{年齢別年間死亡数}}{\text{年齢別年央人口}}$$

となる。けれども、日本の人口動態統計では、粗死亡率と同じく、

$$\frac{\text{年齢別年間死亡数}}{\text{10月1日の年齢別人口}}$$

と定義している。通常は5歳階級だが、5歳未満は1歳階級にする場合が多い。年齢別死亡率は、後述する生命表解析にも、他の指標を計算する際にも基本となるので、きわめて重要である。

**DSMR** Directly Standardized Mortality Rate の略。邦訳は直接法訂正死亡率、直接法年齢調整死亡率、直接標準化死亡率などいろいろあるが同じものを指す。日本の官庁統計で、たんに「年齢調整死亡率」と表記されている場合は、間接法ではなく、こちらを指す。

$$\frac{\sum \text{標準人口年齢別人口} \times \text{ADR}}{\text{標準人口}} \times 1000$$

と定義されている。標準人口は、比較したい集団全部に対して共通の年齢階級別人口を用いる。高齢者が多い集団は、ADR の高い年齢層が多いので、たとえ同じ年齢層で比較した場合の ADR が同じでも、若者が多い集団に比べて総死亡数が多くなるため、CDR は高くなってしまう。人口構造におけるこうした年齢の偏りが CDR に及ぼす影響を調整するため、特定の人口構造を重みとした ADR の重み付き平均をとった値が、この DSMR である。考え方も素直でわかりやすいが、対象とする集団の ADR がわからないと計算できないのが欠点である。

**SMR** Standardized Mortality Ratio の略。日本語では標準化死亡比という。

$$\frac{\text{年間総死亡数}}{\sum \text{年齢別年央人口} \times \text{標準人口の年齢別死亡率}}$$

と定義される。対象集団の年齢別死亡数が未知でも計算できるのが利点である。

**ISMR** Indirectly standardized mortality rate の略。間接法訂正死亡率、間接法年齢調整死亡率、間接標準化死亡率などと呼ばれる。

$$\frac{\text{CDR} \times \text{標準人口} \text{ CDR} \times \text{年央人口}}{\sum \text{年齢別年央人口} \times \text{標準人口 ADR}}$$

という式で表わされる（岡崎陽一「人口統計学」）。

しかし実は

$$\text{標準人口 (CDR)} \times \text{SMR}$$

と同値である（国際人口学会編，1994；Smith 1992）。従って，**SMR**と同様に，対象集団の年齢別死亡率がわからなくても計算できる利点をもつ。

**平均寿命** 英語では Life expectancy という（Average life span という言い方もある）。より正確には，ゼロ歳平均余命と言った方がいい。前述の通り，死亡年齢の平均値ではなく，ある年の ADR に従ってゼロ歳人口 10 万人が死んだ場合に，平均して何年生きるかという値である。グラフィカルには，縦軸に生残率（ゼロ歳のときに 1.0）をとり，横軸に年齢をとって生存曲線を書いた場合の，曲線下面積として把握することができる。計算方法については，5.3「生命表解析」を参照されたい。

**健康余命** 英語では Health expectancy または Healthy life expectancy とよぶ。健康の定義も多様なため，これも多様な定義が可能だが，一般には，日常生活自立度が十分に高い状態で生存している人だけを生存と数えた場合の生命表解析により得られる。

**HALE** Health Adjusted Life Expectancy。QALY（Quality Adjusted Life Years）とも似た考え方で，健康状態を調整した余命の計算値。

**DALE** Disability Adjusted Life Expectancy。DALY（Disability Adjusted Life Years）と同じ（というか裏返しの）考え方で，障害をもって生きる 1 年と障害なく生きる 1 年の価値を等しくないと考え，障害の程度によって調整した余命の計算値だが，一歩間違えば障害者差別につながりかねないので注意が必要である。

**PMI** Proportional Mortality Index あるいは Proportional Mortality Indicator の略。50 歳以上死亡数を総死亡数で除した値であり，国際比較によく用いられる。

**死因別死亡率** 特定死因による年間死亡数を年央人口で除して 100000 を掛けた値を指す。疾病による死亡の場合，死因が単独であることは少なく，複合死因である場合が多いことが問題である。通常，疾病の死因分類は，ICD（International Statistical Classification of Diseases and Related Health Problems；国際疾病分類）によって行われる。現在のところ，第 10 版（ICD-10）が最新である。死因分類が改訂されることは，死因別死亡率の経年変化を見る場合には注意しなければならない<sup>\*4</sup>。日本では 1981 年以降一貫して悪性新生物死亡率がトップである。近年では悪性新生物死亡数は 34 万を超え，心疾患死亡数も 18 万を超えて，さらに増加を続けている。それに対して脳血管疾患死亡数は 12 万 6 千となり，わずかに減少している。

<sup>\*4</sup> 日本において，1995 年から心疾患死亡率が低下した最大の理由は，ICD-9 の死因簡単分類から ICD-10 の死因分類が変わったことによって死亡診断書の様式が変わり，「疾患の終末期の状態としての心不全，呼吸不全などは記載しない」とされたために，それまで心不全とされて心疾患死亡を押し上げていた死因が別の分類になったためである。

通常、疫学的転換の過程において、年齢調整した死因別死亡率の変化を追ってみると、まず感染症による死亡率が低下し、次いで心疾患による死亡率が低下し、最後のがんの死亡率の低下が始まる。がんによる死亡率が低下するのは、喫煙率の低下によって半分が説明できるという報告がある (Staetsky, 2009)。また、がんを胃がん、子宮頸がん、肝臓がんという感染性の病原体が関係するがんと、それ以外のがんに分けて、それぞれの年齢調整死亡率の年次推移をみるとパターンが異なる (Wilmoth ら, 未発表, 2009 年 12 月 22 日, 厚生政策セミナー「長寿革命」での基調講演による)。

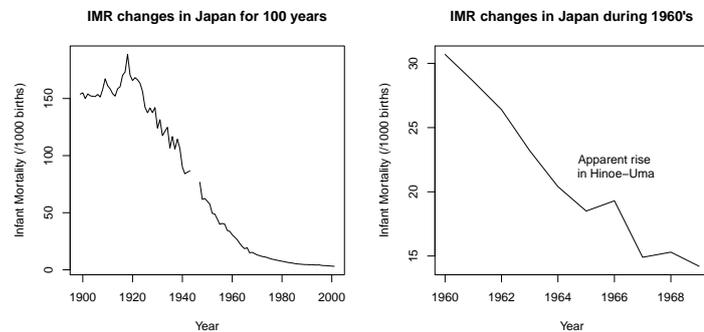
**死因別死亡割合** 特定死因による年間死亡数を年間総死亡数で除した値である。全死因のうち、どの死因による死亡が相対的に多いのかを示す。PMR (Proportional Mortality Ratio) ともいう。増減はその疾患の増減だけでなく、他の疾患の増減とも連動する。

**YLL** Yeas of Life Lost の略で、日本語では損失余命と呼ばれる。個々の死因について、生命表で、その死因による死亡がなかったら平均余命がどれだけ延びるかを計算して、その死因による損失余命と考えるのが、蒲生らの方法である (別に Graham らが開発した方法もある)。R では Epi パッケージの `erl()` 関数の `y11=TRUE` オプション<sup>\*5</sup>で計算できる。世界各国のデータについては、DALY パッケージで Global Burden of Disease 2010 (GBD2010) に基づいた計算結果を参照できる。

**IMR** Infant Mortality Rate の略で、乳児死亡率を意味する。出生 1000 当たりの、生後 1 年未満の死亡件数。分母に死産は含まないし人口でもないことに注意。たんに死亡水準の指標というよりも、とくに途上国で衛生状態や栄養状態を敏感に反映する指標なので、公衆衛生学的に重要である (生活文化水準を反映する指標ともいえる。日本はスウェーデンやスイスと並んで、世界で最も乳児死亡率が低い国の 1 つであり、日本での乳児の死因は、1979~1984 年のみ出産時外傷等が 1 位だが、その後は一貫して先天異常が 1 位である)。なお、分母と分子の期間がずれていることにも注意すべきである。つまり、実際の計算は、ある年の生後 1 年未満の死亡をその年の出生数で除して 1000 を掛けることによって行われるので、分子には前年に生まれた子供も含まれるし、分母には翌年亡くなる子供も含まれる。それが相殺されると見なして、ある年の指標値として使うのである。このことから考えると、例えば日本の丙午のように、極端に出生数が少ない年があると、その年の乳児死亡率は過大評価される可能性がある<sup>\*6</sup>。

<sup>\*5</sup> 考え方の詳しい説明が <https://cran.r-project.org/web/packages/Epi/vignettes/y11.pdf> にある。

<sup>\*6</sup> 実際、その前後はずっと低下傾向にあるのに、1966 年 (丙午) の乳児死亡率はわずかに上昇している。これが見かけの上昇であることは、30 年も前に指摘されている (近藤, 1980)。



**新生児死亡率** 生後4週未満の死亡を新生児死亡という。新生児死亡率は出生1000当たりの新生児死亡数である。

**早期新生児死亡率** 生後1週未満の死亡を早期新生児死亡という。早期新生児死亡率は出生1000当たりの早期新生児死亡数である。これは母体の健康状態の影響と、NICUの普及度など医療水準の影響が大きい。

**周産期死亡率** 周産期死亡については、1995年から新しい定義「妊娠満22週以後の死産と生後1週未満の早期新生児死亡を合わせた死亡」（旧定義は22週のところが28週）が採用されている。妊娠後期の死産と早期新生児死亡が母体の影響を受けやすいことと、途上国では早期新生児死亡が死産扱いされることが多いため、そこで誤分類があっても影響を受けない指標となる。現在の周産期死亡率は、ある年の出生数と妊娠22週以後の死産数の和を分母として、その年の早期新生児死亡数と妊娠22週以後の死産数の和を分子として、1000をかけた値である。最近の日本は旧基準では約3、新基準では約5である。

**幼児死亡率** 1～4歳の死亡を幼児死亡という（英語では Toddler Mortality）。2005年の日本の幼児死亡数を分子として、2005年日本の年央人口の1～4歳人口を分母として10万を掛けた値は25.4である（1999年には約33だった）。幼児死亡の死因は不慮の事故（欧米に比べて多い）や先天異常が多い。乳児死亡と幼児死亡をあわせて5歳未満死亡という概念も、全般的な衛生状態の指標としてよく使われる。

**妊産婦死亡率** 妊娠、分娩、産褥に直接関連する疾病や異常によって母性が死亡した場合を「妊産婦死亡」または「母性死亡」という。妊産婦死亡率は出産または出生10万当たりの妊産婦死亡数で表す。2006年日本は出生10万あたり4.9（1985年に15.8、1999年には6.1）できわめて低い。主な死因は出血と妊娠高血圧症候群である（PIHと略記される。かつて妊娠中毒症と言われていたもの。2005年4月から変更。日本産婦人科学会の定義は「妊娠20週以降、分娩後12週まで高血圧が見られる場合、または高血圧に蛋白尿を伴う場合のいずれかで、かつこれらの症状が単なる妊娠の偶発合併症によるものではないものをいう」）。

### 4.5.2 遡及聞き取りデータからの推計法とその指標

計数データで推計できる指標は、遡及聞き取りデータからも推計できる場合が多い。集団サイズが小さい場合には、1年間では十分な代表性をもつ死亡データが得られない場合が多いので、遡及聞き取りによって何年分かをまとめて分析することも可能である。その他に、死亡年齢がわかる場合は、もっと精密な死力の分析が可能になる。

ただし、生没年を記録する慣習のある宗教の信者である場合（歴史人口学では、日本では宗門人別帳や過去帳、欧米ではキリスト教の教区簿冊を使う）を除き、遡及聞き取りで得られる年次情報の信頼性は低い。信頼性を上げるための工夫については、中澤 (2007) などを参照されたい。

### 4.5.3 平均余命と生命表

かつての先進国や現在の途上国のように、ゼロ歳平均余命が50年であるような状態から70年まで上げるときに大きく寄与するのは乳幼児死亡の減少であり、そのためには母子保健や周産期医療の改善、一般的な衛生状態や栄養状態の改善が有効であることは、多くの国の歴史が示してきた通りである。一方、ゼロ歳平均余命を70年から80年やそれ以上に上げるときに大きく寄与するのは高齢者の死亡率低下であり、基本的には高度医療や延命医療を含む医療水準の超高度化と、健康増進・疾病予防の推進である。

延べ生存期間の総和は、年齢別死亡率  $q_x$  が変化しないとして、 $l_x(1 - q_x/2)$  によって  $x$  歳から  $x+1$  歳まで生きた人口  $L_x$ （開始時点の人口が決まっていれば死亡率も変化しないので  $x$  歳の静止人口と呼ばれる）を求め、それを  $x$  歳以降の全年齢について計算して和をとることで得る。グラフィカルな見方をすれば、0歳における生残率を1とした生残率曲線の積分値（曲線下面積そのもの）ともいえる。ゼロ歳平均余命の延伸にともなって、かつては三角形 (triangular) に近かった (De Moivre の公式が成り立つような) 生残率曲線が矩形化 (rectangularization) することがよく知られている (Kannisto, 2000)。

平均寿命とは0歳平均余命のことだが、これは、ある時点での年齢別死亡率（後述する  $q_x$ ）に従って、ゼロ歳児10万人が死んでいったとすると、生まれてから平均してどれくらいの期間生存するのかという値である\*7。

一般に  $x$  歳平均余命は、 $x$  歳以後平均してどれくらいの期間生存するのかという値なので、 $x$  歳以降の延べ生存期間の総和  $T_x$  を  $x$  歳時点の個体数  $l_x$  で割れば得られる。延べ生

\*7 誤解されることが多いが、死亡年齢の平均ではないので注意されたい。また、死亡年齢の中央値である半数死亡年齢は、ゼロ歳折半余命（または静止人口中位数年齢）という名前がある。ともに平均寿命とは別の指標値である。

存期間の総和は、ちょうど  $x$  歳に達した者が  $x+1$  歳に達しないで死亡する確率、すなわち年齢階級  $[x, x+1)$  における死亡率  $q_x$  が変化しないとして、 $l_x(1-q_x/2)$  によって  $x$  歳から  $x+1$  歳まで生きた人口  $L_x$  (開始時点の人口が決まっていれば死亡率も変化しないので  $x$  歳の静止人口と呼ばれる) を求め、それを  $x$  歳以降の全年齢について計算して和をとることで得られる。

ヒトの人口学では、通常の年齢別死亡率  $m_x$  (ある年に  $x$  歳で死亡した人数  $d_x$  をその年の  $x$  歳年央人口で割った値) から  $q_x$  を  $q_x = m_x/(1+m_x/2)$  として求めて生命表を計算する。この式の意味は次の通りである。

1. 年央人口で  $x$  歳だった集団  $N_x$  の実際の年齢は、 $x$  歳以上  $x+1$  歳未満である。
2. もし死亡が一定速度で起こるなら、期首人口で  $x$  歳だったのは  $N_x + d_x/2$  人となるはずである。
3. この人たちが  $x+1$  歳になるまでに  $d_x$  人死亡することになるので、死亡率  $q_x$  は、

$$q_x = \frac{d_x}{N_x + d_x/2} = \frac{d_x/N_x}{N_x/N_x + d_x/2/N_x} = m_x/(1+m_x/2)$$

となる。

4. ゼロ歳のところは短期間での死亡率の変化が大きいのので、最初の1ヶ月は1週間ずつ、次いで2ヶ月、3ヶ月、6ヶ月と刻んで、1歳未満を7つの階級に分けて計算するのが普通である。
5. また、年齢5歳階級で計算する生命表 (abridged life table) を作ることがあるが、その場合は死亡の線型性を仮定するのは無理なので別の補正を使う。よく用いられる Greville の方法は、

$${}_5q_x = \frac{{}_5m_x}{1/5 + {}_5m_x \left[ 1/2 + 5/12 \left[ {}_5m_x - \frac{\ln({}_5m_{x+5}) - \ln({}_5m_x)}{5} \right] \right]}$$

である (Ng and Gentleman, 1995)。この式は、5歳階級での通常の年齢別死亡率の自然対数をとった値の傾き  $\left( \frac{\ln({}_5m_{x+5}) - \ln({}_5m_x)}{5} \right)$  が  ${}_5m_x$  と等しければ、上の式と一致する\*8。

生物一般について考えるときは、同時に生まれた複数個体 (コホート) を追跡して年齢別生存数として  $l_x$  を直接求めてしまう方法 (コホート生命表) とか、たんに年齢別個体

\*8 Greville の式については、文献によって表記方法が若干異なるが、この、Ng and Gentleman (1995) による説明が一番わかりやすかった。オリジナルの Greville TNE (1943) Short methods of constructing abridged life tables. *Rec. Am. Inst. Actuar.*, 32: 29-43. が入手できないので確認できないのだが、館 (1963) によれば、

$${}_nq(x) = \frac{{}_nm(x)}{\frac{1}{n} + {}_nm(x) \left[ \frac{1}{2} + \frac{n}{12} \{ {}_nm(x) - \ln c \} \right]}$$

で、 $c$  は  ${}_nm(x)$  が Gompertz 法則に従うものとして  ${}_nm(x) = Bc^x$  としたことにより、US の経験から  $\ln c$  の

数を  $l_x$  と見なしてしまう方法（静態生命表，偶然変動で高齢の個体数の方が多い場合があるので平滑化するのが普通）がよく行われる。ヒトの場合は個体数変動が大きいので，現在人口しかわからないような無文字社会での短期的な研究でも，年齢別人口を少なくとも2回調べ，その差から年齢別死亡率を推計して死亡率から計算する（Timothy Gage や William Brass が開発した方法で，Two census method と呼ばれる）。人口移動が少ない場合は精度の良い方法である。

他には，Gompertz 曲線を使って補正したり<sup>\*9</sup>，Siler モデルを使って補正する方法もある（非線形最小二乗法でパラメータ推定すればよい。R での非線形最小二乗法は，`nls()` 関数や `optim()` 関数で実行できる）。なお，Siler モデルの式は，

$$h(t) = a_1 \exp(-b_1 t) + a_2 + a_3 \exp(b_3 t)$$

である。

また一方， $l_x$  から

$$\mu_x = -\frac{1}{l_x} \frac{dl_x}{dx}$$

として求められる，ちょうど  $x$  歳における瞬間の死亡率  $\mu_x$  は死力と呼ばれる。これは，次に述べる，個人ごとの生存年数をコックス回帰する場合のハザード関数に他ならない。

#### 4.5.4 生存時間解析

個人ごとの死亡までの間隔データ，つまり死亡年齢を用いれば生存時間解析を行うことができる。集団の死亡パターンは個人個人の frailty（生まれつきの脆弱さ），あるいは病気

---

値は 0.080 ~ 0.104 で，通常 0.09 が用いられるとされている。一方，和田 (2006) では，

$${}^nq_x = \frac{{}_n m_x}{\frac{1}{n} + {}_n m_x \left[ \frac{1}{2} + \frac{n}{12} \{ {}_n m_x - \log_e \left( \frac{{}_n m_x + n}{{}_n m_x} \right)^{\frac{1}{n}} \} \right]}$$

と書かれている。これは Ng and Gentleman とまったく同値である。生命表解析の専門書である，Namboodiri and Suchindran (1987) によれば，

$${}^nq_x = \frac{{}_n M_x}{(1/n) + {}_n M_x [(1/2) + (n/12)({}_n M_x - k)]}$$

かつ  $k$  は生命表によってわずかに異なるかもしれない定数で，0.09 に等しいとしても大きな誤差はないだろうとされている。数理人口学者 Keyfitz による “Applied Mathematical Demography” では，

$$\int_0^n \mu(x+t) dt = n {}_n m_x + \frac{n^3}{12} {}_n m_x^2 (\log_n m_x)'$$

が Greville の良く知られた結果であると書かれているが， $n^3$  の部分が他の式と違っていて，何を意味するのかよくわからないが，それを除けば，どの本も違いがあるのは 1カ所，そこをどの程度近似的な表現にしているかの違いなのではないかと思う。

<sup>\*9</sup> <http://www.toukei.metro.tokyo.jp/seimei/2005/sm-gaiyou.htm> に示されている東京都の生命表は，高齢者の死亡を Gompertz 曲線で補正している。ここに書かれている解説は丁寧で参考になる。

や死亡の特定原因への感受性の分布を反映して変わってくるので、これらの変数と死亡年齢の関連を分析することによって、死亡水準を、集団レベルの分析だけの場合よりも詳細に把握することができる。

古典的な統計手法としては、ロジスティック回帰分析やダミー変数を使った回帰分析があるが、死亡の場合は事象発生までの期間が長いために、これらの手法では観察期間中にエンドポイントが得られないとか、観察開始時が特定できないといった問題があった。これら打ち切りデータを分析から除外してしまうと、相対的に事象発生までの期間が長めのデータで除外が生じやすくなるため、推定される半数死亡年齢を過小評価してしまう危険がある。

この問題を解決したのがカプラン=マイヤ法による生存曲線の推定やハザード解析である。カプラン=マイヤ法は右側打ち切り（観察期間中に死亡が起ころなかったり、観察途中で研究参加をやめてしまったり、転居してしまったりして、エンドポイントが不明なもの）をうまく考慮することができる、ノンパラメトリックな生存曲線の推定法であり、半数死亡年齢（死亡年齢の中央値）を推定することができる\*10。具体的には、イベントが起こる可能性のある人口を分母、まだイベントが起こっていない人口を分子として、時間の長さを横軸にして、イベントが起こった場合は分子から1を引き、打ち切りが起こった場合は分母からも分子からも1を引くという形で、時間と人数の積和（Kaplan-Meierの積・極限推定量）を求め、まだイベントが起こっていない人口がちょうど全体の半分になるまでの時間をイベントが起こるまでの時間の中央値とする。

また、比例ハザード性を仮定したコックス回帰はきわめてロバストな（比例ハザード性だけを仮定し、死亡ハザードについて特定のパラメトリックな分布は仮定しないという意味で、いわばセミパラメトリックな）解析法である\*11。若齢での危険因子と高齢での危険因子は異なっているため、経時的に変化する説明変数を扱える必要があることと、死亡の生物学的メカニズムを考えるべきであることから、加速モデルなど、病因論的なハザードモデルも開発されてきている。これらを総称して、生存時間解析と呼ぶ。Rでは survival ライブラリを使うと、これら各種の生存時間解析を実行できる\*12。

生存時間解析は、通常、フォローアップ研究または歴史人口研究（数百年前でも宗門人別帳やキリスト教の教区簿冊 [parish record] には信者個人の名前と結婚、出生、死亡、疾病罹患、家族構成など、ほとんどすべての人口動態イベントが記録されているので、そのデータを再構成することによって、個人ごとの人口動態イベントが起こるまでの時間は把握できる）から得られるデータに対して用いられる。

\*10 Rでは survival パッケージの survfit() 関数で実行できる。

\*11 Rでは survival パッケージの coxph() 関数で実行できる。

\*12 <http://minato.sip21c.org/swtips/survival.html> の解説を参照されたい。

### 4.5.5 コホート分析

生存時間解析同様、フォローアップ研究または歴史人口研究のデータによく用いられる解析法として、コホート単位で変化を追うことで、世代間差を評価する「コホート分析」がある。例えば、コホート生命表（世代生命表）というものも作られているし、疾病罹患や死亡といった人口動態イベントの発生には、年齢の効果、時期の効果、コホートの効果が別々に（あるいは相互に関連しながら）作用すると考えられることから、これら3要因の効果を分解して考える、Age-Period-Cohort モデル（APC モデルと略記される）やレキスダイアグラムもしばしば用いられる。R では、Epi ライブラリに入っている `apc.fit()` や `apc.plot()` によって前者が、`plot.Lexis()` によって後者が実行できる（Carstensen, 2007）。

## 4.6 死亡の特性

死亡がさまざまな要因の集積による多面的現象であることは出生よりも広く認識されているが、同時に、年齢別死亡に何らかの規則性が存在することも古くから指摘されており、集団レベルでのそれを明らかにしようとする試みが多数なされてきている。なかでも生命表による死亡の分析は形式人口学のバックボーンとも言われる。人口転換にともなう死亡の集中という現象を個人差の減少とみなせば、個人レベルの現象と集団レベルの現象がはっきりとつながっていることがわかるが、発生までの期間が長いために、個人レベルの現象としての死亡は、なかなか捉えにくいという問題がある。

## 4.7 死因の分解

分解してみたところで、これらが独立でないのは自明である。例えば、乳児期の感染症を完全に防御できたら、感染症があつたら死んでいたはずの人は、年をとってから何か別の要因で死ぬのだ。通常は比例ハザードを仮定するが、そうでない考え方もある。

**偶発的な死亡** 相対的に若齢では事故死が、高齢では病死が多い社会がふつうであろう。

活動性が高い分、若齢の方が事故に遭う可能性が高いとも考えられるが、事故に遭った場合死亡に至る確率は、高齢者の方が高いだろう。従って、事故死に代表される偶発的な死亡は年齢とは無関係に起こるのだが、若齢成人ではそれ以外の死因による死亡がほとんどないことが、事故死割合の年齢差の原因と考えられる。

**初期故障** 初期故障は、乳幼児期に限定された死因である。それ以外の乳幼児期の死因としては、大部分が感染症である。乳児死亡に関しては、公衆衛生上の観点及び出生

への影響から、特別な意味があるので、さまざまな指標が提案されている。

感染症 感染症は、乳幼児期と高齢者において影響が大きい。

故障蓄積による死亡 がんを始めとして多くの成人病はこれである。

## 4.8 健康余命

公衆衛生学や医療行政の分野では、平均余命よりも、なるべく疾病負荷（ひいては医療費）の小さい生存期間としての健康余命の延伸が目標とされる。R では SURVEYLIFE パッケージ (<http://www.geog.leeds.ac.uk/projects/migrants/SummerSchoolFiles/QMSWednesday.html> から Word のドキュメントがリンクされている) や [http://www.eurohex.eu/pdf/Reports\\_2010/2010TR7.1\\_Decomposition%20tools.pdf](http://www.eurohex.eu/pdf/Reports_2010/2010TR7.1_Decomposition%20tools.pdf) がある。日本では 2012 年に厚生労働省が「都道府県別健康寿命」を発表してメディアの注目が集まったし、2015 年 12 月に 2 度目の発表があって各都道府県の保健当局は一喜一憂したように見受けられるが、実はいくつもの定義や計算方法があって、メディアなどでも混同されることがあって注意が必要である。複数の定義があるのは、「健康」の捉え方や、そのために使うデータの取り方が多様であることに起因している。

### 4.8.1 「健康」の扱いの違いによる違い

2015 年 12 月に厚生労働省から発表された都道府県別健康寿命は、健康日本 21（第二次）が、「健康格差の縮小」（具体的には「日常生活に制限のない期間の平均の都道府県格差の縮小」）を目標としているのを受けて計算されたもので、死亡についてはもちろん生命表だが、健康については国民生活基礎調査のデータを使っている。

2012 年「静岡 1 位」の根拠は、[http://www.mhlw.go.jp/bunya/kenkou/dl/chiiki-gyousei\\_03\\_02.pdf](http://www.mhlw.go.jp/bunya/kenkou/dl/chiiki-gyousei_03_02.pdf) に書かれているように、厚生労働科学研究費補助金「健康寿命における将来予測と生活習慣病対策の費用対効果に関する研究」であった。グラフを見ると、たしかに男性愛知県 1 位、女性静岡県 1 位となっている。この出所はかつて東大疫学の助教授だった橋下修二先生の研究班（以下橋本班と記す）で、研究班の web サイトの、健康寿命の算定結果（平成 22 年）（エクセルファイル、0.01 MB [http://toukei.umin.jp/kenkoujyummyou/syuyou/kenkoujyummyou\\_h22.xls](http://toukei.umin.jp/kenkoujyummyou/syuyou/kenkoujyummyou_h22.xls)）をみると、活動制限-都道府県というシートに載っているデータが、厚労省のグラフの元であることがわかる。ここでは、「日常生活に制限のない期間の平均」が健康寿命と考えていることに注目しなければならない。同じサイトでリンクされている算定方法を読むと、この計算方法が、「国民生活基礎調査と生命表を基礎情報とし」、「サリバン法（広く用いられている健康寿命の計算法）を用いて算定」という、かつてはほぼ標準的な算定方法

だったことがわかる。今回公開された資料でも、データが新しいものになっただけで計算方法は同様なので、愛知、静岡、山梨が上位に来ている傾向は変わっていない。

けれども、国民生活基礎調査の代わりに要介護認定の「要介護度2以上」が非自立であるという基準で「日常生活動作が自立している期間の平均」を出した表が、前述の研究班のサイトからリンクされている健康寿命の算定方法の指針（平成24年9月）[http://toukei.umin.jp/kenkoujyummyou/syuyou/kenkoujyummyou\\_shishin.pdf](http://toukei.umin.jp/kenkoujyummyou/syuyou/kenkoujyummyou_shishin.pdf)の付表3-2に載っている。これだと、男女とも長野県が日本一になっていて、長野県保健当局などは、こちらを参照して、長野県は平均寿命も健康寿命も日本一だと主張することがある。

どちらもそれなりに根拠はあるけれども、両者は「健康でない」とする定義が違うので、結果が異なるのも当然といえる。長野県は要介護度2以上の人が少ない一方で、国民生活基礎調査で「日常生活に支障がある」とか「自覚的には不健康」と答えた人は相対的に多かったと考えられる。一病息災という言葉もあるし、自分の健康状態を過信しない方が無理な生活をしないので自立期間は長くなるということかもしれない。

橋本班の最近発表された報告書 [http://toukei.umin.jp/kenkoujyummyou/houkoku/H26\\_toku.pdf](http://toukei.umin.jp/kenkoujyummyou/houkoku/H26_toku.pdf) を読むと、同研究班員で、公衆衛生学会編集委員会の副編集委員長でもある村上義孝さんが国内外の状況をレビューした結果（42ページ付近）には、最近の研究の動向として、2010年頃から要介護認定データを使ってサリバン法で算定したものに集約されてきたと書かれているので、国民生活基礎調査のデータを使う方法は、健康日本21（第二次）目標設定時には標準的だったけれども、若干古い方法と考えられ、意味から考えても、多少の順位の変動で騒ぐようなものではないだろう。

実は、国民生活基礎調査では、主観的健康度や通院についても尋ねていて、全国レベルではこれらを用いた健康余命の計算も行われている（別府，高橋，2016<http://www.paoj.org/taikai/taikai2016/abstract/1075.pdf>）ので、これを都道府県別比較に使ってみたら、また別の県が高い値を示すかもしれない。

#### 4.8.2 国際比較のための HALE

GBD 2013 DALYs and HALE Collaborators (2015) Global, regional, and national disability-adjusted life years (DALYs) for 306 diseases and injuries and healthy life expectancy (HALE) for 188 countries, 1990-2013: quantifying the epidemiological transition. Lancet, online publication on 26 Aug 2015 (doi:10.1016/S0140-6736(15)61340-X)<http://www.thelancet.com/journals/lancet/article/PIIS0140-6736%2815%2961340-X/abstract>

HALE（健康余命）の計算法はいろいろあるが、この研究は基本的に Sullivan 法を使っ

て、生命表の  $L_x$  から不健康な割合を取り除いて延べ平均健康生存年数を求めるという計算をしている。注意すべきことは、ここで「不健康な割合」の捉え方が日本の健康寿命とはまったく違うことである。世界共通の「不健康」の調査は不可能なので、日本の健康寿命の計算方法を世界ですることも当然不可能である。代わりに、Global Burden of Diseases にリストされている 300 余の病気への罹患を積み上げて「不健康な割合」を推定しているため、病気の罹患率や有病割合の推定精度に国や地域によって大きな差があることから、健康寿命の推定精度も国や地域によって異なることには注意が必要である。この論文では 1990 年から 2013 年の間に日本の健康余命は約 3 年伸びていて、相変わらず世界のトップであるが、伸び自体は先進国の平均的なレベルである。医療費削減の立場からすると HALE を伸ばすことよりも DALYs を短くすることが重要という意見もある（もちろん、中西準子さんが批判されていたように、この考え方にも問題はあるし、そもそも健康とは何なのかをもっと深く考えた方が良いと思うが）。

## 4.9 Age-Period-Cohort モデル

age 効果は普通の生命表をみればわかる。多くの死亡モデルはこれだけみている cohort 効果だけみるのが雪崩モデル period 効果は年次別変動。突発的なパンデミックの影響とか天変地異の影響とか戦争とか→それぞれ図示

Epi パッケージに apc という関数があり、これら 3 つの効果を同時に分析することができる。普通に重回帰するのと比べると、このモデルの意味がわかるだろう。もっと面白いのは残差か。

apc または apc+geography で 4 次元にしても面白い（g は国より、社人研の石井太さんがやっている JMD で都道府県データが入れば）。

レキシス図についても書く

【からだにいいもの R】 R と解析：人口統計学、レキシス図の描写「Lexis-PlotR」パッケージ……という tweet ([https://twitter.com/datasci\\_blogs/status/680118533213138944](https://twitter.com/datasci_blogs/status/680118533213138944)) から当該ページ (<http://www.karada-good.net/analyticsr/r-359/>) に飛んでみたら、確かに LexisPlotR というのは、grid グラフィクスでレキシス図を描くためのパッケージのようだった。

レキシス図はコホート研究のデータ提示に適していて、とくに age-period-cohort の 3 要因の効果を統合して眺めるのに良い作図法だが、base グラフィクスでは、既に Bendix Carstensen の Epi パッケージ (<http://bendixcarstensen.com/Epi/>) に Lexis.diagram() という関数があるし（使い方を知りたければ、Epi パッケージがインストール済みであれば、library(Epi); example(Lexis.diagram) の出力を見たらだいたいわかると思う）、3 要因を分解して捉えるための apc.fit() や apc.plot() という関数も用

意されているので、ggplot2の愛用者でなければ、敢えてLexisPlotRを導入するメリットは感じられない。

## 4.10 死亡の数学モデル

Graunt (1662) がロンドンの人口構造を推測するのに年齢依存の死亡スケジュールを仮定したのが、最初の数学的アプローチであった。次いで、DeMoivre (1725) が生存関数  $\ell(x)$  を年齢  $x$  の一次関数として、

$$\ell(x) = \ell(0) \times (1 - x/a)$$

として定式化した。 $\ell(x)$  は、その集団のうち正確な年齢  $x$  歳まで生き延びる割合を意味し、 $a$  は生存する可能性がある最高年齢を意味する。この定式化は明らかに誤りだが、昔は（例えば明治生まれの日本人男性のコホート生命表に対しては）そんなに悪い近似でもなかった。

年齢別死亡秩序については、Graunt 以来、さまざまなモデルが立てられてきた。代表的なものとして、老化に伴う死亡増加と年齢とは無関係な偶発的死亡を定式化した Gompertz-Makeham モデル（既に書いたが、生命表の高齢部分の補正にも用いられる）、年齢を3区分してそれぞれの死因の違いを考慮した Thiele モデル、それを改良した5パラメータの Siler モデルや8パラメータの Mode-Busby モデル、さらに乳児期の死亡の説明力を高めた Helligman and Pollard モデル、ワイブル分布によって小児期の死亡遞減を表した Mode-Jacobson モデル（Gage and Mode, 1993）、探索的に発見した3つのパラメータからなる簡単な関数でどんな死亡パターンも表現できることを示した Denny のモデル（Denny, 1997）<sup>\*13</sup> などがある。これらは、精緻化に伴って適合精度は向上したが、すべてアприオリに集団レベルでの年齢別死亡パターンを与えるものであった。もっとも、Siler モデル

$$h(t) = a_1 \exp(-b_1 t) + a_2 + a_3 \exp(b_3 t)$$

に関しては、Gage (1991) が解釈を与えている。Gage は、第1項は若年での感染症による死亡を意味し、第2項は事故死と妊娠出産に関連した死亡を意味し、第3項は中高年での癌や心血管疾患での死亡を意味するとしている。また、衛生状態の改善と  $a_1$  の低下が対応するなど、ある程度、パラメータが現実的意味をもつことも利点である。

<sup>\*13</sup> <http://minato.sip21c.org/demography/denny.html> を参照されたい。3つのパラメータを  $a, b, c$  とし、

$$\ell(x) = \frac{1}{(1 + a(\frac{x}{105-x}))^3} + b \sqrt{e^{\frac{x}{105-x}} - 1} + c(1 - e^{-2x})$$

例えば、fmsb に Jlife として入っている日本の完全生命表データを使って、1955 年男性と 2005 年男性の  $q_x$  に Siler モデルの当てはめをすることは、以下のコードで可能である。

```
library(fmsb)
Ages <- 0:111; FR <- Ages+1
qx1955 <- Jlife$qx1955M[FR]; qx2005 <- Jlife$qx2005M[FR]
rS1955 <- fitSiler(, qx1955); i <- 0
while (rS1955[7]>0 & i<10000) {
  rS1955 <- fitSiler(rS1955[1:5], qx1955)
  i <- i+1 }
rS2005 <- fitSiler(, qx2005); i <- 0
while (rS2005[7]>0 & i<10000) {
  rS2005 <- fitSiler(rS2005[1:5], qx2005)
  i <- i+1 }
LEGENDS <- c("qx1955(data)",
  sprintf("a1=%5.3f, b1=%5.3f, a2=%5.3f, a3=%5.3f, b3=%5.3f",
  rS1955[1], rS1955[2], rS1955[3], rS1955[4], rS1955[5]),
  "qx2005(data)",
  sprintf("a1=%5.3f, b1=%5.3f, a2=%5.3f, a3=%5.3f, b3=%5.3f",
  rS2005[1], rS2005[2], rS2005[3], rS2005[4], rS2005[5]))
LEGENDS[2] <- ifelse(rS1955[7]==0, LEGENDS[2], paste(LEGENDS[2], "[nofit]"))
LEGENDS[4] <- ifelse(rS2005[7]==0, LEGENDS[4], paste(LEGENDS[4], "[nofit]"))
par(cex=0.8, las=1)
plot(Ages, qx1955, xlim=c(0, 120), ylim=c(0, 1), pch=18,
  col="black", ylab="", axes=FALSE, frame.plot=FALSE,
  main="Fitting Siler's model for qx of Japanese males'
  complete life table in 1955 and 2005.")
axis(1,0:6*20,0:6*20)
axis(2,0:5/5,sapply(0:5/5,function(z) sprintf("%3.2f",z)))
lines(Ages, Siler(rS1955[1], rS1955[2], rS1955[3], rS1955[4],
  rS1955[5], Ages), lty=1, col="black")
points(Ages, qx2005, pch=1, col="red")
lines(Ages, Siler(rS2005[1], rS2005[2], rS2005[3], rS2005[4],
  rS2005[5], Ages), lty=2, col="red")
legend("top", pch=c(18, NA, 16, NA), lty=c(NA, 1, NA, 2),
  col=c("black", "black", "red", "red"), legend=LEGENDS)
```

これに代わるアプローチとしては、実際の生命表に基づいた Coale and Demeny の生命表<sup>\*14</sup>、基準となるモデル生命表から 2 つのパラメータでどんな生命表も表わせるとした

<sup>\*14</sup> 1966 年発表，1983 年改訂。4 つの異なるパタン [乳児，子供，成人，高齢者の死亡割合パタン] が異なる，東・西・南・北] と 25 の異なる死亡水準をもつモデル生命表を立てた。しかし，このアプローチの致命的

Brass (1968) のロジットモデル<sup>\*15</sup>などが知られている。これらの死亡モデルは、パラメータを増やすほど実際の年齢別死亡曲線へのあてはまりは良くなるので、老年期における死亡率のばらつきの補正や短期間の死亡データの補正には使えるが、個人の死亡のメカニズムを表現したものではないため、(1) 年齢別死亡率の偶然によるばらつきを評価することが難しい、(2) 社会変化にともなう疾病構造の変化など、死亡構造へのダイナミックに変動する影響を取り込むことが難しい、(3) 死にやすさの個人差を無視している、という欠点があった。

そこで、最近注目されているのが、雪崩モデルである。これは、個人の死亡過程を定式化した死亡モデルである (Gavrilov and Gavrilova, 1991)。中身を簡単に説明する。

- (1)  $s_0, s_1, s_2, \dots, s_n, \dots$  が  $0, 1, 2, \dots, n, \dots$  個の故障をもつ個人の人数とし、故障数と独立に速度  $\lambda_0$  で増加し、故障とは無関係に速度  $\mu_0$  で死亡が起こって総人数から除かれる、
- (2) 故障数に応じて故障の増加速度が  $\lambda$  増し、死亡率が  $\mu$  増す、と仮定する。これは、染色体損傷による死亡モデル (Le Bras, 1976) にバックグラウンド死力  $\mu_0$  を加味した形である。
- (3) この仮定の下で各状態の変化速度についての微分方程式を解いて合計すると生存関数が得られ、そこから死力を導出すると、

$$\mu(x) = \mu_0 + \mu\lambda_0(1 - \exp(-(\lambda + \mu)x))/(\mu + \lambda \exp(-(\lambda + \mu)x))$$

となり、 $\lambda \ll \mu$  ならば年齢  $x$  が小さいところでは年齢によらない死亡の成分と、加齢に伴う死亡増加成分の和という Gompertz-Makeham 型で近似できる。

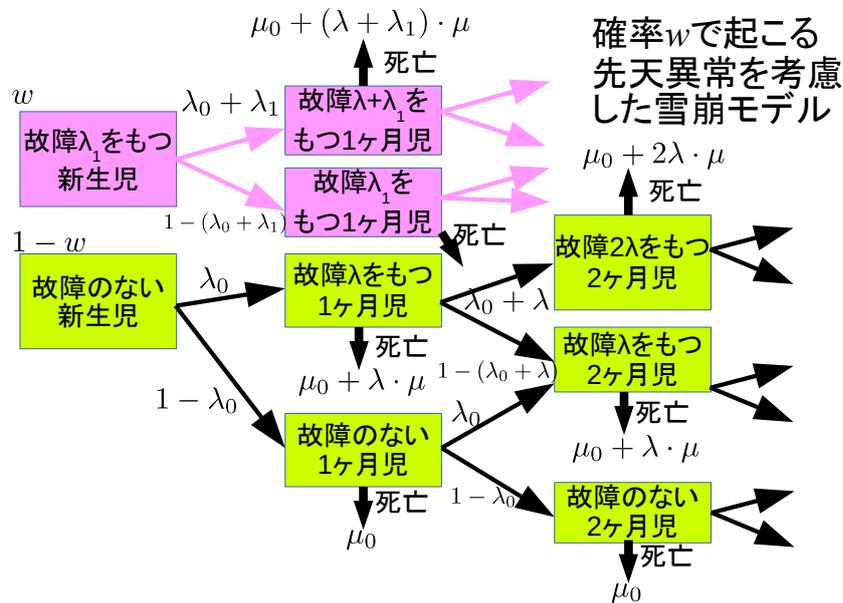
しかし、この形では乳幼児期の高死亡率とその遞減を表せないという大きな欠点があり、実際の死亡秩序への適合は高くなかった。雪崩モデルの改善には、出生時から故障を抱えている（つまり、先天異常をもつ）個人が集団内に存在するという frailty の不均質性を導入するとか（下図参照）、故障蓄積が死亡へと突き進む閾値を設定する、あるいは故障蓄積以外にも、事故や感染など突発的なリスクによって死亡してしまう可能性を考慮するといった方法がある (Mori and Nakazawa, 2003)。

な弱点は、柔軟性がないことであった。あくまで、現実のデータに「近い」モデルを選ぶしかなかった。

<sup>\*15</sup> ターゲットとなる集団の生命表の  $l_x$  を  $l_0(x)$ 、基準となる生命表の  $l_x$  を  $l_s(x)$  と書き、2つのパラメータを  $a, b$  とすれば、常に

$$\frac{1}{2} \ln\left(\frac{1 - l_0(x)}{l_0(x)}\right) = a + \frac{b}{2} \ln\left(\frac{1 - l_s(x)}{l_s(x)}\right)$$

と表わせる。



平成11年日本人男性の生存曲線については、オリジナルの雪崩モデルではどうしてもフィットしないのだが、先天異常割合を0.4%入れるだけで、劇的にフィットが改善する。しかし1891年生まれ日本人男性コホートのデータについては、先天異常を30%入れても初期の激しい生存確率低下にフィットしなかったため、感染症の影響を組み込むことによって、漸くフィットさせることができた。しかし、それでも、中高年での生存曲線が直線的に低下するところにはフィットしなかったため、故障蓄積に対して線形に死亡が上昇するのではなく、S字状の反応カーブを示すようにしたら、ほぼ完璧にモデルがデータにフィットした (Mori and Nakazawa, 2003)。この解析結果により、例えば1920年から1995年までの日本人男性の生存曲線にフィットさせたモデルの $\lambda_0$ の変化が、脳血管疾患による死亡が、戦後から高度経済成長期の塩分の高い食事、過重労働、過密な居住環境など（それらが $\lambda_0$ ）によって上昇し、その後、それらの改善によって低下したことと呼応していることが示された。

## 4.11 死亡データ分析の例

### 4.11.1 集計済の政府統計データをまとめて分析するには

日本のデータを使って実際に計算してみる。Excelでもいいのだが、せっかくなのでRを使ってみよう。Excelではセルに式をコピーすることによって計算を行うが、Rではベクトル演算ができるので、式の書き方がより簡潔明瞭になる。

なお、R で人口学的な分析を行うためのライブラリとしては `demogR`<sup>\*16</sup> や `Epi`<sup>\*17</sup> (の一部) や `demography`<sup>\*18</sup> があり、`demogR` では生命表分析をする関数や Coale and Demeny のモデル生命表、サンプルデータなどが提供され、`Epi` では Age-Period-Cohort 分析を行う関数やレキシスダイアグラムを描く関数などが提供されている。

---

\*16 <http://www.jstatsoft.org/v22/i10/paper> に詳細な説明がある。

\*17 <http://staff.pubhealth.ku.dk/~bxc/Epi/> から関連文書へのリンクを辿れる。

\*18 <https://github.com/robjhyndman/demography> で開発されている。主としてモデルによる予測のためのパッケージ。

```
mortality.R(1)

# mortality.R
# References:
# http://www.stat.go.jp/data/nenkan/02.htm
# 平成20年日本統計年鑑 第2章 人口・世帯
# Mortality data of Japanese in Japan.
# 2006年「国民衛生の動向」厚生統計協会（S60modelpopJ =昭和60年モデル人口）
#
S60modelpopJ <- c(8180,8338,8497,8655,8814,8972,9130,9289,9400,8651,7616,
  6581,5546,4511,3476,2441,1406,784)*1000
AC <- c(paste("0:16*5","0:16*5+4","[85-]"),sep=""),"[85-]")
# same as follows:
# AC <- c("[0-4]", "[5-9]", "[10-14]", "[15-19]", "[20-24]",
# "[25-29]", "[30-34]", "[35-39]", "[40-44]", "[45-49]",
# "[50-54]", "[55-59]", "[60-64]", "[65-69]", "[70-74]",
# "[75-79]", "[80-84]", "[85-]")
names(S60modelpopJ) <- AC

S60M <- c(6042, 1155, 1011, 3179, 3397, 3167, 4237, 7110, 10234, 15063,
  24347, 30747, 30884, 38240, 55100, 65593, 59125, 48786)
names(S60M) <- AC
H02M <- c(4532, 844, 760, 3204, 3466, 2916, 3264, 5449, 9769, 14218,
  20161, 32925, 42742, 42664, 51737, 69320, 67916, 67451)
names(H02M) <- AC
H07M <- c(3929, 752, 716, 2413, 3640, 3203, 3297, 4413, 8236, 15616,
  21905, 30491, 47188, 59828, 60927, 68504, 77924, 87750)
names(H07M) <- AC
H12M <- c(2933, 438, 493, 1721, 2875, 3271, 3749, 4621, 6840, 13141,
  24103, 31848, 42214, 60962, 76413, 73947, 73533, 102177)
names(H12M) <- AC
H17M <- c(2291, 409, 361, 1220, 2303, 2887, 3915, 4915, 6806, 10577,
  19546, 34233, 43403, 55261, 80198, 99338, 89502, 127261)
names(H17M) <- AC
```

```
mortality.R(2)

S60F <- c(4792, 636, 638, 1033, 1272, 1558, 2496, 4017, 5650, 7644,
  11504, 14828, 19961, 26490, 40891, 55657, 64448, 80930)
names(S60F) <- AC
H02F <- c(3451, 533, 482, 1149, 1329, 1361, 1774, 3102, 5542, 7510,
  10097, 14616, 19986, 27267, 38076, 58203, 71633, 110407)
names(H02F) <- AC
H07F <- c(3111, 483, 468, 949, 1447, 1393, 1832, 2426, 4578, 8520,
  11041, 14241, 21122, 29261, 41516, 56924, 79939, 141519)
names(H07F) <- AC
H12F <- c(2336, 300, 251, 676, 1160, 1546, 1847, 2425, 3639, 6595,
  11740, 14144, 18466, 28096, 40115, 57053, 73527, 171735)
names(H12F) <- AC
H17F <- c(1811, 246, 229, 582, 1067, 1283, 2037, 2554, 3432, 5177,
  9418, 15346, 18855, 25568, 40627, 60024, 84683, 225778)
names(H17F) <- AC
S60P <- c(7459, 8532, 10042, 8980, 8201, 7823, 9054, 10738, 9135, 8237,
  7933, 7000, 5406, 4193, 3563, 2493, 1433, 785)*1000
names(S60P) <- AC
H02P <- c(6493, 7467, 8527, 10007, 8800, 8071, 7788, 9004, 10658, 9018,
  8088, 7725, 6745, 5104, 3818, 3018, 1833, 1122)*1000
names(H02P) <- AC
H07P <- c(5995, 6541, 7478, 8558, 9895, 8788, 8126, 7822, 9006, 10618,
  8922, 7953, 7475, 6396, 4695, 3289, 2301, 1580)*1000
names(H07P) <- AC
H12P <- c(5904, 6022, 6547, 7488, 8421, 9790, 8777, 8115, 7800, 8916,
  10442, 8734, 7736, 7106, 5901, 4151, 2615, 2233)*1000
names(H12P) <- AC
H17P <- c(5578, 5928, 6015, 6568, 7351, 8280, 9755, 8736, 8081, 7726,
  8796, 10255, 8545, 7433, 6637, 5263, 3412, 2927)*1000
names(H17P) <- AC
```

ここで、男女を合計した死亡数を求めるには、ただベクトルの足し算をすればいい。年齢別死亡率 ASMR（上の説明では ADR としてある）を求めるには、年齢別の死亡数を年齢別の年央人口で割ればいい。粗死亡率は、死亡数合計を人口合計で割れば求められる。以上の R コードは次の枠内の通り。

```
mortality.R(3)
S60T <- S60M+S60F
H02T <- H02M+H02F
H07T <- H07M+H07F
H12T <- H12M+H12F
H17T <- H17M+H17F

S60ASMR <- S60T/S60P; S60CDR <- sum(S60T)/sum(S60P)
H02ASMR <- H02T/H02P; H02CDR <- sum(H02T)/sum(H02P)
H07ASMR <- H07T/H07P; H07CDR <- sum(H07T)/sum(H07P)
H12ASMR <- H12T/H12P; H12CDR <- sum(H12T)/sum(H12P)
H17ASMR <- H17T/H17P; H17CDR <- sum(H17T)/sum(H17P)

CDRs <- c(S60CDR, H02CDR, H07CDR, H12CDR, H17CDR)
```

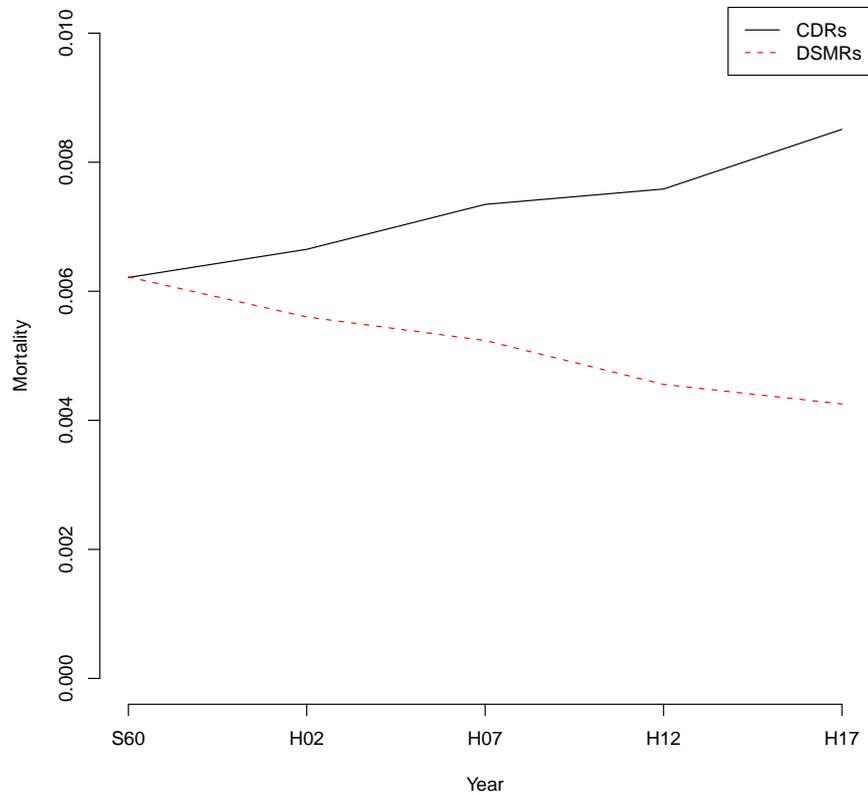
直接法年齢調整死亡率は、次の枠内のように関数定義でき、年齢別死亡率を与えると計算できる。

```
mortality.R(4)
DSMR <- function(ASMR) {
  if (length(ASMR)!=18) { print("age class is inadequate."); NA } else {
    sum(ASMR * S60modelpopJ)/sum(S60modelpopJ)
  }
}

DSMRs <- c(DSMR(S60ASMR), DSMR(H02ASMR), DSMR(H07ASMR),
  DSMR(H12ASMR), DSMR(H17ASMR))
```

以上求めた粗死亡率と直接法年齢調整死亡率の年次変化をグラフにして pdf ファイルに出力するには、次のコードでできる。

```
mortality.R(5)
pdf("compareDR.pdf", width=8, height=8)
plot(1:5, CDRs, type="l", col="black", xlab="Year", axes=FALSE,
  ylab="Mortality", ylim=c(0, 0.01))
axis(1, 1:5, c("S60", "H02", "H07", "H12", "H17"))
axis(2, seq(0, 0.01, by=0.002))
lines(DSMRs, col="red", lty=2)
legend("topright", col=c("black", "red"), lty=1:2, legend=c("CDRs", "DSMRs"))
dev.off()
```



次に生命表を作ってみる。例えば 2010 年の日本の男女別生命表の `qx` は、`fmsb` パッケージの `Jlife$qx2010M` と `Jlife$qx2010F` という変数になっているので、これらから生命表を計算するのは、下記のように容易である。

```
lifetable.R(1)
clifetable <- function(qx) {
  nc <- length(qx)
  lx <- numeric(nc)
  dx <- numeric(nc)
  Lx <- numeric(nc)
  lx[1] <- 1e+05
  for (i in 1:(nc - 1)) {
    dx[i] <- lx[i] * qx[i]
    lx[i + 1] <- lx[i] - dx[i]
    Lx[i] <- (lx[i] + lx[i + 1])/2
  }
  dx[nc] <- lx[nc]
  Lx[nc] <- lx[nc]/2
  Tx <- rev(cumsum(rev(Lx)))
  ex <- Tx/lx
  return(data.frame(qx, lx, dx, Lx, Tx, ex))
}

library(fmsb) # includes the definition of clifetable() above
# missing values have to be omitted before applying clifetable()
clifetable(Jlife$qx2010M[!is.na(Jlife$qx2010M)])
clifetable(Jlife$qx2010F[!is.na(Jlife$qx2010F)])
```

調査データとして得られるのは、通常、年齢別人口と年齢別死亡数なので、年齢別死亡数を年齢別人口で割った年齢別死亡率（ $ADR=ASMR$ ，生命表関数としては  $mx$ ）から生命表を求める関数定義が必要になる。ついでに年齢各歳でも5歳階級でも計算できるように（乳児死亡と1-4歳死亡が分かれている場合にも対応できるように）年齢階級幅を可変にし、 $mx$  から  $qx$  を求める際の補正方法も複数パターンをサポート（`mode` というオプションで1のとき単純な線形補間，2のとき Greville の補正）したコードを `fmsb` パッケージに `lifetable()` 関数として実装した。

## lifetable.R(2)

```

lifetable <- function (mx, ns = NULL, class = 5, mode = 1) {
  nc <- length(mx)
  qx <- numeric(nc)
  if (mode > 10) {
    mode <- mode%10
    grev <- TRUE
  } else {
    grev <- FALSE }
  if (is.null(ns)) {
    n <- rep(class, nc)
    ages <- c(0, cumsum(n)[1:(nc - 1)])
    if (mode %in% 4:5) {
      mode <- mode - 2 }
  } else {
    n <- ns
    ages <- c(0, cumsum(n)[1:(nc - 1)])
    if (mode %in% 2:3) {
      mode <- mode + 2 }
  }
  if (mode == 1) {
    ax <- c(rep(0.5, nc - 1), 1/mx[nc])
  } else if (mode == 2) {
    ax <- c(0.1, rep(0.4, 4), rep(0.5, nc - 6), 1/mx[nc])
  } else if (mode == 3) {
    ax <- c(0.3, rep(0.4, 4), rep(0.5, nc - 6), 1/mx[nc])
  } else if (mode == 4) {
    ax <- c(0.1, 0.4, rep(0.5, nc - 3), 1/mx[nc])
  } else if (mode == 5) {
    ax <- c(0.3, 0.4, rep(0.5, nc - 3), 1/mx[nc])
  } else if (mode == 6) {
    if (mx[1] < 0.107) {
      ax <- c(0.045 + 2.684 * mx[1], (1.651 - 2.816 * mx[1])/4,
        rep(0.5, nc - 3), 1/mx[nc])
    } else {
      ax <- c(0.33, 1.352/4, rep(0.5, nc - 3), 1/mx[nc]) }
  } else if (mode == 7) {
    if (mx[1] < 0.107) {
      ax <- c(0.053 + 2.8 * mx[1], (1.522 - 1.518 * mx[1])/4,
        rep(0.5, nc - 3), 1/mx[nc])
    } else {
      ax <- c(0.35, 1.361/4, rep(0.5, nc - 3), 1/mx[nc]) }
  } else {
    ax <- c(rep(0.5, nc - 1), 1/mx[nc]) }
  if (!grev) {
    qx <- n * mx/(1 + n * (1 - ax) * mx)
    qx[nc] <- 1
  } else {
    for (i in 1:(nc - 1)) {
      qx[i] <- mx[i]/(1/n[1] + mx[i] * (1/2 + n[i]/12 *
        (mx[i] - (log(mx[i] + 1) - log(mx[i]))/n[i]))) / n[i] }
    qx[nc] <- 1 }
  px <- dx <- lx <- Lx <- numeric(nc)
  lx[1] <- 1e+05
  px <- 1 - qx
  for (i in 1:(nc - 1)) {
    dx[i] <- lx[i] * qx[i]
    lx[i + 1] <- lx[i] - dx[i]
    Lx[i] <- n[i] * (lx[i + 1] + ax[i] * dx[i]) }
  dx[nc] <- lx[nc]
  Lx[nc] <- lx[nc]/mx[nc]
  Tx <- rev(cumsum(rev(Lx)))
  ex <- Tx/lx
  return(data.frame(ages, n, ax, mx, qx, px, lx, dx, Lx, Tx, ex)) }

```

このように定義しておけば (fmsb パッケージで定義済みなので、実際はパッケージを

ロードするだけで良い), 例えば, 上記 `mortality.R` に与えられている昭和 60 年の年齢別死亡率から, Greville の方法で補正した生命表を求める計算は 1 行で済み, 0 歳平均余命が 78.15 年であることがわかる。

```
lifetable.R(3)

library(fmsb)
source("http://minato.sip21c.org/ldaR/mortality.R", encoding="CP932")
options(digits=3) # avoiding too detailed digits
lifetable(S60ASMR, class=5, mode=2)
```

	ages	n	ax	mx	qx	px	lx	dx	Lx	Tx	ex
[0-4]	0	5	0.10	0.001452	0.007215	0.993	100000	721.5	496753	7815323	78.15
[5-9]	5	5	0.40	0.000210	0.001049	0.999	99278	104.1	496080	7318570	73.72
[10-14]	10	5	0.40	0.000164	0.000821	0.999	99174	81.4	495628	6822490	68.79
[15-19]	15	5	0.40	0.000469	0.002342	0.998	99093	232.1	494769	6326862	63.85
[20-24]	20	5	0.40	0.000569	0.002842	0.997	98861	280.9	493462	5832094	58.99
[25-29]	25	5	0.50	0.000604	0.003015	0.997	98580	297.3	492157	5338632	54.16
[30-34]	30	5	0.50	0.000744	0.003711	0.996	98283	364.8	490502	4846476	49.31
[35-39]	35	5	0.50	0.001036	0.005168	0.995	97918	506.0	488325	4355974	44.49
[40-44]	40	5	0.50	0.001739	0.008656	0.991	97412	843.2	484952	3867649	39.70
[45-49]	45	5	0.50	0.002757	0.013689	0.986	96569	1321.9	479539	3382698	35.03
[50-54]	50	5	0.50	0.004519	0.022344	0.978	95247	2128.2	470913	2903159	30.48
[55-59]	55	5	0.50	0.006511	0.032032	0.968	93119	2982.8	458136	2432246	26.12
[60-64]	60	5	0.50	0.009405	0.045946	0.954	90136	4141.4	440325	1974110	21.90
[65-69]	65	5	0.50	0.015438	0.074320	0.926	85994	6391.1	413994	1533785	17.84
[70-74]	70	5	0.50	0.026941	0.126205	0.874	79603	10046.3	372901	1119791	14.07
[75-79]	75	5	0.50	0.048636	0.216818	0.783	69557	15081.2	310082	746890	10.74
[80-84]	80	5	0.50	0.086234	0.354701	0.645	54476	19322.6	224072	436808	8.02
[85-]	85	5	6.05	0.165243	1.000000	0.000	35153	35153.2	212736	212736	6.05

## 4.12 文献

- Alvarez HP (2000) Grandmother hypothesis and primate life histories. *American Journal of Physical Anthropology*, 113: 435-450.
- Carstensen B (2007) Age-period-cohort models for the Lexis diagram. *Statistics in Medicine*, 26(15): 3018-3045.
- CDC: Deaths and Mortality. [<http://www.cdc.gov/nchs/fastats/deaths.htm>]
- Chamberlain AT (2006) *Demography in Archaeology*. Cambridge University Press.
- Charnov EL (1992) *Life History Invariants*. Oxford University Press.
- Denny C (1997) A model of the probability of survival from birth. *Mathematical and Computer Modelling*, 26: 69-78.
- Gage TB (1991) Causes of death and the components of mortality: Testing the biological interpretations of a competing hazards model. *American Journal of Human Biology*, 3(3): 289-300.
- Gage TB, Mode CJ (1993) Some laws of mortality: How well do they fit? *Human*

- Biology*, 65: 445-461.
- Gavrilov LA, Gavrilova NS (1991) *The Biology of Life Span: A Quantitative Approach*. Harwood Academic Publisher, New York.
  - Hawkes KJF, O'Connell NG, Blurton-Jones HA, Charnov EL (1998) Grandmothering, menopause, and the evolution of human life histories. *Proceedings of National Academy of Sciences, USA*, 95: 1336-1339.
  - Kannisto V (2000) Measuring the Compression of Mortality. *Demographic Research*, 3(6). (<http://www.demographic-research.org/volumes/vol13/6/3-6.pdf>).
  - 川上憲人, 小林廉毅, 橋本英樹 [編] (2006) 『社会格差と健康：社会疫学からのアプローチ』, 東京大学出版会.
  - Keyfitz N, Caswell H (2005) *Applied Mathematical Demography. 3rd Ed.*, Springer Science+Business Media, Inc., New York.
  - 国際人口学会編, 日本人口学会訳 (1994) 『人口学用語辞典』, 厚生統計協会.
  - 近藤良 (1980) 1966年(ひのえうま)における先天異常乳児死亡率の見かけ上の増加について. *先天異常*, 20: 7-16.
  - Mori Y, Nakazawa M (2003) A new simple etiological model of human death. *The Journal of Population Studies (Jinko-Gaku-Kenkyu)*, 33: 27-39.
  - 本川達雄 (1992) ゾウの時間・ネズミの時間. 中公新書
  - 中澤 港 (2007) 「第9章 小集団人口学」 In: 稲葉 寿編『現代人口学の射程』, ミネルヴァ書房, 京都, pp.172-195.
  - Ng E, Gentleman JF (1995) The impact of estimation method and population adjustment on Canadian life table estimates. *Health Reports*, 7(3): 15-22.
  - Smith DP (1992) *Formal Demography*. Plenum Press.
  - Staetsky L (2009) Diverging trends in female old-age mortality: A reappraisal. *Demographic Research*, 21: 30.  
[<http://www.demographic-research.org/volumes/vol21/30/21-30.pdf>]
  - 杉本正信, 古市泰宏 (1998) 老化と遺伝子. 東京化学同人
  - 館稔 (1968) 『人口分析の方法』, 古今書院.
  - 和田光平 (2006) 『Excelで学ぶ人口統計学』, オーム社.
  - Westendorp RGJ, Kirkwood TBL (1998) Human longevity at the cost of reproductive success. *Nature*, 396: 743-746.



## 第5章

# 出生の分析

出生は、移動がなければ、人口変化に正の寄与をする唯一の人口動態であり、「人口変化」と同じ次元の現象である。つまりは、人間の集団の複雑な営為が引き起こした結果として見える現象を対象としているわけで、単一の側面だけでは説明しきれないのが当然であり、その理解には多面的及び学際的なアプローチが必要となる。

しかし一般に、例えば、「日本の少子化はなぜ起こったか?」「少子化を改善するにはどうしたらよいか?」などと政策的問題設定が行われる際に、この多面的アプローチの必要性が意識されることは稀である。多くは社会経済的側面からだけの対策が立てられているようである。仕方ない面もあるのだが、少子化対策などがうまく行かない原因の一端は、そこにあると思う。ここで「仕方ない」という理由は、多面的アプローチをする専門家が不足していることだ。

日本に限らず、世界中を見渡しても、人口学者の多くは社会経済的側面からのアプローチを主としており、生物人口学の専門家は数えるほどしかいない。では、生物学者の方はどうかと見れば、分子生物学が花盛りであり、個体以上のレベルを対象とする研究は、相対的に手薄である。生態学や人類学の中では、生物人口学的な研究をする人もいるわけだが、人類学会の発表を見ても大半が遺伝か形質か考古で、生態とか人口といった研究はごく稀である。それゆえ、生物人口学的な知見の集積が足りないということはいえると思う。逆に、産婦人科学的な、不妊治療、生殖医療といった観点からアプローチされる医師の方々には、社会経済的側面や文化的側面が相対的に不足していたり誤解があったりするものがまま見られる。

本稿の目的は、出生の人口学的分析方法を紹介することであるが、分析方法に入る前に、まず出生という人口学的プロセスを構成している要因を包括的に理解しておく必要がある。ついで、人口学分野で、出生を捉えるために、どういうアプローチが取られてきたかを紹介し、最後に出生の指標（分析方法）を調査方法別に紹介する。

## 5.1 出生というライフイベントの多面性

### 5.1.1 ヒトの生殖戦略～マクロな、あるいは生物学的視点

ヒトの女性が一生の間に産む子どもの数（完結出生力）は、記録によれば0人から69人である（長谷川，1993）。69人というのは双生や品生を多く含む例外的なデータであるが、集団の平均でみても1～11程度ときわめて大きなばらつきをもっており、しかもチンパンジーの平均完結出生児数が2～4であるのに比べると、かなり高い方にシフトしている。自然状態で生活している動物の出生力を調べるのはきわめて難しく、データが乏しいのでバイアスがかかっている可能性があるにしても類人猿の中では高い方である。

しかし、哺乳類全体にまで比較の範囲を広げると（完結出生力での比較はデータが乏しく難しいので一腹仔数 [= 1度の出産で生まれる子どもの数] での比較になるが）、ヒトを含めて霊長目は概ね1から2であり、平均4のネコや平均8のマウスなどに比べてずっと少ない。性的に成熟してから老化によって生殖能力を失うまでの期間、つまり再生産期間は長い、その一方で月経周期が長いために受胎のチャンスが少なく、在胎期間と産後授乳中の不妊期間からなる「1回の出産にかかる期間」が長いために出産間隔も長く、性成熟に要する期間も長いので、出産可能な回数は少ないのである。少産ということ自体に進化的な直接の利点はないにもかかわらず、霊長目が少産となったのは、一種の副作用である。つまり、大きい子を産むことや、産んだ子の世話（授乳したり離乳食を与えたりすることを含む）をすることに大きな進化上の利点があったから、代償として1個体の子を産むことに大きなコストがかかるようになり、多産が不可能になったのである。

実際には、意図的な出産制限が無い場合に健康なヒトが一生の間に産む子どもの数は、後で述べるように栄養状態などによって異なるが、4人から10人程度である。生態学者のマッカーサー (R.H. MacArthur) とウィルソン (E.O. Wilson) が考え出した  $r-K$  淘汰という概念によれば、このレベルの出生力は、ヒトが  $K$  戦略者、すなわち少数の子どもを大事に育てる戦略をとる動物であることを示すものである。 $K$  戦略は比較的安定した自然環境のもとで外敵の少ない種が発達させる戦略であり、霊長目が外敵の少ない森林で進化したことを考えれば、もっともらしい説明である。しかし、これだけでは、なぜ他の類人猿が霊長目の中でも少ない完結出生力を発達させているのにヒトの完結出生力が大きくなっているのかということや（意図的な出産制限が無い場合に健康な女性が一生の間に産む子どもの数は平均して3から10程度であり、2から4と推定されるチンパンジーよりは明らかに多く、かつ分散が大きい）、ヒトだけが閉経後にも長い期間生存することを説明できない。この点については、チャーノフ (E. Charnov) の生活史戦略の観点から、寿命が長くなるような戦略という指摘がなされている。そもそも寿命とは老化によって個体の生命

が尽きるまでの時間である。老化の意味についてはいくつかの説があって論争中であるが、有力な仮説の一つに体細胞廃棄説がある。何十億もの体細胞すべてのエラー修復を長い間続けることは非効率なため、生殖細胞だけのエラー修復をして、そこから再び個体発生をして次世代を構築することで全体としてのエラー蓄積を避けることができるわけであるが、その代償として廃棄すべき体細胞にはエラーがたまっていくのは避けがたい。次世代の生産を終えた閉経後の生存が次世代の生存に寄与しないなら、閉経後の生存は淘汰を受けないことになり、遺伝的な影響はないことになる。したがって、同じように 50 歳で閉経するならば、寿命が 60 歳でも 100 歳でも子ども数とは関係しないわけであり、それならば寿命が 60 歳で尽きたとしても多産にした方が適応的であると考えられる。実際、イギリス貴族の家系図を使った分析では、60 歳で死んだ人の子ども数の方が 70 歳で死んだ人の子ども数より多く、80 歳、90 歳と死亡年齢が高くなるに連れて子ども数が少なくなる傾向があったことが指摘されている。しかし、ヒトの場合は子どもが再生産年齢に達するまでの養育コストが大きいので、孫や曾孫を育てるという形による閉経後の家系の生存への寄与は無視できない。ユタ大学のホークス (K. Hawkes) らは、チンパンジーと再生産期間がそれほど変わらないのにヒトの再生産完了後の生存年数が長いのはおばあちゃんが忙しい娘の子育てを手伝うためである、という説を唱えている (Hawkes et al., 1998)。以上をまとめると、マクロな視点からみたヒトの生殖戦略は、他の多くの霊長目と同じく 1 腹 1 子を数年の間隔をあけて生み、その子どもたちを祖父母世代も一緒になってコストをかけて大事に養育し、なるべく子どもの死亡率が低くなるような方向に進化してきた一方で、そのことが閉経後の寿命を延ばしてきたということができる。

このシステムの数学的極限を考えると、長生き世代と多産世代が交互に繰り返すことができれば合理的であるが、ヒトの遺伝子はそれを実現するよううまいメカニズムをいまだ発明していないようである。しかし、実孫でなくとも養育するような文化的システムがあれば、世代交代の代わりに、多産で短命な系統と少産で長寿な系統が二極分化して並立することもまた合理的である。ヒトが生涯に産む子ども数の分散が大きい原因の一端は、このあたりにあるのかもしれない。

### 5.1.2 もう一つのマクロな視点～有性生殖のコストと利点・配偶者選択

ヒトの生殖戦略についてもう一つ考えねばならないマクロな視点は、それが有性生殖であるということである。有性生殖には進化的にみていくつかの利点がある。木村資生とともに集団遺伝学の基礎を築いたクロー (J.F. Crow) によれば、次の 3 つが最も重要だとされている。第 1 に、まれで有益な突然変異が結集することである。無性生殖では別の系統に生じた変異は決して同一個体にとりこまれることがないが、有性生殖で組み換えが起これば二つの有利な突然変異が同一個体にもたらされることが起こりうる。このことはマ

ラー (H.J. Muller) が指摘し、進化の初期においてとくに重要な意味をもっていたと考えられている。第2に、変化する環境への対処をする上で有利だという点があげられる。無性生殖で大きな環境変化に対処するほど変異を起こしやすくすると、同時に有害な型の負荷も大きくなるのに対して、有性生殖で組み換えが起こることによって現在の集団構成員のどれよりも適応度の高い型を生ずることができるので全体としての突然変異発生率を低く抑えることができる。第3に反復突然変異に対処する上で組み換えが有利だという点があげられる。無性生殖集団では有害な突然変異が排除される機会は復帰突然変異が起こらない限りないわけだが（「マラーのラチェット」と呼ばれる）、有性生殖では組み換えによって突然変異のない型が容易に再構成されるのである。これら3つに代表される利点によって、組み換えがある有性生殖集団が無性生殖集団に打ち勝ったという群淘汰が現在の有性生殖集団を構成してきたのだろうとフィッシャー (R.A. Fisher) は示唆している。

それでもなお、有性生殖には大きなコストがかかる。無性生殖に比べると繁殖効率が悪いばかりでなく、無性淘汰が全遺伝分散に作用するために最適な遺伝子型が一度現れたら存続するのに対して、有性生殖ではライト (S. Wright) のいう多重峰問題が存在し、最適な遺伝子型が存続するとは限らない。これは上に挙げた利点と表裏一体である。繁殖効率が悪いというのを言い換えれば、配偶者選択が必要になるということである。ヒトに限らず哺乳動物はすべて有性生殖であるが、ヒトの場合は脳が大きくなったことや個体数が多くなったことなどが原因となって社会が複雑化しているので、配偶者選択にきわめて大きなコストがかかることが特徴的である。社会の複雑化にともなって成立したと考えられる複雑な婚姻規制により、通婚圏がある程度大きくなると適切な配偶者が見つからないことがきわめて容易におこりうる。狩猟採集を主な生業にしていたバンド社会のヒト集団では、100人前後のバンド社会の中での婚姻が大部分だったと想定されているので、人口規模100人でシミュレーションしてみると、結婚適齢に達してもちょうどいい年齢の相手が見つからないために結婚できないというケースが10%近く起こるのである (Nakazawa and Ohtsuka, 1997)。

それを乗り越えたとして、さてどういう配偶者を選ぶのか、という問題が残る。これに関しては、同類婚と雑種強勢の綱引きではないかと考えられる。同類婚は希な遺伝子の保持に寄与するが、変動する環境下では多様な遺伝子をもつ方が個体の生存に有利であるため、集団遺伝学的にはどちらもそれなりに理にかなっている。雑種強勢は、ある意味では自分に似ていない相手を選ぶことであるが、これについてはフェロモン支配という説が有力である。動物実験では1980年代から確かめられていたが、1997年になって、ヒトでも、ヴェデキンド (C. Wedekind) らの実験により、HLAの型が自分と違う異性の汗の臭いに惹かれるという結果が得られたのである (栗原, 1998; 山元, 2004)。次節でも触れるが、生殖において嗅覚が果たす役割は、配偶者選択ばかりでなく、きわめて大きいことがわかってきている。

### 5.1.3 ヒトの生殖の生物学的機構～ミクロな、あるいは医科学的視点

ヒトの生殖のミクロなメカニズムを考えると、女性に卵子を供給する能力、男性に精子を供給する能力があって、適切なタイミングで性交が行われ、受精し、受精卵が子宮に着床し、在胎期間を生き延びて出産に至るといった一連の過程が必要である。この過程は、すべてにわたって神経とホルモンの作用によって精密にコントロールされている。さらにいえば、胎児期の女性あるいは男性への性分化や思春期に生殖機能が成熟する過程もホルモンによる制御を受けている。これらの制御メカニズムの研究は生殖内分泌学 (reproductive endocrinology) と呼ばれ、微量なホルモンの測定方法の発達や、in situ hybridization のようなレセプターの局在を同定する方法の開発により、新たな知見が続々と得られてきている分野である。本節の残りの部分では、生殖内分泌学の知見を踏まえ、一連の生殖過程を順に説明する。

**女性が卵子を供給する能力** 生殖が成立するためには、まず女性において有排卵月経が定期的になければならない。ヒトの女性は出生時に卵原細胞を卵巣内に約 500 万個もっており、それらは 1 個 1 個が裸の状態ではなくて別の細胞に取り囲まれている (原始卵胞と呼ばれる)。原始卵胞内の卵原細胞は分裂休止期にある。思春期になると月経が始まり、原始卵胞から 1 つずつ成熟卵胞が形成されてゆく。月経のメカニズムは以下のように理解されている。

まず視床下部から GnRH (ゴナドトロピン放出ホルモン) の分泌が起こり、刺激を受けた下垂体から FSH (卵胞刺激ホルモン) が分泌され、血流を通して卵巣を刺激し、数十個の原始卵胞が発育を開始する。これらの原始卵胞からはステロイドホルモンの一種であるエストロゲンが分泌され (この場合、正確にはエストラジオールであるが、本稿では総称のエストロゲンを用いる)、それと同時に子宮では子宮内膜が増殖を開始する。やがて発育中の原始卵胞から排卵まで進む主席卵胞が 1 つ選択され、他の原始卵胞の発育は止まってしまう (なお、自然状態では 2 つ以上が発育を続ける確率は約 1 % である)。主席卵胞は大きく成熟し、卵原細胞も第 1 次卵母細胞となり、エストロゲンの分泌も高まってゆく。この期間は卵胞期 (または濾胞期) と呼ばれ、通常 14 日間前後である。やがて卵胞からのエストロゲン分泌がピークに達すると、血中濃度が 250-450 pg/mL を超え、それがほぼ 2 日間維持されていることを視床下部と下垂体が感知すると、視床下部からの GnRH 分泌が上昇し、下垂体からの FSH 分泌が低下し、代わって LH (黄体化ホルモン) 分泌が上昇する。この作用はエストロゲンを介した LH 分泌の正のフィードバック制御であるが、排卵前の低レベルのプロゲステロンはこのフィードバック制御を増強する

ことが知られている。なお、この時の LH は一気に大量分泌されるので LH サージと呼ばれる。成熟卵胞が LH サージを感知すると、中から卵子が放出される。これが排卵である。なお、排卵の頃に第1次卵母細胞は第1回成熟分裂を行い、第2次卵母細胞と第1次極体になっているので、実際に放出される卵子は第2次卵母細胞と第1次極体がくっついたものの周囲を透明帯と何層かの卵丘細胞が取り巻いている状態である。排卵とともに、卵子を放出した後の卵胞は LH の作用によって黄体へと変化する。黄体期の開始である。黄体から分泌される高レベルのプロゲステロンに刺激された子宮内膜は、受精卵の着床に備えて厚さ 10 mm 程度にまで肥厚し分泌期に入る。同時にこの高レベルのプロゲステロンは脳の体温調節中枢にも作用し、体温を上昇させる。さらに下垂体に対して抑制的に作用し、後述する LH のパルス頻度を低下させ（エストロゲン存在下で LH パルスを高頻度低振幅から低頻度高振幅なパターンに変化させる）、それまで LH サージを起こしていたエストロゲンによる正のフィードバック制御回路そのものを抑制することで LH レベルを低下させる。受精卵が子宮内膜に着床しない限り約 14 日間で黄体の寿命は尽きてプロゲステロンの分泌は低下し、子宮内膜への血液の供給が低下して子宮内膜は基底部分を残して壊死し、はがれ落ちて体外へ放出される。これが月経である。寿命が尽きた黄体は白体へと変化する。ヒトでは1月経周期（月経が終わった日から次の月経が終わる日まで）が約 28 日なので、12 歳から 50 歳まで妊娠せずに月経が続いたとしても排卵に至る卵胞の総数は 500 個足らずであり、出生時の原始卵胞のうち生殖に関与する可能性があるのは 0.01% に満たないのである。

月経周期をつかさどる上記一連の制御機構は視床下部－下垂体－卵巢系 (hypothalamus-pituitary-ovarian axis) と呼ばれる。この系において鍵となるプロセスは、視床下部からの GnRH の分泌である。GnRH の分泌は、通常はパルス状で起こるが、LH サージの直前には一度に大量分泌される（サージ）。これらは別々の制御系によっており、それぞれ GnRH パルスジェネレータ、GnRH サージジェネレータと呼ばれる。パルスジェネレータは男性にもあるが、サージジェネレータは女性にしかなく、月経周期を形成する上で重要な役割を担っている（田中，1998）。神経内分泌系のメカニズムを調べることは技術的に難しいために、これらのジェネレータのメカニズムもまだ解明されたとは言い難いが、最近になっていくつかの知見が得られてきている。

GnRH パルスジェネレータは、卵巢を除去したラットの視床下部正中隆起部に多ニューロン発射活動記録用電極を恒常的に植え込んで電気活動の状態を記録し、同時に心房内に留置したカニューラから 6 分おきに採血して血中 LH 濃度を測定するという川上正澄の実験によって、正中隆起部に約 15～20 分おきに同期して活動電位を発生する（発火する）ニューロン群として確認された。この発火は血中 LH

濃度のパルス状の高まりに先立って起こっていることも確認された。ノビル (E. Knobil) らのアカゲザルを使った実験によって、霊長目ではこの発火が約 1 時間周期であることが推定されている。その後、自然状態ではエストロゲンはパルスジェネレータに影響を与えないにもかかわらず、これらの卵巣を摘出した動物にエストロゲンを与えると数時間のうちに GnRH パルスジェネレータの発火間隔が延び LH パルスも間隔が延びることがわかった。そこで視床下部のいろいろな場所や卵巣摘出したアカゲザルの第 3 脳室に直接エストロゲンを注入したところ、末梢血中の LH 濃度が低下することが確認され、エストロゲンは負のフィードバック制御を行っていることがわかった。つまり、エストロゲンは卵巣から分泌されるので、エストロゲン濃度が高くなるとそれが視床下部や下垂体のエストロゲン感受性ニューロンに感知されて LH 分泌が低下して卵巣への刺激が止まり、エストロゲン濃度は一定範囲に保たれるというわけである。ただし、エストロゲン感受性ニューロンの情報は直接 GnRH パルスジェネレータに届くのではなく、オピオイドニューロンが介在するらしい。オピオイドとは生体内で合成、分泌されるモルヒネ様ペプチドのことで、オピオイドレセプターに結合することによって疼痛抑制、呼吸抑制、催眠などの作用をもつ物質の総称であり、一般に生体内快樂物質として知られている。物質としては、エンドルフィン、エンケファリン、ダイノルフィンの 3 群に分類され、それぞれ  $\mu$ ,  $\kappa$ ,  $\delta$  というレセプターに特異的に結合する。 $\mu$  レセプターの働きを抑制する物質であるナロキソンを投与すると GnRH と LH のパルスが頻回になるという実験結果から、エストロゲン濃度が高くなるとエストロゲン感受性ニューロンから何らかの神経伝達物質が分泌されてオピオイドニューロンを刺激し、分泌された  $\beta$  エンドルフィンが GnRH ニューロンの働きを抑制するというメカニズムが推測されている。なお、 $\beta$  エンドルフィンが GnRH ニューロンの働きを抑制するメカニズムとしては、GnRH の分泌を制御している一酸化窒素を伝達物質とする神経系を止めてしまうことが最近提唱された (Faletti et al., 1999)。

一方、GnRH サージジェネレータについては、ヒトやサルでの知見はまだほとんどないが、ラットでは GABA (gamma aminobutyric acid の略で、日本語ではガンマアミノ酪酸) ニューロンが介在していることがわかっている。GABA は哺乳類の中枢神経系に高濃度に存在する抑制性の神経伝達物質の 1 つであり、生体内には GABAA と GABAB の 2 種類のレセプターがある。ラットの性周期は 4 日間であるが、そのうち発情前期の日の LH サージが始まる前に GABAA の働きを阻止するピククリンという薬物を投与すると LH サージが早まるが、それ以外の時点でのピククリン投与は何の作用も起こさないのである。このことから、次のような推測がなされる。すなわち、通常の状態では GnRH サージジェネレータは GABA ニューロンによって抑制されていて、卵胞が成熟してエストロゲンへの曝露が一定の値と

時間を超えると、それが排卵準備完了のサインとなって GABA ニューロンの作用を止め、GnRH サージが起こるというメカニズムである。これはエストロゲンの正のフィードバック作用と呼ばれる（田中，1998）。

**男性が精子を供給する能力** 生殖のためには、男性において運動性のある精子が十分な濃度で生産され、性交時に射精されることが必要である。ヒトの男性では、Y 染色体上に存在する SRY という遺伝子が胎生期に発現し、未分化の性腺を精巣に分化させる。胎生 7 週くらいに生殖腺隆起の髄質の発達が著しくなり、逆に周辺部の皮質が退化して精巣になってくるのである。精巣には将来精子形成を行う精細管と間質細胞（ライディッヒ細胞）が出現してくる。間質細胞からは胎生期のうちにテストステロンの分泌が始まり、その作用でウォルフ管が精巣上体、輸精管、精嚢、射精管などへと分化し、ミュラー管が退化する。ミュラー管の退化は精細管の中にある支持細胞（セルトリ細胞）から分泌される抗ミュラー管ホルモンの作用による。外性器の性分化も胎生 8 週ころからテストステロンの作用で起こる。視床下部の GnRH パルスジェネレータによる GnRH 放出によって刺激を受けた下垂体から LH と FSH がパルス状に分泌され、これらが精巣の間質細胞を刺激してテストステロンを分泌させる。胎生 14~20 週頃にはテストステロン濃度が急上昇し、脳の性分化が起こると言われている（「アンドロゲンシャワー」と呼ばれる。アンドロゲンはテストステロンに代表される男性ホルモンの総称である）。脳の性分化のメカニズムはヒトのデータは乏しいが、ラットの実験から得られた知見から、次のように推測されている。胎児は常にエストロゲンに曝されているが、胎児の血液中には  $\alpha$  フェトプロテインというエストロゲン結合タンパクがあってエストロゲンは脳血液関門を通過できない一方で、テストステロンは脳血液関門を通過でき、脳内でアロマターゼという酵素によって芳香化されてエストロゲンになり、脳内のエストロゲンレセプターに結合し、視索前野-視床下部や辺縁系の扁桃核に分布するこれらのレセプターをもつニューロンの「予定死」を促進したり阻止したりして雌型神経回路の形成を抑え（この過程で GnRH サージジェネレータが失われる）、その後テストステロンに反応する雄型神経回路が形成されるというのである（田中，1998）。精子形成は卵子形成と異なり、思春期以後は定常的に行われる。精子に発生していく細胞（精子形成細胞）は精細管の中に存在し、支持細胞からの栄養供給を受けて精子を形成する。精細管の最外側にある精子形成細胞は精原細胞である。精原細胞の一部は分裂する能力を保持し続けるが、一部は管腔側へ移動して成長期に入り、第 1 次精母細胞となる。第 1 次精母細胞は第 2 次精母細胞、精細胞をへて順次分裂しながら管腔内部へ移動し（第 2 次精母細胞から精細胞になるときに減数分裂が起こる）、精細胞の形が変わって、最終的に第 1 次精母細胞 1 個あたり 4 個の精子が形成される。最終産物である精子は支持細胞内端にしばらく宿ったあと精細管腔に

放出され、精巣上体に送られる。精子は射精されるまでの間、精巣上体にとどまって成熟する。性交時には、精巣上体から放出された精子が精管やそれに続く射精管、最後に尿道を通る間に、途中で混ざり合った前立腺や尿道球腺からの分泌液とともに、精液として射精される。一日あたりの精子生産量は約1億個で、精巣上体の貯蔵量は約4億個といわれ、1回の射精でその半分程度が放出される（ただし、1992年のWHOの基準では、1回の射精量が2ミリリットル以上、精子濃度が1ミリリットル中に2000万個以上なら正常ということになったので、現代では貯蔵量なども低下している可能性がある）。精子が授精能力をもつためには、射精1時間以内では50%以上の精子が直進運動性を示すか、高速直進する精子が25%以上あり、30%以上の精子が正常形態を示し、かつ子宮頸管粘液内でも運動性を失わないことが必要である。女性によっては子宮頸管粘液内に抗精子抗体をもつ場合があり、その場合は精子自体が正常であってもそのカップルは不妊になってしまう。近年、精子数が減少し、運動性も低下しているという報告がみられる。次節で詳しく触れるが、メタアナリシスによって1992年にこの現象を指摘したデンマークのスカケベック(N.E. Skakkebak)のグループは、環境内分泌攪乱物質の影響を示唆している。胎児期に通常より高いレベルのエストロゲンに曝露した場合、支持細胞の胎児期における増殖が抑えられて、精細管当たりの支持細胞数が減ってしまい、その結果精子数が減ることが動物実験で知られているので、胎児期の環境内分泌攪乱物質曝露が精子数減少を起こしている可能性もあるというのである(森, 1998)。しかし、精子数減少がみられない集団もあるし、精子数の減少がある集団でも、その原因は社会的要因あるいは食生活の変化など、環境内分泌攪乱物質以外である可能性もあるため、今後のデータの蓄積が待たれる。

性交 さて、いくら女性に受胎能力があっても男性に精子供給能があっても、適切な時期に性交をしない限り、生殖には至らない。目的論的に考えれば排卵前後に性行動が活発になれば効率がよく、多くの哺乳動物では実際にそうになっている。では、その生物学的メカニズムはどうなっているのだろうか。

哺乳動物でメスの性行動がはじめて起こるのは、個体が子どもを産んで育てられる大きさまで成長し、春機発動期(ヒトの思春期にあたる)に到達したときとされる。

「思春期に達して恋の悩みに身を焦がす年ごろになると、意中の誰かを思いながら漠然とした不安にさいなまれて食欲がなくなり、夜も寝られずそわそわと落ちつかなくなるのは、何も我々に限ったことではなさそうである」(市川ら『脳と生殖-GnRH神経系の進化と適応』pp.182)

春機発動期以降の牛や山羊のメスは3週間に1度発情し、16~20時間にわたって

特徴的な性行動をし、逆に摂食行動、休息、身繕い行動などの「自分を維持するための行動」は激減する。配偶者と首尾良く出会うためには行動量を増やさねばならないが、そのことは捕食者と出会う危険も増やすので、交尾が確実に受胎につながるような時期にのみ行動量が増えることが適応的である。交尾中の体勢が外敵に対して無防備であることを考えればなおさらである。つまり、行動系と内分泌系のリズムが共調する必要がある。その鍵となるのが GnRH である。GnRH は前述のように視床下部-下垂体-卵巣系を介して性ステロイドホルモンの分泌をコントロールすることによって行動を含むさまざまな生殖機能を制御しているばかりでなく、脳内で直接、神経機能を調節することで性行動の発現そのものを制御しているらしいのである（市川ら, 1998）。このことを初めてはっきりと示した実験は 1980 年頃、ラットを使ってなされたものである。ラットの雌は発情期になるとロードシスという特異的行動を発現する。ロードシスとは、雄に乗りかかられたりヒトに体を掴まれたりして触覚刺激を受けたときに、背を曲げ尻を上げて雄の乗駕行動を受け入れる姿勢を示す反射行動である。ロードシスが GnRH のコントロールを受けていることを示した実験は、卵巣摘除したラットにエストロゲンを投与しておき、中脳中心灰白質や視床下部腹内側核 (VMN) に微量の GnRH をカテーテルで直接投与するとロードシス商（一定の刺激に対してロードシスが起こる割合）が上昇し、逆に GnRH のレセプターに結合してその作用を阻害する物質（GnRH アンタゴニスト）を投与するとロードシス商が激減することを示したものであった。その後の研究から、中脳中心灰白質と視床下部腹内側核の間には密接な神経繊維の連絡があることや、腹内側核がエストロゲンによって引き起こされる発情期特有の生理的変化の中核であることから、中脳中心灰白質は腹内側核からの刺激を受け取ってロードシスという反射につなげる部位であると想定されている。腹内側核は満腹中枢としても知られ、GnRH レセプター、エストロゲンレセプターのみならずレプチンレセプターも存在し、腹内側核に隣接する弓状核 (Arc) にも GnRH レセプターとレプチンレセプターが存在し、弓状核で産生されるニューロペプチド Y (NPY) による食欲亢進シグナルがレプチンの結合によって制御されることが注目されている。思春期になって食欲がなくなるのも、これら神経伝達物質と視床下部ニューロン群の作用の結果と考えられるが、この点については次節で詳しく触れる。

ヒトにおいても排卵前後に性行動が活発になるという報告があるが、社会文化的要素が大きいため、結論はでていない。

**受胎** 適切な時期に妊孕力のある男女の間で性交があれば、受胎が成立する割合は 1 月経周期あたり 30% である（堤, 1999）。適切な時期とは、生きた精子がちょうど排卵された卵子に巡り会う時期であるから、卵子の受精能力が続くのは排卵後 48 時間以内であり、24 時間以内が高いことと、子宮や卵管内での精子の寿命が最大 3 日

程度であることを考え合わせれば、排卵の前後2日くらいと考えてよい。射精後の精子は子宮口から頸管内に入り、子宮腔を経て卵管内に入るが、この女性生殖路に在る間にキャパシテーションと呼ばれる現象が起こり、授精能力を獲得するのである。精子と卵子が会合するのは、主として卵管膨大部である。

先に述べたように卵子と精子は生成過程が異なり、精子は精原細胞から常に新たに分裂し続けているのに対して、卵子は原始卵胞から成熟卵胞になる過程に入るまで卵巣内で何十年も休眠状態にあるので、年齢が受精確率に与える影響は男女で異なっている。20代の女性では妊娠を希望し始めてから実際に妊娠するまでの平均月経周期数が3.3であるのに対して、40代では15.4というデータがある(堤, 1999)。

卵子に出会う精子は平均約200個であるが、それらが競争するように卵子周囲の卵丘細胞を取り崩した後、卵子を包む透明帯を突破した最初の1個の精子が卵子の表面に達し、先端部にある先体を卵表面と融合させて卵黄膜に小腔をあけ、精核を卵子の内部に送り込んでしまうと、卵子の周囲には受精膜ができて他の精子の侵入は阻止される。精核は卵核(卵子の核)と相融合して受精卵となる。受精卵は直ちに分裂を開始しながら、3~4日かかって卵管膨大部から子宮へ到達する。受精卵はそれから約2日の間に分裂を続けて胞胚期となり、同時に透明帯から抜け出て(ハッチング)、栄養胚板の酵素作用で子宮内膜上皮層を融解して子宮内膜内部に侵入する。その後5日程度で融解された子宮内膜は修復されて卵を覆うようになる。これを受精卵の着床という。着床が起こると受精卵表面の栄養胚板は充実性の小突起を発生し、それはやがて絨毛膜となってhCG(ヒト絨毛性ゴナドトロピン)を産生しはじめる。母体の血流を介してhCGは卵巣を刺激し、黄体を妊娠黄体へと変化させ、プロゲステロン分泌を継続かつ増加させる。今度はこのプロゲステロンが子宮内膜の脱落膜を発達させ、とくに発達した基底脱落膜(受精卵が着床した部位)が絨毛膜と一体となって胎盤を形成する。絨毛膜からのhCG分泌はかなり多量かつ特異的に起こるので、尿中hCGレベルの変化を調べることによって着床後7~10日というきわめて早い時期に妊娠を検出することができる。

母体にとってみれば受精卵の着床は、遺伝的に異なる情報をもつものが体内に入るという意味では臓器移植と同じなのに拒絶反応が起こらない(Rh不適合の場合のように拒絶反応が起こるのが希である)のは何故だろうか。これには2つの説明があって、1つは羊水中に胎児の抗原が多量に溶けていて、母体で抗体が作られても羊水中の抗原に消費されるというものである。もう1つは受精後48時間以内に胚(受精卵)から分泌が始まる早期妊娠因子(EPF; Early Pregnancy Factor)が胎盤で局所的に免疫抑制作用を発揮しているということである。EPFは基本的に胚によってしか作られず、しかも早ければ受精後3時間で分泌が始まるという報告があ

り、その意味では受精のマーカーとして hCG より優れている。しかし、EPF は定量が難しく、その本体に関してもチオレドキシソと複数の低分子物質の共役であるという説や、Fc レセプター様分子であるという説があり、論争が続いてきた。最近になって、EPF の本体がラットのミトコンドリアに含まれる熱ショックタンパクの1つであるシャペロニン 10(cpn10) と1残基の違い以外は一致することがわかった。分子としての cpn10 は他のタンパクの機能を修飾する働きをもつことが示唆されているが、胎盤での免疫抑制のメカニズムはまだ解明されていない。今後の研究の進展が期待される。

**在胎期間** 哺乳動物全般でみれば、在胎期間が短いほど産子数を増やすのに向いているが、その分マウスのように新生児が未熟である（晩成性）。一方、ウマやウシのように在胎期間が長いものはかなり個体として完成した状態で生まれてくる（早成性）。ヒト以外の類人猿の新生児も、体毛に覆われていて自分で体温調節ができるし、生まれてすぐに自力で母親にしがみつける。しかしヒトは、在胎期間が長いわりに例外的に晩成性である。これは、脳が巨大化したため、他の部分の成長を待って胎内にとどまっていると胎児の頭が大きくなりすぎて産道を通れなくなるからと考えられている（榎本,1999）。

妊娠中の内分泌機構は、妊娠の成立・維持及び胎児の発育のために大きく変化する。変化の主役は胎盤である。先に述べた hCG の分泌は妊娠 60~70 日頃にピークに達した後、急に 10 分の 1 のレベルに低下し、以後それを維持することが知られている。hCG レベルが低下した時点で妊娠黄体の内分泌機能も閉止し、以後は胎盤自身から分泌されるプロゲステロンが妊娠を維持する。エストロゲンも妊娠初期には卵巣からの分泌（エストラジオールが主）が大部分であるが徐々に胎盤からの分泌に移行する。胎盤から分泌されるエストロゲンは、妊娠末期にはかなり大量になり、血中濃度も排卵時のピークのさらに 1000 倍もの高レベルになるが、その大部分はエストリオールである。

生産に至った 198,408 例の白人のデータから計算されたヒトの在胎期間は、平均 37.4 週、分散が 4.98 週であり、やや左に歪んでいるものの正規分布に近い分布を示している (Wood, 1994)。一方、hCG や EPF を用いた早期の妊娠診断技術の発達の結果、早期胎児死亡の確率がかなり高いことがわかってきている。1976 年の Population Studies に掲載されたルリドン (H. Leridon) の妊娠歴の遡及聞き取りによる胎児死亡率は 12% と推定されたが、hCG を用いたフォローアップ研究では 31% から 62%、EPF を用いたフォローアップ研究では 89% という値が報告されている。バングラデシュの集団で hCG を用いた研究の結果では 86.7% という値が出ていることから、集団による違いも示唆されている。胎児死亡には母の年齢やパリティの影響もあり、年齢やパリティが高くなると胎児死亡率も上昇することが報告

されている。早期胎児死亡後の無排卵期間はバングラデシュのデータでは最頻値が29日、次いで58日である (Holman, 1996)。人工流産や自然流産後の無排卵期間の平均は3.4~4.1週であり (Wood, 1994)、ほぼ同等である。早期胎児死亡は無自覚に起こるが、無排卵期間ができることにより、実質的に月経周期が延長したのと同等の効果をもたらす、出生率の低下に寄与する大きな要因である。

#### 5.1.4 生殖戦略のまとめ

ヒトは、歴史の中で「育児コストのかかるサル」としての進化を遂げてきた。その生殖メカニズムは、実に巧妙に視床下部-下垂体-性腺系によってホルモン制御されており、外部からの刺激に対して比較的ナイーブである。このことは、ヒトの寿命が長く、産子数が哺乳類の中では少ない方であることと無関係ではない。つまり自己の生命維持についての安全装置をかけた上で、できるだけ子どもを残そうというのがヒトの生殖戦略であるといえよう。

#### 5.1.5 出生力と妊孕力

既に述べたように、ヒトの女性が一生の間に生む子ども数の分布には、集団によって大きな違いがある。いいかえれば、ヒトの出生力 (fertility) には集団による多様性がみられる。出生力の多様性を説明する要因として考えられることは、1つには、意図的な出産抑制や結婚年齢、あるいは通い婚とか結婚形態に依存して変わる性交頻度など、社会的な因子である。ここでは詳しくは触れないが、社会的因子の出生力への影響はきわめて大きいと考えられる。しかし、社会的な因子による制約がまったくないとしても、なお出生力の多様性は残る。北米 Hutterites とパプアニューギニアの Gainj は同じように若くして結婚し、意図的な出産抑制はないにもかかわらず、平均完結出生力は Hutterites が10以上であるのに対して Gainj は4.3と、大きな違いがある。この違いを説明するのは、生物学的な妊孕力 (fecundity) の差異である。妊孕力を規定する要因は、大きく分けると遺伝要因と環境要因からなる。遺伝要因は、一般にはヒトという生物に共通のものと考えられるが、細かくみれば地域集団ごとに各々の地域環境に対して独自の適応を果たし、多様になってきている面も見逃せない。環境要因は、地域集団ごとに異なっていることはもちろん、地域集団内でも多様であるし、時間とともに変化する点が大きな特徴である。

なお、妊孕力を厳密に定量化するには、妊孕力のある受胎待ち時間 (fecund waiting time)、あるいはその逆数である受胎能力 (fecundability) という概念が用いられる。一般に受胎能力といえば、自然受胎能力、すなわち、避妊を行わない場合の、1月経周期あたりの受胎確率をさす。これらの値には実際には性交頻度や男性因子も寄与するのであるが、

女性の妊娠歴データをもとに推定されるので、女性の生物学的な受胎の能力の指標とみなされる。

### 5.1.6 妊孕力の遺伝要因

遺伝要因としては、単純に考えると多産の遺伝子や少産の遺伝子、あるいは性比を偏らせる遺伝子といったものが思いつく。子沢山の家系とか女系家族は実在するから、そういう遺伝子があっても良さそうであるが、まだ特定されていない。環境要因の影響が大きいことやヒトの再生産期間全体を含むような長期的なデータが必要なこともあって、そういう gene-targeting を行うのがきわめて難しいことも一因である。

無視できないのは、遺伝的な低妊孕力である。子宮頸管内の抗精子抗体の存在や、Rh-の女性と Rh+ の男性の間の第2子の胎児死亡など、夫婦の組み合わせによって不妊になる場合も考慮する必要がある。この点については、1つの集団内でも個人差があることに注意せねばならない。例えば、北米に住み意図的な出産抑制をせずに伝統的な生活をしているアーミッシュと呼ばれる人々の子ども数分布のハザード解析においては、集団内の妊孕力が均質でないと仮定した方が、均質としたモデルよりも高い説明力が得られた。この結果は、結婚時点から妊孕力が低いカップルと、そうでないカップルが存在することを示すものである。低妊孕力をもたらす遺伝因子はいくつも示唆されているのだが、ここでは最近新たな知見が相次いでいるレプチンに焦点をあてることにする。

### 5.1.7 内分泌系機能の遺伝子支配

肥満マウスとして知られる ob/ob マウスでは遺伝的に生殖能力が低い。ob/ob マウスが交配中に偶然生じたのは1950年だが、遺伝子としての変異部位とその物質が同定されたのは1994年のことである。レプチンと呼ばれるこの物質は主として脂肪組織から分泌されるホルモンで、正常なレプチンがないと視床下部の満腹中枢が活動しないためにいくらでも食べてしまうばかりでなく、交感神経系の働きも低くなるために産熱が低下してエネルギー過剰状態になることから、極端な肥満になってしまうのである（蒲原，1998）。ヒトも含め、ほとんどの脊椎動物でレプチンの配列は保存されており、食欲を抑制しエネルギー消費を増大させる機能をもっている。では、正常なレプチンがないと生殖能力が低くなるのは何故なのだろうか。

この問いに答えるには、レプチンが脂肪組織で作られ、分泌されるという特徴に注目する必要がある。生殖にはたくさんのエネルギーが必要であり、女性は第二次性徴の発現にともなって体脂肪率を増やしてゆくことが知られている。つまり、レプチンは十分な生殖に耐えうるだけの十分な体脂肪の蓄積が完了したことを脳に伝えるシグナルなのではない

だろうか、ということである。女性スポーツ選手などで体脂肪が低いと思春期の遅延が認められる場合があることは、この仮説の傍証となる。ob/ob マウスに実験的にレプチンを投与すると生殖機能が回復することも確認されている。さらに、栄養状態がよくなった先進国で初経年齢が早まっていることも、この仮説を支持する。マウスの実験では、レプチンを投与して自由に摂食させた場合はレプチン投与なしで自由に摂食させた場合に比べて摂食量は 80% 程度になり体重も軽くなるが膣開口時期 (VO) は同程度である。一方、レプチン投与なしで摂食量を 80% に減らすと体重減少はレプチン投与群と差がないが VO は有意に遅れることがわかっている。したがって、十分なレベルのレプチンがあることは、生殖の必要条件といってよい。しかし、レプチンは生殖を起こす十分条件ではない。というのは、いくらレプチンを投与しても、摂食量を自由摂食時の 36% あるいは 63% にまで減らすと有意に VO が遅れるからである (Cunningham et al., 1999)。

レプチンが視床下部-下垂体-卵巣系に影響を与えるメカニズムは完全にはわかっていないが、最近急速に研究がすすみ、多くの知見が得られてきている。レプチンの作用部位にはレプチンレセプター (Ob-R) があるわけだが、Ob-R はクラス I のサイトカインレセプターファミリーに属し、1 度だけ細胞膜を貫通する膜タンパクである。大きく分けると細胞内領域が長いものと短いものの 2 種類があるが、細胞外領域は共通である。長いタイプを刺激するとシグナル伝達系を活性化させ、また転写タンパクの活性化因子を作動させるが、短いタイプには少なくとも見かけ上はこの活性がない。正常なレプチンを産生できない変異だけでなく、Ob-R が未熟なままに切断されてしまう変異でも肥満や性腺発育不全が起こることが知られている。Ob-R とそのメッセンジャー RNA (mRNA) の局在の分析結果から、Ob-R は視床下部-下垂体-卵巣系のすべてに渡って発現していることがわかっている。レプチンは、卵巣において LH の刺激によるエストロゲン産生に影響すること、下垂体における LH や FSH 分泌に影響することが *in vitro* の実験からわかっているが、それ以上に視床下部において GnRH パルスジェネレータに直接作用することが重要である。

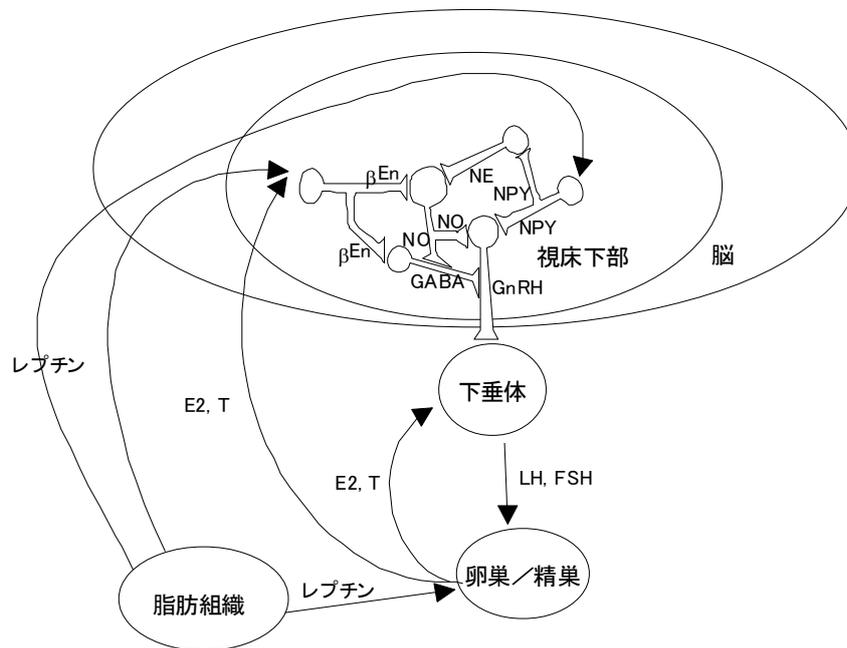


図1. レプチンが生殖機能に影響を与えるメカニズムの模式図  
 (E2:エストラジオール, T:テストステロン, βEn:ベータ  
 エンドルフィン, NE:ノルアドレナリン, NPY:ニューロ  
 ペプチドY, NO:一酸化窒素, GABA:ガンマアミノ酪酸,  
 GnRH:ゴナドトロピン放出ホルモン, LH:黄体化ホルモン,  
 FSH:濾胞刺激ホルモン)

すべての哺乳動物で Ob-R mRNA は視床下部の弓状核 (Arc) と腹内側核 (VMN) に発現している。前節で述べたように弓状核には NPY を産生する神経細胞体があり、分泌された視床下部性 NPY は外側野 (LHA) の摂食中枢に作用して食欲を亢進させ、同時にノルアドレナリンニューロンからの交感神経系を抑制してエネルギー消費を減らす働きがある。レプチンが欠損している ob/ob マウスでは弓状核における NPY の発現量が增大していることと、このマウスにレプチンを投与すると NPY 発現量が低下することから、レプチンの食欲抑制作用とエネルギー消費亢進作用は NPY の働きを抑制することで実現されていると考えられる。レプチンレセプターに変異があってレプチンを結合できない db/db マウスと呼ばれる肥満マウスにおいても NPY の発現量が增大していて、レプチンを投与してもそれが変化しないことから、この仮説は支持される。飢餓状態の動物でも弓状核における NPY の発現が亢進しており、レプチンを与えると NPY レベルが低下することから、NPY の産生自体がレプチンによって制御されていることが示唆される。一方、弓状核においては、POMC (副腎皮質刺激ホルモンやβエンドルフィンなどいろいろなメラノコルチンの前駆体) を含むニューロンの 30% が Ob-R も発現していて、弓状核における POMC の mRNA 発現量はレプチンの作用によって増加することがわかっている。POMC からはβエンドルフィンが産生される。前節で述べたように、GnRH パルスジェ

ネレータはオピオイドニューロンを介してエストロゲンの負のフィードバック作用を受けているが、レプチンの作用によって $\beta$ エンドルフィンレベルが上昇すれば、この経路が修飾されることは明らかである。つまり、弓状核におけるレプチンの作用は、NPY産生を抑えてノルアドレナリンニューロンを活性化すると同時にPOMC産生を増やしてオピオイドニューロンを刺激することである。それら2つが一酸化窒素ニューロンを刺激して、直接GnRHパルスジェネレータを制御する一方で、オピオイドニューロンと一酸化窒素ニューロンはGABAニューロンを介してもGnRH分泌を制御すると考えられるのである(Faletti et al., 1999; Cunningham et al., 1999)。レプチンレベルが脂肪細胞の量によって決まってくることを考えれば、脂肪蓄積に関連するすべての遺伝素因が生殖に影響するといつてよいことになる。例えば、ピマインディアンでエネルギー節約遺伝子ではないかと提唱された $\beta 3$ アドレナリンレセプター遺伝子の変異も、生殖機能に影響を与えている可能性がある。中立でない変異には急速に淘汰がかかるのが普通だが、エネルギー摂取が限られた条件下では生存にとっては有利だが生殖にとっては不利だという変異が選択される境界条件もありうるのである。

ところで、肥満マウスでの生殖能力が低いことから連想すると、先進国における肥満の増加を出生力低下と結びつけて考えたいところである。しかし、ヒトの肥満においては、95%の症例で体脂肪量と血中レプチン濃度の間に正の相関が見られており、今日の先進国でみられるヒトの肥満の多くはレプチン不足に起因するものではないことがわかっている。直接の関連を考えるのはやや早計である。ただし、レプチン不足でないにもかかわらず肥満になってしまうということは、脳内におけるレプチンレセプターあるいはその作動系の異常が想定されるので、関連がないともいいきれない。

レプチンレベルが上昇しても生殖機能が亢進するわけではないのに対して、レプチン低下時の生殖機能低下は非常に敏感である。人類の長い歴史においては飢餓状態の方が長かったため、飢餓に陥ったことを脳に知らせ、生殖機能や甲状腺ホルモンの分泌を抑えてできるだけ消費エネルギーを節約しようとするシグナルの存在は重要であったと思われる。脊椎動物というレベルでレプチンの配列が強く保存されていることは、このシグナルの重要性を示すものであろう。

なお、最近の研究で、レプチンは胎盤でも産生され、胎児の発達にも何らかの役割をもっていることが示唆された。また胃でも産生され、コレシストキニンの分泌を促進することで短期的食欲抑制も行っているらしいが、これらの機能の生殖との関係はまだ明らかでない。

### 5.1.8 フェロモン

環境要因に入る前に、遺伝要因とも環境要因とも区別しがたい要因について触れねばならない。すなわちフェロモンである。フェロモンというと、アリの道標フェロモン、ゴキブリの集合フェロモン、ミツバチの女王物質といった昆虫のものが有名である。要は、動物の個体が体内で生産し、体外に少量分泌して離れたところの同種他個体に生理学的あるいは行動上の影響を与える物質、ということである。フェロモンという名前は、ギリシア語の *pherein* (運ぶもの) と *horman* (刺激するもの) からの合成語である。アリの道標フェロモンだったら、あるアリが出した道標フェロモンの後を伝って他のアリがついてくるわけである。本節の主題である妊孕力に関連するフェロモンは、一般に性フェロモンと呼ばれる。カイコガの処女メスから分離されている「ボンビコール」は化学構造までわかっていて (*trans-10, cis-12-hexadecadien-1-ol*)、オスを誘引する能力をもつことが知られている。逆に、マダラチョウの1種 *Lycorea ceres ceres* のオスはヘアペンシルという器官から *2,3-dihydro-7-methyl-1H-pyrrilizin-1-one* を出すことによってメスを誘引する。オスが出す物質の方が近距離で作用するものが多く、催淫物質とも呼ばれる。フェロモンが妊孕力に与える影響は実に広範であり、ヒト以外の動物においては古くから知られていた。たとえば、マウスのブルース効果という現象は、妊娠マウスが胎児の父親と違う MHC の型をもつオスマウスの尿の臭いを嗅ぐと流産してしまうというものである。

近年になって、ヒトでもフェロモンは無視できない作用をもっていることが明らかになってきた。前節で述べたように女性が MHC の型が異なる異性の分泌物の臭いを好むことから配偶者選択にもフェロモンが作用していると考えられることに加え、月経周期もフェロモンによる調節作用を受けているらしいことがわかってきている。

月経周期へのフェロモンの影響を示した実験研究は、1998年の3月に発表されている。このデータはかなりクリアに揮発性の物質による月経周期の同調を示している (Stern and McClintock, 1998)。著者の一人マクリントック (M.K. McClintock) が、寄宿舎生活をしていて同じ部屋に住んでいる人たちの月経周期が同期してくることに気づき、4~6ヶ月も一緒に住んでいると月経開始日が2日以内に近づくことを発見して、「女性は月経周期を同期させるフェロモンを出している」という主旨の論文を1971年の *Nature* に発表したのが、問題の発端である。彼女は、それ以後もずっと「他人の月経周期を左右するフェロモン」というアイデアにこだわり続けてきたが、批判は多かった。中でも、マリ共和国に居住するドゴンと呼ばれる人々の調査データに基づいて「月経は個人ごとに独立に起こっているとした場合からの有意なずれはない」というストラスマン (B.I. Strassmann) の批判は痛烈であった。数々の批判に答えるべく、とうとう実験に踏み切った結果が、1998年の報告なのである。彼女としては「月経周期の同期」よりも「他人の月経周期に影響を

与えるフェロモンの存在」を示すことが先決なので、動物実験でやったように（既にラットでは証明済み）、個体の分泌物を他個体に嗅がせて、月経周期に影響するかどうか見ればよいのである。

ある一つの大学の、20歳から35歳までの学生とスタッフから29人のボランティアを募り、9人を分泌物ドナー兼キャリアコントロール（分泌物の臭いを脳で判断されて影響が出ると困るので、臭い消しの意味もあって70%イソプロピルアルコールをキャリアとして分泌物に添加してから被験者に嗅がせているため、キャリア自体の効果を調べる必要があってドナーにキャリアを嗅がせたのである）とし、残り20人を被験者とした。

ラットの実験結果から、仮想「フェロモン」は2つあると予想されていた。前節で述べたように女性の月経周期は排卵を境にして濾胞期と黄体期からなっているが、マクリントックの仮説は、濾胞期には他の女性の月経周期を短くするフェロモンが出て、排卵時には他の女性の月経周期を長くするフェロモンが出て、この2つの相互作用によって結果的に全個体の月経周期が同期してくる場合があるということである。

実験デザインは、以下の通りである。ドナーには毎日香料を使わずに入浴してもらって、その後少なくとも8時間、腋窩にコットンパッドをあててもらい、イソプロピルアルコールを4滴垂らして凍結保存する。濾胞期と排卵時別々にラベルする。便宜上これらを「濾胞期フェロモン」、「排卵期フェロモン」と呼ぶことにする。フェロモンの効果にもベースラインの月経周期にも個人差が大きいことが予想されるため、デザインはクロスオーバー法が用いられた。つまり、被験者20人を2群にわけ、群Aにはベースラインの1月経周期の後、2月経周期の間「濾胞期フェロモン」を与え、続く2月経周期の間「排卵期フェロモン」を与える。一方、群Bにはベースラインの1月経周期の後、2月経周期の間「排卵期フェロモン」を与え、続く2月経周期の間「濾胞期フェロモン」を与える。ドナーにはサンプル採取の1月経周期の後、2月経周期の間、イソプロピルアルコールを与え続ける。見るべきデータは月経周期の長さである。

結果は非常にクリアで、濾胞期フェロモンを与えると月経周期は2日弱短縮し、排卵期フェロモンを与えると月経周期は2日弱延長していて、これらの差は統計的に有意だった。キャリアのみでは月経周期の長さには何の影響もなかった。周期の中身を細かく見ると、濾胞期の長さが変化していた。つまり、排卵が早くなったり遅くなったりしていた、という結果であった。

月経周期の変化は妊孕力に大きな影響を与える要因であり、これがフェロモンの影響を受けているということは、居住形態などが妊孕力に影響する可能性があるということである。今後の研究では考慮すべき要因であろう。

### 5.1.9 妊孕力の環境要因

ヒトの妊孕力に影響する環境要因としては、性病などの病気、授乳、栄養状態、心理社会的ストレス、季節変化、および環境内分泌攪乱物質の影響が主なものである。いずれもヒトの社会の複雑化にともなって個人間差や集団間差が大きくなってきていると思われる。以下、これら6つの要因について詳しく説明する。

**病気** クラミジアや淋病など、いくつかの性病は卵管閉塞（あるいは閉塞までいかなくとも卵の自然な移動を妨げるような卵管の傷）などの後遺症を引き起こし、不妊の原因となる。性病以外にも IUD の使用などによる外傷や、産褥敗血症、あるいは性器結核といった疾患でも後遺症として卵管の異常につながる場合がある。女児の割礼がある集団や、非衛生的な人工流産法を実行している集団では、それらも原因となりうる。質の良い疫学データは少ないが、多くの自然出生集団では性病は少なかったと考えられている。ただし、中央アフリカやいくつかの太平洋の島々のような例外もある。性病のある集団では、完結出生児数の分布が、性病のない自然出生集団とははっきりと異なることが指摘されている。例えば、サハラ以南のアフリカでは、不妊を訴えるカップルの平均年齢が他地域に比べて有意に若いのである。その結果、完結出生児数が0のカップルの頻度が高くなってしまふ。しかし、サハラ以南のアフリカでさえ性病の蔓延は最近100年以内に、より広い世界と接触して起きたことと考えられる。産褥敗血症による二次的不妊の影響の大きさを指摘した報告は多いものの、それを定量した研究はない (Wood, 1994)。

性病の中で最近注目されているのはクラミジアである。クラミジアは以前より人畜共通感染症として、失明を引き起こすこともあるトラコーマの病原体として知られていたが、実は卵管炎を起こす可能性があつて不妊に大きな影響をもつことがわかってきたのである。クラミジアは細胞内に封入体を作り検出が困難な上、感染の自覚症状に乏しいために、性行為によって感染し、気づかないうちに骨盤内や腹腔内で増殖して各部で炎症を引き起こし、卵管の閉塞や卵管周囲や腹腔内の癒着をきたして卵管性不妊症になるリスクが高いのである。日本において現在、不妊症を訴えて外来を受診した女性について抗クラミジア抗体を調べると数パーセントから十数パーセントの割合で陽性であり、既に無視できないほどの影響があるものといつてよい。男性においても尿道炎や前立腺炎などを発症して精子の運動性を低下させるので、妊孕力低下の原因となる。

2004年に旭川医大の今井助教授らによって発表された調査結果（尿検査でPCRによりクラミジアDNAを増幅して検出、自覚症状があるものは除く）によれば、性

交経験のある高校生の約1割からクラミジアが検出されており、従来考えられていたよりも、日本人の若者のクラミジア感染割合は1桁高いことがわかった。

**授乳の影響** 先進国の女性は栄養状態がよいので授乳中でも月経が再開するということがいわれてきた。しかし、栄養状態がよくても頻回な授乳があれば吸乳刺激の影響で GnRH パルスジェネレータが抑制され、月経が起こらないことがわかっている。したがって授乳には、母体からのエネルギー損失による栄養状態の低下という意味だけでなく、GnRH 分泌を抑制する別のメカニズムがあるといえる。古くは吸乳刺激によって下垂体前葉から分泌されるプロラクチンが GnRH や LH 分泌を抑制するものと考えられていたが、プロラクチン分泌を抑制しても吸乳刺激があれば LH パルスが抑制されることから、現在ではプロラクチンが決定的であるとはみなされていない。吸乳刺激によって乳汁が排出されるメカニズムは、乳房や乳頭付近に分布する知覚神経が脊髄を上行し、視床下部室傍核及び視索上核に辿り着くことがわかっており、そこから何らかの信号を受けた下垂体後葉からオキシトシンが分泌され、オキシトシンが乳腺に作用して乳腺胞を取り巻く筋上皮細胞を収縮させ、乳汁を乳腺外に排出するということがわかっている（市川ら, 1998）。後で述べるように室傍核はストレスを受けたときにも活性化されることから、それと同様な機構で GnRH 分泌が抑制されている可能性はある。しかし矛盾する実験結果もあって、オピオイドニューロンの関与も含め、この経路については決定的なことはわかっていない。

マクロな視点で見ると、いろいろな集団で授乳による無月経の影響を調べた報告が多数蓄積されており、授乳が月経周期を延長する効果をもつことは間違いない。とくに授乳強度が決定的であることは、1982年に報告されたエディンバラでのフォローアップ研究から明らかになった。授乳頻度が1日あたり6回未満になり、授乳時間が1日あたり60分未満になると排卵が再開したのである。1日あたり80分未満で再開という報告もあるが、注目すべきなのはボツワナの!Kung と呼ばれる人々やパプアニューギニアの Gainj と呼ばれる人々における、1日あたり40~50回という高頻度で、かつ1回あたりの授乳時間がきわめて短いパターンでも、2年にわたって排卵が抑制されるという報告である (Wood, 1994)。授乳強度のなかでも頻度と時間のどちらが決定的かということについては長い間論争が行われてきたが、最近得られた知見によれば、とくに夜間に長い授乳間隔が空くと排卵が再開することから、頻度の方が決定的である可能性が強いようである (Wood, 1994)。

**栄養状態の影響** 前節で述べたように、哺乳類の排卵は、視床下部、下垂体、卵巢が密接に情報を交換することによって調節されるのであるが、この排卵調節のメカニズムは、絶食や激しい運動の影響を強く受ける。先進国の女性の排卵調節機構は比較的良好に保たれ約28日ごとの排卵が保持されるが、カラハリ砂漠中央部に住む San

と呼ばれる部族の女性は激しい労働と栄養不足の影響で、1年のうち初春にしか出産がみられない。初夏から秋にかけての食物を得やすい時期にだけ排卵が起こっていると考えられる（田中, 1998）。

若い女性が無理なダイエットから無月経になることが時折みられるが、カニクイザルを使った実験の結果からも、絶食をさせている間は無月経となり、食事を再開するとすぐに月経が回復することが報告されている。インスリンの投与によっても同様な現象が引き起こされることから、絶食によって血糖の利用性が低くなり、LHサージが起こらなくなると考えられる。いまのところ、このメカニズムは次のように考えられている。

2-デオキシグルコース (2DG) を投与すると血糖値を下げずに血糖利用を阻害することができる。この場合にも LH パルス頻度は低下し、発情周期も阻害される。しかし、糖尿病の患者では血糖値が高くても LH 分泌は抑制されている。これは、糖尿病ではインスリンレベルが低いかインスリン抵抗性（インスリンに細胞が反応しない状態）のため、細胞が血糖を取り込む能力が低く、飢餓状態にあるからと考えられる。しかし、血糖値の低下がそのまま脳内での細胞の活動を低下させるわけではなく、おそらく延髄最後野に存在すると考えられる糖センサーが血糖値低下を感知し、隣接する弧束核から視床下部の室傍核に投射するノルアドレナリン作働性ニューロンを刺激して室傍核でのノルアドレナリン分泌を高め、これが摂食行動を誘起すると同時に GnRH 分泌を抑制するものと考えられている（市川ら, 1998）。

**心理社会的ストレスの影響** ストレスが生殖機能を抑制するメカニズムとしてわかっているのは、1つはオピオイドニューロンの活性化にを通じての経路である。社会心理的ストレスばかりではなく、痛み、激しい運動、低栄養、マラソンなどによる肉体的なストレスによってもオピオイドニューロンは活性化され、前節に述べた GnRH パルスジェネレータの働きが抑制されるため、LH や FSH のパルス状分泌も抑制され、性ホルモンの分泌も低下し、全体として生殖機能が抑制されるのである。

また、ストレス負荷時には視床下部室傍核で産生される副腎皮質刺激ホルモン放出ホルモン (CRH) の下垂体門脈への分泌が増加し、続いて下垂体前葉から副腎皮質刺激ホルモン (ACTH) が分泌されて副腎皮質を刺激し、副腎皮質ホルモン分泌が増大することが知られている。このとき脳内でも CRH 分泌が高まるが、室傍核に細胞体をもつ CRH ニューロンの神経終末は視索前野の GnRH ニューロン細胞体あるいは正中隆起の終末と隣接しており、これらの部位で直接 GnRH 分泌を抑制しているものと考えられている（市川ら, 1998）。

**季節変化の影響** 生殖全体を考えれば、排卵前後に性交することだけでなく、在胎期間を経て出生した子どもが生存しやすい時期を選ぶことが繁殖成功につながることは明らかである。高緯度地方では食物の乏しい冬よりも、食物が豊富で気温も過ごしや

すい春から秋に生まれた子どもの方が死亡率が低くなるであろうことは想像に難くない。実際、ヒト、ボノボ、コモンマーモセット、ワタボウシタマリン、シシオザル、ブタオザルなど例外はあるものの、多くの霊長目は発情季をもち、繁殖頻度が季節によって変化する。たとえばニホンザルでは、3月から8月にかけてが出産季であり、交尾季は9月から2月である。非交尾季にはメスの下垂体から分泌されるゴナドトロピンレベルも卵巣から分泌される性ホルモンレベルも低く、排卵が起こらないし、オスでも精巣から分泌されるテストステロンが少なく、精子形成が不活発で精細管径も小さいので、自然状態での非交尾季には、たとえ交尾が行われたとしても妊娠する可能性はきわめて低いのであるが、そもそも排卵前にしかメスは発情しないのが普通である（ニホンザルの場合は1週間から10日程度前から発情する）。発情とは、(1) オスを性的に引きつける「魅力」があり、(2) オスへの誘い行動をする「積極性」があって、(3) 性交渉が実際に行われる「受容性」があるという3つの行動要素が揃ったメスの状態であるとされているが、ニホンザルの場合、発情期になると外見からはっきりそれとわかるのである。まず顔や尻の性皮が赤く色づき、子宮頸部から粘液が多量に分泌されて尻が濡れ、特有の臭気を発するようになる。同時に活発にオスに接近し、発情期特有の「ウアー—」という「発情鳴き」をする。オスは排卵の1~2日前まではこれに攻撃的な行動をするが、それを過ぎてメスがエストロゲン分泌のピークを迎えると攻撃が抑制されて性的な接近をし、交尾をするようになる。シシオザルやブタオザルでは性皮の最大腫脹がちょうど排卵時にあたっていて、口をもぐもぐさせる「リップ・スマッキング」など、オスがメスを交尾に誘う行動も活発になる。霊長目ではこうした生理的要素だけで交尾相手を選択されるのではなく、「親しさ」や「優劣順位」といった社会的要素の影響も大きくなる。たとえば非交尾季に毛づくろいを頻繁に交わしあう「親しい」相手とは、交尾季になっても交尾頻度が低く、近親相姦を避けるためのメカニズムとして作用している（榎本, 1999）。

ヒトにおいては、北極近くに住むイヌイトの女性は一日中暗闇の冬季には排卵が抑制されるという報告がある他、ヨーロッパ、カナダ、日本などでの歴史人口についての詳細な解析から出産には季節性があり、しかもそのパタンが夏に多くなるアメリカ型や早春と秋に多いヨーロッパ型など地域によって異なることが明らかになっている。これは性行動が生理的及び社会的に季節性の影響を受けることばかりでなく、精子の質や早期胎児死亡の季節差などとの相互作用によることを示している（三浦, 1983）。なお、季節的な発情の生理的メカニズムとしては、日長変化が網膜の桿体細胞から網膜視床下部管と呼ばれる神経を經由して視床下部の視交叉上核に伝えられ、室傍核、上顎交感神経節を經由して松果体を刺激してメラトニン産生をコントロールし、メラトニンによって視交叉上核や下垂体隆起部が刺激され GnRH

や LH のパルス分泌がコントロールされるという経路が想定されている。メラトニンは夜の長さを示すホルモンであり、ヒツジでは日長が短くなると循環血中のメラトニンレベルが上昇し、LH パルス頻度が上昇することがわかっている（市川ら、1998）。

**環境内分泌攪乱物質の影響** 最近、いわゆる環境ホルモン、すなわち環境中に極微量に存在し、それを生体に取り込むことによって生体の内分泌系が攪乱される物質の生殖への影響を示唆するデータが続々と発表されている。正式には環境内分泌攪乱物質と呼ばれるが、この中にはビスフェノール A のようにエストロゲン類似物質で直接エストロゲンレセプターに結合し、エストロゲンが結合したときと同じ作用を示したり、あるいはエストロゲンが結合できなくなるためにエストロゲンの作用を抑制すると思われるものや、ダイオキシンのようなアリルヒドロカーボン (Ah) レセプターに作用することで間接的に性ホルモンの精密な調節機構を狂わせる物質が含まれるが、詳しいメカニズムはわかっていない。

環境内分泌攪乱物質の生殖への影響として注目されているのは、ヒトの精子数が減少しているという報告である。ショッキングな論文が発表されて、研究過熱の引き金が引かれたのは 1992 年のことである。研究グループのリーダーの名をとって、俗に「スカケベック論文」というが、最近 50 年間にわたって、ヒトの精子数は減少しつづけているという内容である。当然のことながら、大議論が沸き起こった。著者たちは直接環境内分泌攪乱物質の影響だと断じたわけではなく、「真の減少があった」ことと、「精巣ガンや停留辜丸や尿道下裂のような尿生殖器の異常の発生率が上昇していることと併せて考えると、男性の生殖腺の機能に深刻な影響を与える要因が増大していることが示唆される」と言っており、本文最終パラグラフでは、「エストロゲンかその化合物、またはエストロゲン様の作用をもつ化学物質、あるいは他の環境あるいは内因性の要因が精巣の機能に障害を与えるかどうかは、今後はっきりさせねばならない」と書いているだけであるが、彼らが 1998 年に書いたレビューでは環境内分泌攪乱物質が原因で精子数が減少していることと、子宮内膜症も同じ原因で増加していることが主張され、やや強い調子になっている。

この論文は過去の文献値を再集計したメタアナリシスなので、(1) サンプルに偏りがあるか、あるいは統計処理がまずい、(2) フィンランド人の精子は 1958 年から 1992 年の間まったく変化が無く、今も 1 億 3300 万個を維持している、(3) 米国では過去 25 年間減っていない、といった批判もあったが、(1) スコットランドでの 3792 人の男性の調査で 1940 年生まれの男性は 1 億 2800 万個/mL だが 1969 年生まれでは 7500 万個/mL に減少していた、(2) ベルギーで 1977 年から 1980 年に採取された精子と 1990 年から 1993 年に採取された精子の比較の結果、正常な形の精子の割合が 39.6% から 27.8% に減少し、正常に運動する精子の割合も 53.4% か

ら 32.8% へ低下した, (3) パリの精子銀行のサンプルで 1945 年生まれと 1962 年生まれの男性の 30 歳時点での精子を比較した結果, 1945 年生まれは 1 億 200 万個/mL, 1962 年生まれは 5100 万個/mL とほぼ半減し, 正常形態割合, 運動率とも低下していた, (4) ドイツでは 7800 万個, 香港では 6200 万個, ナイジェリアでは 6400 万個, パキスタンでは 7900 万個, といずれも 1 億個未満のデータで低下が示唆される, といった「精子数減少」を支持する報告もあり, 大論争となった。

その後, スカケベック論文と同じデータを厳密に再分析することで, 精液の質の低下が artifact なのか現実なのか検証することを目的に少なくとも 4 つの論文が発表されているが, そのうちアメリカ国立環境健康科学研究所 (NIEHS) の研究者であるスワン (S.H. Swan) が報告したものがかなり厳密に行われていて, かつ有意な減少という結果だったため, 環境内分泌攪乱物質の研究者にはよく引用されている (例えば, 森, 1998)。

しかし実はスワン論文は決定的とはいえないことに注意すべきである。分析対象はスカケベック論文の 61 文献から非英文論文 3 篇と不妊治療後に授精に成功した男性のデータを含む 2 篇を除いた 56 文献, 14947 人 (もとの 97%) を使用し, 元文献から抽出した, 平均精子密度 (あるいは中央値), 出版年, 研究地域 (US, Europe / Australia, Non-Western), 研究目的, 授精能力があるとわかった男性の割合, 精液採集法, 年齢, サンプル前の禁欲期間である。方法としては, これを地域別に層別化し, 平均精子密度と傾きの地域間差を評価するための交互作用項を含めた重回帰分析を行ったのである。重回帰は, 線型, 階段, スプライン, 2 次式のモデルについて SAS の GLM プロシージャで実施し, サンプル数で重み付けしている。結果としては, 有意な地域間の差があり, かつ欧米においては精子数が単調減少しているというものであった。しかし, 例えば禁欲期間について, 3 日後より 10 日後の方が精子濃度が 50-69% 高いという報告や, 平均禁欲期間が 7.5 日 (1956 年) から 4.4 日 (1986 年) に短縮したという報告があることを踏まえた上で, プロトコルが推奨する禁欲期間を守っていた文献が 30 しかなかったことなどを考え合わせると, サンプルバイアスが存在する可能性は否定できない。さらにスプラインモデルと線型モデルの適合度があまり変わらないことから, 50 年間の単調減少をいうにはあまりにナイーブな結果なのである。それゆえ結論は決定的とはいえず, 減少している集団を特定して減少の要因を探ることが重要であろうと思われる (中澤, 1999)。

### 5.1.10 社会経済的側面

日本の人口学者は社会学や経済学出身者が多いので、この分野の研究は盛んに行われているが、多くは海外の研究者が開発した理論を日本のデータに適用できるかどうかを検討したというレベルにとどまっている。

「出生力転換」という言葉があって、近代化に伴ってどの国でも同じように出生力が低下したといわれているが、それを検証するためには、経時的な国間比較のための出生力指標が必要である。プリンストン研究（Ansley J. Coale が中心となってヨーロッパの出生力転換を研究）で、Coale はデータが不十分な 18 世紀、19 世紀のヨーロッパ各地の出生力、結婚、結婚出生力の統一的な尺度として、いくつかの間接的出生力指標を考案した。一名ハテライト指標と呼ばれるこれらの指標は、人口史上の最大自然出生力集団であるハテライトの年齢別出生率を基準とした相対指標なので、母の年齢別出生数のデータが必要ないという利点があり、出生力転換をデータで示すときに広く使われている。

記号	説明
W <sub>x</sub>	当該人口集団の x 歳女子人口 (x in 再生産期間)
W <sub>x</sub> [m]	当該人口集団の x 歳有配偶女子人口
W <sub>x</sub> [um]	当該人口集団の x 歳無配偶女子人口
B	当該人口集団の年間の総出生数
B[m]	当該人口集団の年間の嫡出出生数
B[um]	当該人口集団の年間の非嫡出出生数
f <sub>x</sub> [H]	Hutterite 集団（人類集団がこれまでに経験した中で最も高い出生力をもつ自然出生集団）の年齢別出生率

として、

- 総合出生力指標 (If) =  $B / \sum (W_x * f_x[H])$
- 嫡出出生力指標 (I<sub>g</sub>) =  $B[m] / \sum (W_x[m] * f_x[H])$  (有配偶出生力指標)
- 非嫡出出生力指標 (I<sub>h</sub>) =  $B[um] / \sum (W_x[um] * f_x[H])$  (無配偶出生力指標)
- 結婚指標 (I<sub>m</sub>) =  $\sum (W_x[m] * f_x[H]) / \sum (W_x * f_x[H])$  (I<sub>m</sub> = If/I<sub>g</sub> \* B[m]/B)

と定義すると、 $If = I_g I_m + I_h(1 - I_m)$  が成立する。それゆえ、出生力を結婚と結婚出生力に分けて考えることができる利点がある。つまり、晩婚なら I<sub>m</sub> が低くなり、早婚なら I<sub>m</sub> が高くなるので、晩婚化が出生力低下の主な要因であるならば I<sub>g</sub> はそれほど低下せずに I<sub>m</sub> が小さくなることによって If が小さくなるという現象が起きる。19 世紀末から 20 世紀にかけてのヨーロッパ諸国の I<sub>g</sub> の変化を見ると、半減以上に減っていることがわかる。

出生率の低下については、社会・経済分野でも、いろいろな仮説が提唱されてきたが、

結論はついていない。以下、主な仮説を紹介しておく。なお、途上国での 20 世紀後半については（実は日本の戦後 20 年間の急速な TFR 低下についても）、政府の介入の役割も無視できない。家族計画キャンペーンを行って避妊実行率を上げてきたことは、間違いなく出生率低下に大きく寄与したし、「子どもは授かるもの」という考え方から、「子どもを何人もつか」という能動的でコントロール可能なものという考え方に意識が変わったことも出生率低下への影響が大きかったことは自明であるが、十分に理論化されていない。

**シカゴスクール（Becker ら）の理論** 純粹経済学的なアプローチである。大前提として、転換前の社会も転換後の社会も合理的な行動をとっていたと見なす。さらに、（ケインズ流の）近代経済学の金科玉条である「嗜好一定の公準」を認める。

すると、産業革命によって所得水準があがっても、子どもの養育にかかる相対的成本が上昇したことによって、子どもの質を同等に保つために子どもの数に対する需要が減少したと考えることができる。これがシカゴ派の基本発想である。Leibenstein によって、「理論の本質が計量経済学の便宜のためにねじ曲げられている」と批判されている。1973 年の Michael からは供給要因にも触れるようになった。しかし子どもの供給側面は、受精能力、自然出生力、避妊、授乳、禁欲、その他の生理的、生物学的要素、及び出生力抑制を巡る動機づけや慣行など社会文化的要素が多く、経済的要素は少ないので、シカゴ派の純粹経済学的フレームでは扱うことが難しい（大淵, 1988）。

シカゴスクールの理論は、2つのタイプに分けられる。一つは子どもの量と質の相互作用に注目するミクロ経済学的接近であり、もう一つは出生力行動を主に女子の就業行動と結びつけた新家政学的接近である。現在では後者が主流である。「子どもの質」仮説を実証するために、Becker は子ども 1 人あたりの支出を子どもの質の代理変数と捉え、質に対する需要の所得弾力性が量のそれより大きいことを論証しようとしたが失敗した。新家政学的接近の嚆矢は Butz-Ward モデルである。これは出生力行動と労働市場を直接結びつけて所得変化に対する出生力の反応を見るもので、米国、英国、戦後日本のデータをよく説明できることが示されている。

しかし、「嗜好一定」には批判が多い。Pollak と Watkins は、「文化は特定の社会で個人が選択する行動のタイプを規定するから、ヒトの『合理的な振る舞い』自体が適切な範囲も、文化によって規定されるかもしれない」ことから、合理的な振る舞いが不変に保たれるという嗜好一定の公準は無意味であると論じている。

**女性の教育と地位の向上による意図的な出生抑制論** 文化によって嗜好が変わるという論としては、女性の教育の影響を論じた Cleland と Wilson の研究がある。Cleland は、各国の個人データに基づいて、既婚女子の出生児数の都市・農村間格差、教育程度別格差をみると、全般的に都市居住者は農村居住者より、教育年数の長い者は

短い者より出生児数が少ない傾向があることから、近代化が女性の教育と地位の向上を通じて出生力を低下させたと論じている。メカニズムとしては避妊を考えているので、これは reproductive rights の考え方に通じるものと思われる。出生力転換においては、避妊しても良いという新しい考え方の文化伝播が女性の教育によってもたらされたことを重視する。転換のタイミングが言語・文化的な境界に強く影響されていることが論拠である。

**Easterlin の社会経済モデル** しかし、シカゴスクールの主張ともっとも激しく争ったのは Easterlin である。彼は、嗜好一定の公準を否定し、出産コントロールのコスト、及び子どもの需要と供給への近代化の間接的影響を論じた。

当初 (1966-1975) は、子どもの相対価格こそが不変であると仮定し、所得が増えても子どもの数が減るとしたら、それは親の物質的生活水準に対する願望が上昇するためであるとした。1976 年から子ども一人あたりの希望遺産額という概念を導入した点で子どもにかかるコストの変化も取り入れているが、相対価格不変ということと、供給因子をフレームに入れている点がユニークである。

イースタリンの説は、二つの仮説に分けられる。相対所得仮説 (イースタリン仮説) と、需要供給理論である。需要供給理論では、前近代社会において出生力の超過需要状態があり、したがって自然出生力が現実に発現していたが、近代化によって子どもに対する需要 (Cd) が減退し、同時に潜在供給 (Cn) が高まり、調整費用 (RC ; 抑制の心理的不快感と技術修得のための費用や心理的負担) が減るために、超過供給状態への転換が起こって家族制限への動機付けが生じたと説明される。近代化の初期には調整費用が大きいので抑制行動が具体化せず、しばらくは自然出生力が継続し、近代化による自然出生力の上昇によって出生力上昇も見られるが、近代化の進行とともに意図的な出産制限が起こり、超過供給状態が調整されて、現実の生存児数が需要と一致する点まで (Bumpass と Westoff の「完全避妊社会」) 出生力が低下する、と考える。

途上国の場合は技術移転によって初めから調整費用が小さいので、出生力低下はより急激に起こっても不思議はないことになり、現実と合っている。

**Leibenstein の社会的相対所得仮説** Easterlin と同じく嗜好は一定でないとするものとして、Leibenstein の仮説がある。社会経済的地位、階級、集団による出生力水準の違い、つまり差別出生力の原因を、社会的地位に求めるものである。子どもを通常の耐久消費財とみなさず、拘束財と見なす点がユニークである。拘束財は、それへの家計支出がひとたびかかわりができると相当期間継続する点と、限界効用がある程度まで逡増的である点が通常の財と異なる。伝統的なミクロ経済学ではこの種の財は認めていない。

この論の根拠は、(1) 人間は特定の目標に向かって進む種であり消費水準の達成は

多くの個人にとって重要な目標、(2) 人口はいくつかの社会文化的な集団に分割され、各家計はそのいずれかに所属する。その集団は社会的影響力集団 (SIG) と呼ばれ、各家計の目標生活水準や家族規模の選好に影響する。SIG は社会経済的地位の集団と同じではないが単純化のため同一視してもよい、(3) SIG には構成員の共通目標があり、これによってある商品—健康、保険、教育、住宅など—が目標財となる。目標水準に達するまではその限界効用は逡増的でありうる。集団は階層的で下層ほど目標を達成しにくい、(4) 目標を達することは家計にとって一種の債務である。この債務には臨界的最小値があり、その水準までは限界効用の逡増が続く、(5) 家計と個人は支出の一部によってその地位感覚を表現する、(6) 家計所得の標準は SIG によって異なり、各 SIG は主だった財について目標消費水準を決定する (地位財)、(7) 同一 SIG 内の競争は目標債務への要求を増大させ、非債務所得を圧迫する。その傾向は所得上昇によっていっそう強められる、である。

このフレームでは、支出の逡増的な限界効用が大きくなれば、高所得集団に属する家計の子ども数が低所得集団に属する家計の子ども数より小さくなる場合があることは明らかである。しかし、効用関数や需要関数が定式化されていないので、まったく実証を試みられていない (大淵, 1988)。

**Mosk の家父長制仮説** 一方、社会体制の変化に出生力転換の原因を求める別の仮説として Mosk の論がある。これは、19~20 世紀のユーラシアでの出生力低下は、家父長制のもとにあった「前近代的人口体制」から「近代的人口体制」への移行によると論じるものである。有配偶出生力指標 (阿藤訳では「嫡出出生力指標」)  $I_g$  が高く、死亡力が低いところでは、経験的に  $I_m$  が小さくなる (晩婚になるか生涯未婚率が上がる) 傾向があることに着目したモデルである。前近代的人口体制における家父長制のもとでは、小農の家計は、一定の土地を所有し、生産と消費について意思決定を行うが、決定主体は家長であり、家長は自らの効用 (所得) 極大化と家系の継承 (家族の存続) を目的として行動する。家族内部では家長が子どもたちの結婚と職業選択について決定権をもち、遺産と結婚持参金を統制手段として子どもの労働用役を獲得する。子どもは家を離れば家長支配を免れるが、同時に一切の権利を失う。家長は息子の 1 人によって相続され、その時点で他の子どもたちは親元を離れる。娘は稼ぎのある男性と早く結婚して所得の極大化をはかり、老後の保障を得ようとする。

こうした社会で、家長がとれる人口行動は、家計内になるべく多くの男子労働力を確保することを目指すことになる。妻の経済的価値は若いほど不利であるが、受胎能力は適齢期を境に低下するので、遅くなりすぎない限り高い年齢の妻を得て、自然出生力戦略を用いて比較的短い出産間隔で出産を続けることになる。家長自身の結婚年齢は、相続が家長の死亡か引退時にしか起こらないことから晩婚が合理的で

ある。なお、死亡率と疾病罹患率は共同体の健康状態に依存する。

これを集計的なモデルに拡張すると、所得が高くなると健康状態がよくなることを通して自然出生力が上昇し、同時に死亡率が低下するので、女子の平均結婚年齢が上昇することになる。自然出生力があがることは現実の出生力に正の効果をもつが、結婚年齢上昇は負の効果をもつ。それがバランスしていたと考えれる。

近代化が起こると、子どもに対する経済的支配を保つためにかかるコストが増大し、さらに支配から生ずる労働力という便益が減少するので、家族成員間の力関係が変わって家長の権力が衰える。子どもに対する支配の純利益が減少するので、自然出生力戦略をとることが必ずしも合理的でなくなり、出生児数に応じた出産抑制をするようになる、ということで、出生力転換が説明される。

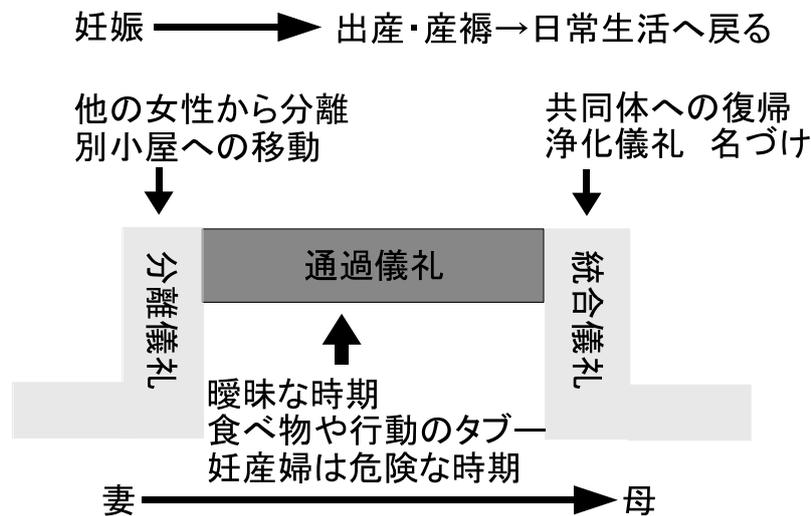
**西洋化原因論** これまでの議論では、近代化に伴う人口動態の変化としてきたが、原因が近代化というよりもむしろ西洋化 (Westernization) と考えるべきだ、というのが Caldwell の意見である。つまり、Westernization が一人の子どもへの投資の集中を伴う核家族化をもたらし、「世代間の富の流れ」が「子どもから親への流れ」から「親から子どもへの流れ」に変わって、それが出生率低下の原因であるとするのである。流れの逆転だけなら、子どもが生産財から消費財に変わったというのと本質的な差はないが、それが Westernization がもたらした社会文化的な諸変化によると考える点がユニークである。

ここでいう Westernization は、進歩、世俗化、大規模教育、環境へのヒトの優位といった概念を含むものである。これは経済的発達に先行し、Bangladesh やアフリカ南部のようなところでも出生力転換が起こったことを説明するのに有効であった。Cain (1982) が拡大家族では老人が子どもの代わりに働けるから、老人が働けるように投資する選択もありうるはずで、核家族化が primary というアイディアには反対しているし、データで実証されていないが (Dow et al., 1994 のケニアのデータで実証しようとする試みは失敗している)、魅力的な仮説ではある。

### 5.1.11 文化的側面

伝統社会では、出産とは、共同体に新しい成員が加わるということであり、一大セレモニーとして村を挙げて祝われるのが普通であった (共同体にとっては、出産そのものよりも、それを経て新しいヒトが共同体に入ってくる方が一大事であったと思われ、そういう共同体では、その意味では出産以上に名付けの式・成人式の方が重視されるであろう)。伝統に近代が入り込んでくるにつれて、こうした伝統は薄れつつある。

## 妊娠・出産の文化的意味



出典: 松岡悦子「出産の文化人類学: 儀礼と産婆[増補改訂版]」海鳴社, 1991年

女性のライフコースにとっては、生命を落とす危険もあり、大イベントであったには違いない。南太平洋の島の村では産後長期間の授乳をするため、産後無月経期間が長くなって、出産間隔は数年あるのが普通だった。出産に伴ってさまざまな文化的なしきたりが行われるのだが、これは、松岡(1991)によれば、一種の通過儀礼と捉えることができる。

現代の日本を初めとする先進諸国では、出産に伴う意識が変化してきている: eg. 柏木恵子「子どもという価値」中公新書:「さずかる→つくる」という意識の変化

## 5.2 出生の近成要因 (proximate determinants)

出生には、上記のようにいろいろな側面があるので、数値として現れる出生力を分析するためには、なんらかの形で整理する必要がある。考えやすい整理の仕方として、出産に直接かかわる要因(近成要因: proximate determinants)を、それ以外の間接的な要因と分けて考えることが Bongaarts という人口学者によって提案された。

- 基本要因
  1. 既婚者の割合
  2. 有配偶出生力の計画的調節要因

3. 避妊
4. 人工妊娠中絶
- 自然出生力要因
  1. 授乳性不妊（授乳による無月経）
  2. 性交頻度
  3. 不妊
  4. 自然胎児死亡
  5. 受胎可能な期間

Wood (1994) によって、上記のうち自然出生力要因についてだけ詳しく再分類したのもも発表されている。

- 曝露要因
  1. 結婚年齢
  2. 初経年齢
  3. 閉経年齢
  4. (もし閉経より早いなら) 病理的不妊の開始年齢
- 感受性要因
  1. 授乳性受胎不能継続期間
  2. (次の受胎確率要因によって規定される) 受胎可能な受胎待ち時間
  3. 性交頻度
  4. 月経周期の長さ
  5. 有排卵性月経割合
  6. 排卵当たりの受胎可能期間
  7. 受胎可能期間中の1度の性交による受胎確率
  8. 胎児損失確率
  9. 早期胎児死亡に伴う非感受性期間の長さ
  10. 生産にいたる妊娠の長さ

この程度まで詳しくすると、生存時間解析のフレームを使って（ハザード解析）、出産を数理モデル化することもできる。

## 5.3 出生の数学モデル

### 5.3.1 年齢別出生力を表すモデル

Coale & Trussell (1978) の  $M$  と  $m$

$$F(x) = G(x)F_m(x), \quad F_m(x) = Mn(x) \exp(mv(x)) \quad [1]$$

[1] 式は、よく知られている Coale and Trussell (1978) の年齢別有配偶出生力のモデルである。 $x$  が年齢、 $F(x)$  が年齢別出生力、 $G(x)$  が既婚割合、 $F_m(x)$  が年齢別有配偶出生力、 $n(x)$  が自然出生力、 $v(x)$  が年齢に依存する自然出生力からの下向きの乖離パターン ( $n(x)$  と  $v(x)$  は Coale と Trussell の論文に表の形で与えられている)、 $M$  は全体としての出生力水準を表す「スケール」パラメータであり、 $m$  は意図的な出産抑制の程度を示す「コントロール度」パラメータである。このモデルは、多くの集団について年齢別有配偶出生力パターンによく当てはまり、かつパラメータがアприオリな意味をもっている点が優れている。

ただし、ここで注意すべきは、前半の  $F(x) = G(x)F_m(x)$  の部分である。これが成り立つためには、非嫡出生が無視できるくらい小さい必要がある（式が厳密に成り立つためには、 $F_m(x)$  の分子に非嫡出生も含まれなくてはならないが、それは論理的におかしい）。

現代日本のデータに Nelder and Mead の方法で当てはめてみると、

年	$M$	$m$
1950	0.792655	0.675629
1955	0.730726	1.196406
1960	0.725194	1.467823
1965	0.711446	1.340176
1970	0.621124	1.113663
1975	0.706662	1.512317
1980	0.753364	1.612551
1985	0.845623	1.69222
1990	0.880761	1.795159
1995	1.067586	2.430152

となる。柏崎 (1990) によれば、いくつかの自然出生力集団について求められている  $M$  と  $m$  は、下表の通りである。

集団と年代	完結出生*	M	m
ハテライト**1921-	10.9	1.148	-0.082
カナダ 1700-	10.8	1.103	-0.152
ハテライト 1921 以前	9.8	1.016	-0.125
ジュネーブ***	9.4	1.186	0.236
イラン農村 1940-50	7.5	0.867	0.079
台湾 1900 年頃	7.0	0.805	0.071
ギニア 1954-65	6.2	0.824	0.390

\* 完結出生の数字は、20歳で結婚した婦人の生涯平均出産児数（完結出生力）

\*\* ハテライトは北米大陸に居住するメノナイト宗派の一分派で伝統的な農業と自然な生活を主義とする。教義によりできるだけ子供をたくさん産むことが是とされている。

\*\*\* ジュネーブの中産階級 1600-49

#### Hadwiger のモデル

$$F(x) = \frac{ab}{c} \left(\frac{c}{x}\right)^{3/2} \exp\left\{-b^2\left(\frac{c}{x} + \frac{x}{c} - 2\right)\right\} \quad [2]$$

[2] 式が Hadwiger のモデルである（原著は 1940 年にドイツ語で書かれているが、[2] 式は Chandra (1999) からの引用である）。年齢別出生力  $F(x)$  を 3つのパラメータ  $a, b, c$  で表す、比較的シンプルな関数である。3つのパラメータには、もともとは人口学的な意味はないが、Chandra ら (1999) は、ヨーロッパのデータについて、 $a$  が TFR とよく相関し、 $b$  が年齢別出生力のピークの高さと相関し、 $c$  が平均出産年齢と相関することを報告している。

戦後日本の有配偶出生パターンに当てはめたときは、

年	a	b	c
1950	4.073464	2.136599	25.412305
1955	3.125217	2.475954	23.72517
1960	2.839094	2.752689	23.237131
1965	2.768205	3.057415	23.597731
1970	2.486332	3.438803	24.141207
1975	2.620117	3.209107	23.203642
1980	2.785815	3.034859	22.987005
1985	3.19451	2.718303	22.625986
1990	3.406838	2.484866	22.191829
1995	5.578103	1.686612	18.863712

Hadwiger のモデルは、有配偶出生だけでなく、配偶関係を無視して全体の年齢別

出生率に当てはめることができる。戦後日本のデータに当てはめてみると、下表のパラメータが得られる。

年	a	b	c
1950	2.095265	3.228236	29.12809
1955	1.332294	3.832099	27.971617
1960	1.114097	4.687673	26.856602
1965	1.188757	4.988628	26.911619
1970	1.174453	5.127005	26.92299
1975	1.04915	5.190715	26.624719
1980	0.963726	5.510413	27.079794
1985	0.97709	5.296869	27.803982
1990	0.849243	5.012705	28.653945
1995	0.780356	4.709112	29.265282

ゴンペルツ関数を使ったリレーショナルモデル It resembles to Brass's logit model as relational model life table because of using transformation for linear regression by standard schedule. This model uses Gompit transformation for the cumulative proportions of age-specific fertility rates among TFR. Let  $p = \text{ASFR}/\text{TFR}$ ,  $\text{Gompit}(p) = -\ln(-\ln p)$ .  $p = \exp(-\exp(-\text{Gompit}))$ .

Brass's fertility polynomial  $f(x) = c(x-s)(s+13.2-x)^2$

$x$ : age,  $f(x)$ : cumulative fertility,  $c$ : level parameter,  $s$ : age at which fertility begins,  $s+13.2$  is the mean age of childbearing.

他 対数正規分布を使ったモデルなども提案されている。

### 5.3.2 ハザードモデル

個人ベースの出産暦データがあれば、生存時間解析の手法を用いて、切れ味よく出生力を分析することができる。たとえば第1子と第2子の出産間隔のデータを使えば、比較的短期間の変化をコホートレベルで捉えることが可能になる。

ハザード解析の基本モデル

## 用語

- 持続期間変数  $T$  により人口学イベント（死亡、出生）のタイミングを表現。
- $t$  を  $T$  の特定の値を示すために用いる。これらの値は、厳密に非負。
- ハザード  $h(t)$  : あるイベントが、それ以前に起こらないという条件付きで、時点  $t$  で発生する瞬間の率
- 確率密度関数  $f(t)$  : イベントの発生までの時間の分布
- 生存関数  $S(t)$  : 時点  $t$  までにあるイベントが「まだ」起こっていない確率
- 余命  $e(t)$  :  $t$  におけるイベントまでの残存時間の期待値

$$h(t) = \lim_{\Delta t \rightarrow 0} \left[ \frac{\text{Prob}(t \leq T < t + \Delta t \mid t \leq T)}{\Delta t} \right]. \quad [3]$$

確率への変換形

$$q(t) = 1 - \exp \left[ - \int_{t+1/2\delta t}^{t+1/2\delta t} h(y) dy \right]. \quad [4]$$

$$f(t) = \lim_{\Delta t \rightarrow 0} \left[ \frac{\text{Prob}(t \leq T < t + \Delta t)}{\Delta t} \right]. \quad [5]$$

$$S(t) = \text{Prob}(T > t) = \int_t^{\infty} f(y) dy. \quad [6]$$

$$e(t) = E(T - t \mid T > t) = \int_t^{\infty} S(y) dy / S(t). \quad [7]$$

ハザード、確率密度関数、生存関数の関係式

$$f(t) = -dS(t)/dt, \quad [8]$$

$$h(t) = -d[\ln S(t)]/dt = f(t)/S(t), \quad [9]$$

$$S(t) = \exp \left[ - \int_0^t h(y) dy \right], \quad [10]$$

$$f(t) = h(t) \exp \left[ - \int_0^t h(y) dy \right]. \quad [11]$$

ハザード解析の手順

1.  $h(t)$ 、 $f(t)$ 、 $S(t)$  のどれか一つについて数学的形式（例えば、指数、ワイブル、ゴンペルツ、対数正規、二次、対数ロジスティック、ガンマなどの分布）を特定する。
2. 最尤推定によってパラメータを決める。
3. 関係式を用いて残りを計算する。

4. 上の2つのプロセスとも解析的には解けないことが多いが、コンピュータにより数値解は得られる。

右側打ち切り標本（ただしランダム打ち切りの場合）を含む場合の最尤推定は以下の通り。

1.  $t_i$  を  $i$  番目のケースについてのイベントが最初に起こるまでか打ち切りまでの時間とし、 $d_i$  をそれに関連した「打ち切りインデックス」（observation が打ち切りだったら 0、打ち切りでなければ 1）とすると、 $n$  個のデータについての尤度関数は、

$$L = \prod_{i=1}^n [f(t_i)]^{d_i} [S(t_i)]^{1-d_i}, \quad [12]$$

式 [9] より  $f(t) = h(t)S(t)$  なので、

$$L = \prod_{i=1}^n [h(t_i)]^{d_i} \cdot \exp\left[-\int_0^{t_i} h(y)dy\right], \quad [13]$$

2. パラメータの最尤推定には、単純に  $f(t)$  と  $S(t)$  についてのモデルから導かれた表現を [12] 式に、あるいは  $h(t)$  についてのそれを [13] 式に代入して、ニュートン・ラフソン法や滑降シプレックス法のような、標準的コンピュータアルゴリズムを用いて数値的に  $L$  を最大化する。
3. モデルの標準誤差を求める。 $\theta = (\theta_1, \theta_2, \dots, \theta_k)$  が、我々があてはめたいモデルの  $k$  個のパラメータからなる一つのベクトルとすれば、情報行列は次の式で与えられる。

$$\mathbf{I} = \begin{bmatrix} \mathbf{E}(-\partial^2 \ln L / \partial \theta_1^2) & \dots & \mathbf{E}(-\partial^2 \ln L / \partial \theta_1 \theta_k) \\ \vdots & \cdot & \vdots \\ \mathbf{E}(-\partial^2 \ln L / \partial \theta_k \theta_1) & \dots & \mathbf{E}(-\partial^2 \ln L / \partial \theta_k^2) \end{bmatrix}, \quad [14]$$

漸近的分散共分散行列は、

$$\hat{\mathbf{V}} = \mathbf{I}^{-1} |_{\theta=\hat{\theta}}, \quad [15]$$

ここで、 $\hat{\theta}$  は、モデルパラメータの最尤推定量のベクトルである。 $\hat{\mathbf{V}}$  の主対角にそっての平方根が、推定されたパラメータの標準誤差を与える。

#### モデルの適合度評価法

1. きわめて単純な、直感に訴えるグラフィカルな方法（カプラン・マイヤ推定量との比較）
2. まったく直感的ではないが最終的には情報が多い、複雑なグラフィカルな方法（マルチンゲール残差のプロット）
3. (一つの方法のパラメータがもう一方のサブセットであるときに限って) 尤度比検定

### Kaplan-Meier 推定

観察された、打ち切りでない待ち時間を、 $t_1 < t_2 < \dots < t_k$  という離散量だとしよう。 $w_j$  を時刻  $t_j$  における打ち切りでないイベントで終了する待ち時間、 $c_j$  を時刻  $[t_j, t_{j+1})$  ( $j = 1, \dots, k$ ) の間に打ち切りで終わる待ち時間とする。このとき、時刻  $t_j$  におけるまだイベントを経験せずに残っている対象者のプールは、

$$R_j = \sum_{i=j}^k (w_i + c_i). \quad [16]$$

$S(t)$  の Kaplan-Meier 推定量は、

$$\hat{S}(t) = \prod_{j|t_j < t} [(R_j - w_j)/R_j], \quad [17]$$

その漸近分散は、

$$\hat{\text{var}}[\hat{S}(t)] = [\hat{S}(t)]^2 \sum_{j|t_j < t} \{w_j/[R_j(R_j - w_j)]\}. \quad [18]$$

### 出生のモデル

Old Order Amish のデータによる受胎能力 (= fecundability ; 一組のカップルが、パートナー両方とも生物学的に受胎能力があるとして、避妊していない性交を続ける 1 ヶ月の間に受胎する確率) の推定

- データ選択の基準
  1. 避妊がない
  2. 病理的な不妊が広まっていない
  3. 結婚前の受胎がほとんどない
- 病因論モデル
  1. 均質モデル : 標本中の全てのカップルが正確に同じ受胎能力をもっていて、この値は時間が経っても変化しない (毎月の受胎ハザード自身は定数であり、 $\lambda$  とする)

$$h(t) = \lambda, \lambda > 0, \quad [19]$$

$$f(t) = \lambda e^{-\lambda t}, \quad [20]$$

$$S(t) = e^{-\lambda t}. \quad [21]$$

$\lambda$  の最尤推定量 :  $\hat{\lambda} = 0.067 \pm 0.001$ 、最初の有効受胎までの期待待ち時間として  $1/\hat{\lambda} = 14.97 \pm 0.26$  ヶ月。式 [4] からこのモデルが与える有効受胎能力の推定値は 0.065。大抵の自然出生力集団では、有効受胎能力の推定量は 0.15 から 0.30 なので、異常に低い。

2. 不均質モデル：集団は2つの均質な部分集団に分けられ、その一方は一定のゼロでない受胎能力をもち、もう一方は完全不妊（すなわち受胎能力がゼロ）であるとすると、

$$\bar{h}(t) = \rho(t)h_1(t) + [1 - \rho(t)]h_2(t)$$

ここで、 $h_1(t)$  と  $h_2(t)$  は、二つの部分集団それぞれのハザードであり、 $\rho(t)$  は時点  $t$  において第1の部分集団に属しているカップルの割合である。第1の部分集団でのハザードを定数  $\lambda$  とおくと、

$$\bar{S}(t) = \rho(0)S_1(t) + [1 - \rho(0)]S_2(t) = 1 - \rho(0)(1 - e^{-\lambda t}), \quad [22]$$

$$\bar{f}(t) = \rho(0)\lambda e^{-\lambda t}. \quad [23]$$

こちらの方がまともな推定値が得られる。

## 5.4 出生調査の方法論

### 5.4.1 断面研究の場合

- 計数データからの推計法とその指標：CBR, ASFR, TFR, ASMFR, TMFR, GRR, NRR, MISG, CWR, RBM

CBR (crude birth rate: 普通出生率, 粗出生率<sup>\*1</sup>) = 年間出生数 / 年央人口 × 1000

ASFR (age specific fertility rate: 年齢別出生率<sup>\*2</sup>) = 年齢別出生数 / 年齢別女性人口 × 1000

TFR (total fertility rate: 合計出生率<sup>\*3</sup>) = 全年齢(階級)について ASFR を合計したもの

ASMFR (age specific marital fertility rate: 年齢別有配偶出生率) = 年齢別嫡出生数 / 年齢別有配偶女性人口 × 1000

TMFR (total marital fertility rate: 有配偶合計出生率) = 全年齢(階級)について ASMFR を合計したもの

GRR (gross reproductive rate: 総再生産率) = ある仮設女子出生コホートについて、再生産完了まで死亡がゼロであるという仮定の下で、その ASFR が現在のものに従った場合の平均女兒数 (実は TFR に女兒出生性比をかけたものと同値である)

\*1 複数年について計算し、平滑化する場合もある

\*2 年齢別というより、年齢5歳階級別で計算するのが普通である

\*3 分母が女性人口であるために、「合計特殊出生率」と呼ばれてきたが、日本人口学会の用語辞典において「合計出生率」でよいとされている。

- NRR** (net reproductive rate: 純再生産率) = 死亡率を考慮した **GRR**。つまり、ある仮設女子出生コホートが現行の **ASFR** と年齢別死亡率に従う場合の、母親がその娘を産んだ年齢まで生存する平均女兒数 (計算には生命表の  $LF(x)$  を用いる)
- MISG** (mean interval between successive generations: 平均世代間隔) = 女子の平均世代間隔は、現行の **ASFR** と年齢別死亡率の下で娘を産んだ母親の平均年齢に等しい。男子も同様。
- CWR** (child woman ratio: 婦人子ども比率) = 5歳未満人口 / 再生産年齢 (通常 15-49) 女子人口 × 1000
- RBM** (ratios of births to marriages: 出生結婚比) = ある年の出生数 / その年の結婚数 (あるいはその年を含むそれ以前の結婚数の加重平均)
- SBR** (standardized birth rate: 標準化出生率) = 標準化死亡率の場合と同じく、母親の年齢別出生率を基準集団の女性の人口構造で重み付けした値。ただし、死亡率の場合ほど使われない。
- 遡及聞き取りデータからの推計法とその指標: **MCP**, **PPR**, **PD**, **BI**, **ABLC**, **PSBP**, **TLFR**, **DMR**
- MCP** (mean completed parity: 平均完結パリティ) = 再生産を完了した女性の既往出生児数の平均値
- PPR** (parity progression ratios: パリティ拡大率) =  $n+1$  人の子どもをもつ女子人口 /  $n$  人の子どもをもつ女子人口
- PD** (parity distribution: 子ども数分布) = 既往出生児数のヒストグラム (ポアソン分布 / 負の二項分布)
- BI** (birth interval: 出生間隔) = 結婚と第1子出生の間隔 / 各出生間隔 / 平均出生間隔
- ABLC** (age at the birth of the last child: 末子出産年齢) = 再生産を完了した女性が末子を出産した年齢
- PSBP** (parity-specific birth probabilities: パリティ別出生確率) = ある期間に起こった第  $x+1$  子出生数 / その期間の期首におけるパリティ  $x$  の女子人口
- TLFR** (total legitimate fertility rate: 合計嫡出出生率) = 結婚持続期間別出生率の合計値
- DMR** (daughter mother ratio: 女兒 / 母親比) = 末子が再生産年齢を超えた母親一人当たりの結婚した娘の数

### 5.4.2 フォローアップ研究の場合

基本的には遡及聞き取りデータの場合と同等である。

**WTFR** (wanted total fertility rate: 希望合計出生力) = ASFR のうち、これ以上子どもは  
いらなと思っていない場合を合計したもの。cf. DFS (desired family size: 希望子  
供数) これについてはバイアスの問題や聞き取りの難しさをめぐってさまざまな議  
論があったが、一応これがマシということになっている。

**Unwanted birth** これ以上子どもはいらなという時点で到達して以降の出生。  
wanted birth と排反。→ wanted fertility は直接指標ではなく、unwanted でない  
ものとして計算される。(cf) intended number of children / ideal family size

**Desired family size** 個人またはカップルが、自分たちにとって持ちたいと思う子  
ども数。WFS では、既婚女性に「もし生涯に生む子どもの数を正確に選べる  
としたら、何人ですか」と訊いた回答。バイアス、攪乱要因多し。

**Desired TFR** Westoff et al. が提案。desired family size を超えた出生を計算から  
除いた TFR。意図せぬ出生制限、非数値回答、子どもの死亡、タイミングの変  
化という攪乱要因は除けるがまだ不完全。

**Wanted status of recent births** WFS では、最近 5 年間の各出生について、それが  
望んだものだったか否かを尋ねる→ Reported wanted TFR は TFR の計算から、  
この質問で not wanted と回答されたものを除いて行えば得られるが、2つの  
バイアス（社会規範などにより大きめに申告する、質問項目の違いの影響）が  
ある。

**Wanted TFR** want-more にバイアスが少ないことに着目して Bongaarts が改良し  
た値

**累積出生率** = 年齢別出生率あるいは結婚持続期間別出生率を、コホートの出生リスクの  
開始年齢からある一定年齢まで積み上げたもの

**完結出生力** = コホートの全成員が再生産を完了したときの累積出生率

**累積純出生率 (cumulative net fertility)** = コホートの年齢別出生率とその年齢までの女子  
人口の生存確率との積和

### 5.4.3 歴史人口研究の場合

長い時間軸での人口再構築特有の問題：資料からの脱落、資料の記録形式の変化

#### 5.4.4 二次資料を利用する場合

二次資料の利点と欠点：多数のデータ／資料の取り方の不均質性

### 5.5 出生分析の実際

#### シミュレーション研究の例

Nakazawa, M. and R. Ohtsuka (1997) Analysis of completed parity using microsimulation modeling. *Mathematical Population Studies*, 6: 173-186.

小集団での人口再生産には、個人間の関係が本質的である。ヒトが生まれてから死ぬまでに経験する人口動態イベントは、基本的には年齢依存する確率として決定することができるが、小集団の場合は、結婚する確率が如何に高くなっても、婚姻規制を満たす適当な相手がいなかったために結婚できないという事態が容易に起こり得る。

そのために、初期人口を決め、各個人に人口動態イベントが起こるかどうかを(0,1)の一様乱数が各々のイベントの生起確率より低いかどうかで決定し、一定の年数の間シミュレートしてみることによって、実際にさまざまなイベントが起こっていった場合に、その集団の人口がどのように変化してゆくかを観察する、マイクロシミュレーション分析が有効になる。

本研究では、Hill and Trussell (1977) の年齢別死亡確率モデル、対数正規分布による年齢別結婚確率モデル、Coale and Trussell (1974) の年齢別有配偶出生確率モデルを組み合わせて人口再生産のマイクロシミュレーションモデルを構築した。パラメータについては、パプアニューギニアの南側、フライ河の河口部に居住する人口約2,000人のギデラと呼ばれる一言語族のデータを用いた。

出生と死亡について多産多死、少産少死とその中間の3パターン、婚姻規制について一夫一妻、一夫多妻の2パターンの組み合わせで、初期人口1000人で、300年間のシミュレーションを、乱数の初期値のみ変えて各々100回ずつ行った結果、一夫多妻の方が夫婦の年齢差が大きくなり、同時に適当な相手が見つからないケースが減った。また、多産多死よりも少産少死の方がギデラの完結パリティや娘母親比からみた人口再生産パターンによく符合していた。

## 小集団の実証的研究の例

Nakazawa, M., A. Ishii and J. Leafasia (2000) Demographic effects of modernization in a small village of Solomon Islands. 人口学研究, 27: 7-13.

ソロモン諸島の首都ホニアラの東 50km に位置する村落の住民 187 人に対して、インタビューに基づいた人口調査を行った。この地域は 1980 年代半ばに橋と舗装道路の整備によってホニアラとの交通の便が良くなり、急激な近代化が起こったと考えられるため、近代化前後で人口学的特徴に違いが見られるかどうかを明らかにすることを目的とした。

調査対象地域の 1 人でも子どもをもつ女性全員に、本人の誕生日と出産暦を尋ねたが、年単位でしか特定できないデータについては区間打ち切りレコードとして扱い、不均質さを取り入れたハザード解析 (カプラン=マイヤ法) によって分析した。その結果、近代化前に比べ、出生力に上昇傾向があることが示唆された。この傾向が主に産後不妊期間の短縮によるとすれば、食生活の近代化が影響している可能性がある。

## マクロデータへのモデルの当てはめの例

中澤 港 (2003) 日本の年齢別出生力パタンのトレンドの分析 山口県立大学看護学部紀要, 7: 67-76.

近年の出生力の低下は日本においては大きな人口問題の一つである。多くの先行研究は、この低下が結婚タイミングの遅れあるいは非婚率の上昇に起因し、有配偶出生力自体は低下していないとしているが、年齢別有配偶出生力のパタンを数学的に詳細に検討したものは少ない。本研究は、戦後の日本における年齢別有配偶出生力パタンについて Coale and Trussell のモデルと Hadwiger のモデルを、年齢別出生力パタンについて Hadwiger のモデルを適用し、そのパラメータの経時的なトレンドを三次関数で近似したものである。Hadwiger のモデルに基づく将来予測によれば、日本の出生力のトレンドは、1950 年以降のデータをすべて使うか、1960 年以降のデータだけを使うかで結果が違っていた。1960 年以降のデータだけを使った場合、今後やや上向くと予想された。

## 要因を探る大規模なサンプル調査

国立社会保障・人口問題研究所 (2003) 第 11 回出生動向基本調査

(独身者調査 [<http://www.ipss.go.jp/Japanese/doukou11/P00-22.pdf>],

夫婦調査 [<http://www.ipss.go.jp/Japanese/doukou11/doukou11.pdf>]) .

出生の要因を探るために定期的に行われているサンプル調査である。かなり突っ込んだ質問項目もあって興味深い。web から全文入手できるので、関心がある方は是非ダウンロードして読みたい。

## 5.6 文献

### 5.6.1 出生全般

- 柏木恵子 2001 年「子どもという価値」(中公新書)
- 松岡悦子 1991 年「出産の文化人類学 儀礼と産婆 [増補改訂版]」(海鳴社)
- 榎本知郎 1999 年「繁殖と性行動」(西田利貞, 上原重男編『霊長類学を学ぶ人のために』世界思想社)
- Faletti, A.G., C.A. Mastronardi, A. Lomniczi, A. Seilicovich, M. Gimeno, S.M. McCann and V. Rettori, 1999, " $\beta$ -Endorphin blocks luteinizing hormone-releasing hormone release by inhibiting the nitricoxidergic pathway controlling its release", Proceedings of National Academy of Sciences, USA, Vol. 96, No. 2 (February)
- Hawkes, K., J.F. O'Connell, N.G. Blurton Jones, H. Alvarez and E.L. Charnov, 1998, "Grandmothering, menopause, and the evolution of human life histories", Proceedings of National Academy of Sciences, USA, Vol. 95, No.2 (February)
- Holman, D.J. 1996, Total Fecundability and Fetal Loss in Rural Bangladesh, Doctoral Dissertation, Pennsylvania State University.
- 市川眞澄, 岡 良隆, 小林牧人, 武内ゆかり, 東村博子, 西原真杉, 朴 民根, 前多敬一郎, 村上志津子, 森 裕司 1998 年『脳と生殖-GnRH 神経系の進化と適応』学会出版センター
- Nakazawa, M. and R. Ohtsuka, 1997, "Analysis of completed parity using microsimulation modeling", Mathematical Population Studies, Vol.6 No.3
- 田中富久子 1998 年『女の脳・男の脳』NHK ブックス
- 堤 治 1999 年『生殖医療のすべて』丸善ライブラリー
- Wedekind, C. and S. Furi, 1997, "Body odour preferences in men and women: do they aim for specific MHC combinations or simply heterozygosity?", Proceedings of Royal Society, London B: Biological Sciences, Vol.264, No. 1387
- Wood, James W., 1994, Dynamics of Human Reproduction: Biology, Biometry, Demography, New York, Aldine-de-Gruyter
- 長谷川真理子 1993 年『オスとメス=性の不思議』講談社現代新書
- 栗原堅三 1998 年『味と香りの話』岩波新書
- 岩本光雄 1977 年「現生霊長類の分類と分布」(伊谷純一郎 編『人類学講座第2巻・霊長類』雄山閣出版)
- 杉本正信, 古市泰宏 1998 年『老化と遺伝子』東京化学同人

- Napier, J.R. and P.H. Napier, 1985, *The Natural History of the Primates*, London, British Museum. (伊沢紘生訳 1987年『世界の霊長類』どうぶつ社)
- 林 壽郎 1968年『標準原色図鑑全集／別巻「動物II」』保育社
- Harvey, P.H., R.D. Martin and T.H. Clutton-Brock 1987, "Life Histories in Comparative Perspective" in B.B. Smuts, D.L. Cheney, R.M. Seyfaith, R.W. Wrangham and T.T. Struhsaker (eds.), *Primate Societies*, Chicago, Univ. Chicago Press
- Hawkes, K., J.F. O'Connell, N.G. Blurton Jones, H. Alvarez and E.L. Charnov, 1998, "Grandmothering, menopause, and the evolution of human life histories", *Proceedings of National Academy of Sciences, USA*, Vol. 95, No.2 (February)
- Adashi, Eli Y., J.A. Rock and Z. Rosenwaks (eds.), 1996, *Reproductive endocrinology, surgery, and technology*, Philadelphia, Lippincott-Raven Publishers.
- Dettwyler, K.A. 1995, "A Time to Wean: The Hominid Blueprint for the Natural Age of Weaning in Modern Human Populations" in P. Stuart-Macadam and K.A. Dettwyler (ed.), *Breastfeeding: Biocultural Perspectives*, New York, Aldine de Gruyter
- 三浦悌二 1983年『生まれ月の科学』篠原書店
- 蒲原聖可 1998年『ヒトはなぜ肥満になるのか』岩波書店
- Cunningham, M.J., D.K. Clifton and R.A. Steiner, 1999, "Leptin's actions on the reproductive axis: Perspectives and mechanisms", *Biology of Reproduction*, Vol. 60, No. 2
- Stern, K. and M.K. McClintock, 1998, "Regulation of ovulation by human pheromones", *Nature*, Vol. 392
- 山元大輔 2004年『男と女はなぜ惹きあうのか：「フェロモン」学入門』中公新書ラクレ

### 5.6.2 低出生力について

- 森 千里 (1998) 「ヒトの精子への影響」(シーア・コルボーン, 養老孟司, 高杉 暹, 田辺信介, 井口泰泉, 堀口敏宏, 森 千里, 香山不二雄, 椎葉茂樹, 戸高恵美子『よくわかる環境ホルモン学』環境新聞社)
- 森千里 (2002) 胎児の複合汚染, 中公新書
- 中澤 港 (1999) 「ヒトの精子数は“全世界で”減少しているか-減少地域で危険因子の追究を」*科学*, Vol. 69, No. 1
- Coale AJ, Watkins SC [Eds.] (1986) *The decline of fertility in Europe*, Princeton Univ. Press, New Jersey.

- Kirk D (1996) Demographic transition theory. *Population Studies*, 50: 361-387.
- Knodel JE (1977) Family limitation and fertility transition: Evidence from age patterns of fertility in Europe and Asia. *Population Studies*, 31(2)
- Borgerhoff-Mulder M (1998) The demographic transition: are we any closer to an evolutionary explanation? *Trends in Ecology and Evolution*, 13(7): 266-270.
- 大淵寛 (1988) 出生力の経済学, 中央大学出版部
- Kaplan H, Lancaster JB, Tucker WT, Anderson KG (2002) Evolutionary approach to below replacement fertility. *American Journal of Human Biology*, 14: 233-256.
- 柏崎浩 (1990) 「自然出生力と抑制された出生力」生物の科学・遺伝, 44(12): 20-24.
- Chandola T, Coleman DA, Hiorns RW (1999) Recent European fertility patterns: Fitting curves to 'distorted' distributions. *Population Studies*, **53**: 317-329.
- Coale AJ, Trussell TJ (1978) Technical Note: Finding the two parameters that specify a model schedule of marital fertility. *Population Index*, **44**: 203-213.
- 廣嶋清志 (2001) 出生率低下をどのようにとらえるか? 一年齢別有配偶出生率の問題性— 数理社会学会『理論と方法』, 16: 163-183.
- Hoem JM et al. (1981) Experiments in modelling recent Danish fertility curves. *Demography*, **18**: 231-244.
- Nelder JA, Mead R (1965) A simplex method for function minimization. *The Computer Journal*, **7**: 308-313.
- 赤川学 (2004) 子どもが減って何が悪いか!, ちくま新書

## 第 6 章

# 結婚と離婚

実際のデータの扱い。年次変化とか国際比較とか年齢パターンとか  
初婚年齢については，Coale and McNeil のモデルや，金子隆一さんが Coale and McNeil  
を一般化したモデルがある。離婚についてはモデル化が困難。

### 6.1 結婚のモデル

Comparing with mortality models, nuptiality and fertility models are more complex and less developed.

The complexity mostly comes from (1) fertility may differ by marital status, (2) marriage, divorce, and widowhood may occur more than once (death must occur only once). Thus the model fit the data less closely than mortality models.

#### 6.1.1 Coale-McNeil nuptiality model

$$g(a) = (0.174/k) \exp(-4.411 \exp(-(0.309/k)(a-a_0)))$$

a: age, g(a): the risk of first marriage at age a, a<sub>0</sub>: earliest age at which there are a significant number of first marriages, k: parameter for the shape of the curve. Ranges of a<sub>0</sub> are from 12 (past India) to 18 (developed countries), because it's minimum age of marriage. Ranges of k are from 0.2 (very rapid increase rate of marriage after a<sub>0</sub>) to 1.7 (very slow).

k=0.5 indicates a rate of marriage that is double that in the standard population (k=1.0). Modified version of Coale-McNeil model (1972) include the scale parameter C, the proportion of women who eventually marry within her whole life. The model is,

$$g(a) = 0.19465 C \exp(-0.174 ((a-a_0)/k - 6.06) - \exp(-0.288 ((a-a_0)/k - 6.06)))$$

In fmsb package, the function `CM(scale=0.8, a0=18, k=2)` gives proportion ever

married for ages 10 to 60.

`fitCM()` can automatically fit the model to actual data of proportion ever married.

See the R code given as <http://minato.sip21c.org/demography-special/Chap14.R>

## 第7章

# 人口移動と人口分布

人口地理学分野で研究が進んでいる。現代は人口移動の影響が大きく、東京都の TFR はどの道府県よりも飛び離れて低いにもかかわらず、東京だけが人口増加しているのは、ひとえに人口流入が大きいからである。ここではごく基本的なデータ分析と、分布を地図上に示す方法だけ概説する。



## 第 8 章

# 人口推計とシミュレーション

### 8.1 人口推計とは何か？

ここでいう人口推計は、人口の将来推計の意味である。不完全なデータの補正も人口統計の推計には違いないが、それは出生、死亡等の各章に譲る。将来推計の意味での人口推計には、大別すると2つのアプローチがある。1つは、人口変化のトレンドを外挿するやり方である。世界人口の推計には良く用いられる。もっとも、トレンドの外挿であるにしても、ロジスティック成長曲線のように、定式化自体は密度依存の死亡というメカニズムに基づいている場合もある。

もう1つは、出生、死亡、結婚、移動などの人口動態のメカニズムを考慮して、その差し引きとして人口を予測するアプローチ、即ちシミュレーションによる方法である（一般に、コウホート要因法と総称される）。後で触れるように、シミュレーションによる人口予測は強力なツールであるが、最大の問題は妥当性の保証がないことである。

#### 8.1.1 密度依存の死亡と出生

・密度がある程度低いうちは、死亡率は密度と独立。ある程度密度が上がると、密度とともに死亡率が上昇する。これを密度依存の死亡率と呼ぶ。しかし密度依存で死亡率が上昇しても暫くは出生数が死亡数を上回るので密度は上がり続ける。さらに密度が上がると死亡率の上昇が急になり、「過補償」状態になって密度は低下する。ちょうど補償するところで密度は定常になる。

・出生についても、密度がある程度低いうちは独立で、ある程度上がると出生率が下がり始め、ちょうど補償する点に収束する。

### 8.1.2 人口支持力 (carrying capacity)

・ちょうど補償する場合，密度は一定値  $K$  に収束する。この値を人口支持力または環境支持力と呼ぶ。

### 8.1.3 ロジスティック成長

・人口密度が十分に低いとき，個体群の人口増加速度は，現存量，すなわち現在人口に比例する。 $dN/dt=rN$  となるので，指数関数的な人口増加（マルサスの成長）が起こる。

・密度が高くなってきて人口支持力に近づくと，人口増加速度は，現存量だけでなく，人口支持力と現存量の差の人口支持力に対する割合にも比例するようになる。これによって人口増加は S 字状のロジスティック曲線になる。これをロジスティック成長という。

$$dN/dt=rN(K-N)/K$$

・上記は連続時間で考えていいときに成り立つが，離散時間で考えねばならない場合は，差分方程式になるので， $r$  の値によって大きく人口増加の様子は変わる。いわゆるカオスになる。

・さらに，密度効果がすぐに出生率や死亡率に影響せず時間遅れがある場合も，数理生物学の分野でいろいろ考えられている。基本的に減衰振動となる場合が多いように思う（ちょっと自信なし）。

## 8.2 コウホート要因法

コウホート要因法とは，Cannan(1895)が開発した人口予測の方法で，主要な人口学研究施設が出している予測はほぼすべてこの原理に基づいている。コウホート要因法では，出生や死亡の将来予測が必要となる。出生でも死亡でも，2時点でのデータからトレンドを計算する Lee-Carter 法で行われるのが普通である。R では `demography` パッケージの `lca()` 関数で Lee-Carter 法が実施可能。

- まず，最新の調査データとして，性・年齢階級（現在では 5 歳階級が多いが，Cannan は 10 歳階級）別人口を得る。以下，通常は女性人口についてのみ考える。
- 次に，最近 5 年間の年齢階級別の死亡率を求めて，年齢階級別生残率を得る。
- それが次の 5 年間にも維持されるとか 10 % 改善されるとか仮定すれば，5 歳未満人口を除く，5 年後の年齢階級別人口が推定できる。
- 5 年後の 5 歳未満人口は，再生産年齢（通常は 15 歳から 49 歳）の女性について，年齢階級別人口と年齢階級別有効出生率（生残する子どもについての出生率）を掛

ければ得られる（年齢階級別有効出生率は、一番単純には、5年前に各年齢階級別女性人口から生まれた現在の5歳未満女児人口を5年前の各年齢階級別女性人口で割れば推定できるが（生残率と同様、それが維持されると考えたり、一定割合で変化すると仮定したりする）、他の方法もある。）。

- 以上で5年後の年齢階級別女性人口が得られたので、後はそれを繰り返せばもっと先が推定できる。
- 男性人口については、女性人口を2倍する方法の他、5歳以上人口は女性人口と同様に年齢階級別生残率から求めて、5歳未満人口に関しては女性人口を1.05倍するという方法などがある。
- これに移住の影響を組み合わせて将来予測をする。

最大の問題は、妥当性のある出生率予測だと言われている。死亡率に比べ、出生率に影響する要因は不明な点が多く、社会文化的な複雑な要因の影響を受けると考えられるし、法制などによって急に変わることがあるためである。最近では避妊実行割合の予測と組み合わせたり、出生率や生残率の予測に確率的な変動を組み込んだりすることも行われている。

コウホート要因法は、過去は現在を通じてのみ未来に影響するという意味でマルコフ的である。よく使われているように予測の関数を線型にしまうと、数学的解析は容易だが、生態学的な制約が考慮されなくなってしまう。しかし、非線型の予測は難しく、今後の研究の進展が待たれる分野である。

移動を含まないコウホート要因法で日本の人口推計をするコードの例を以下に示す。fmsb パッケージに入っている日本の人口データの、Jfert と Jlife から出生と死亡に使う年齢パターンを選び、それが今後100年間変化しないとして、Jpopl の2010年人口から今後100年の人口推計をして1秒ごとに人口ピラミッドを書き換えていく。人口ピラミッドの描画のため、pyramid パッケージもインストールしておく必要がある。

```
library(fmsb)
library(pyramid)

wx <- length(Jpopl[, "Age"])
mp <- fp <- matrix(rep(0, wx*100), wx, 100)
mp[, 1] <- Jpopl[, "M2010"] # set 2010 males as initial males
fp[, 1] <- Jpopl[, "F2010"] # set 2010 females as initial females
# select lifetable for males
mqxn <- select.list(choices=
  names(Jlife)[grep("qx([0-9]+)M", names(Jlife))],
  preselect="qx2010M", title="Select the lifetable for males")
mmx <- qxtomx(Jlife[, mqxn])
mcx <- length(mmx[!is.na(mmx)])
if (mcx < wx-1) {
mmx <- c(mmx[!is.na(mmx)], rep(0, wx-mcx-1))
} else {
mmx <- mmx[1:(wx-1)]
}
# select lifetable for females
fqxn <- select.list(choices=
  names(Jlife)[grep("qx([0-9]+)F", names(Jlife))],
  preselect="qx2010F", title="Select the lifetable for females")
fmx <- qxtomx(Jlife[, fqxn])
fcx <- length(fmx[!is.na(fmx)])
if (fcx < wx-1) {
fmx <- c(fmx[!is.na(fmx)], rep(0, wx-fcx-1))
} else {
fmx <- fmx[1:(wx-1)]
}
# select fertility schedule
FERTN <- select.list(choices=
  names(Jfert)[grep("ASFR([0-9]+)", names(Jfert))],
  preselect="ASFR2010", title="Select the fertitiliy schedule")
asfr <- c(rep(0, 15), Jfert[, FERTN], rep(0, wx-55))
```

```
# Population projection for coming 100 years
# Ignoring migration
for (j in 2:100) {
  baby <- sum(fp[, j-1]*asfr)
  babym <- as.integer(baby*1.06/2.06+0.5)
  babyf <- as.integer(baby*1/2.06+0.5)
  mp[1, j] <- babym
  fp[1, j] <- babyf
  mp[2:wx, j] <- as.integer(mp[1:(wx-1), j-1]*(1-mmx))
  fp[2:wx, j] <- as.integer(fp[1:(wx-1), j-1]*(1-fmx))
}

par(cex=0.8, las=1)

# Actual changes
for (i in c(1:24*2, 52, 56)) {
  pyramid(data.frame(M=Jpopl[, i], F=Jpopl[, i+1], A=Jpopl[, 1]),
    Laxis=0:4*500000, AxisFM="d", Cstep=10,
    main=gsub("M([0-9]+)", "Year \\1", names(Jpopl)[i]))
  Sys.sleep(1)
}

# Projected changes
for (i in c(2:100)) {
  pyramid(data.frame(M=mp[, i], F=fp[, i], A=Jpopl[, 1]),
    Laxis=0:4*500000, AxisFM="d", Cstep=10,
    main=sprintf("Year %d: projected with\n %s, %s and %s",
      i+2009, mqxn, fqxn, FERTN))
  Sys.sleep(1)
}
```

## 8.3 システムダイナミクスモデル

人口予測に、出生、死亡、移動だけではなく、農業、環境、経済、政治、文化その他の要因の影響をモデル化して統合しようという試み。条件付き予測（下記）にならざるを得ないし、長期的には無力である。メドウズらの「成長の限界」「限界を超えて」で用いられた World3 の結果として得られた「どんなシナリオでも、人口と資本が指数的に増加し、その後崩壊する」という予測は世界中に議論を呼んだが、結局合意に至らなかった。World3 の予測が現実と合わなかったのは、1972 年以降に起こった出生率と死亡率の急速な低下

を予見できなかったからである。つまり、システムダイナミクスモデルでは、非線型の関係もモデル化できるが、その時点で手に入るデータのすべてを説明できるようなレベルで検討されたモデルを作ることは現実的に不可能に近いというのが、このアプローチの1つの限界である。もっといえば、ある時点で手に入るデータを説明できたとしても、新しい資源の発見とか、新技術の発明といったことを予測することは原理的に不可能であり、それは長期予測にとっては決定的な限界といえる。ローカルなシステムモデルは役に立つかもしれないと期待されているが、まだ十分な検討はなされていない。

無条件の予測は無理であるとなってきたので、なんらかのシナリオを想定した上での予測、つまり家族計画によって出生率が何パーセント抑制されると仮定すれば人口はどうかなる、といった形の、条件付き予測をするのが現実的である。

## 8.4 人口学におけるシミュレーションの意義

シミュレーションとは、日本語でいえば模擬実験、つまり、ある事象を実際にやる代わりにモデル（模型）で実験してみることをいう。中でも人口学では、ふつう、人口モデルの数値計算による予測あるいは過去の推計をさす。これらは、メカニズムよりも数値間に現れる関係式に基づいてアプリオリに立てたものが多く、観測不可能な変数も多い。この推計においてもっとも重視されるアウトプットは、幅はついているにせよ「推計値」である。しかし、考えてみれば、アプリオリに立てたモデルのアウトプットというものは、正しい（あるいは現実に即している）という保証がない。モデルが正しいかどうかを評価する手段がないのに推計値に意味をもたせるのは危険である。そうすると、人口学におけるシミュレーションには意味がないのだろうか？

この方面の先駆者である Bennett Dyke は、次の2つの用法でシミュレーションを使うなら意味があるといっている。一つは、シミュレーションによるデータの生成である。本質的に観測不可能な事象でも、シミュレーションによってデータを生成すれば、完全なデータセットを得ることが可能である。たとえば、聞き取りによって200年以上にわたる個人レベルの詳細な人口動態を完全に復元することは不可能である。したがって、信頼性のないデータが解析モデルにどのようなバイアスを与えるかを評価する必要がでてくる。こういう場合にシミュレーションを使えば、出生、死亡などについていろいろなレベルを設定した上で、完全な人口動態データを短時間のうちに得ることができる。この完全なデータに対して解析モデルを適用して得られる結果と、何らかの脱落を仮定して得られるデータに対して同じモデルを適用して得られる結果がどの程度一致するかによって、聞き取りの不完全さに対する解析モデルの頑健性を評価できるわけである。さらに、乱数の初期値だけを変えて何度もシミュレーションを実行すれば、与えた初期条件の下でのデータの分布を得ることができる。これは一般にコンピュータ集約型統計学と呼ばれる方法で、解析

的に求めにくい分布関数を，シミュレーションならば容易に求めることができる。

もう一つは，単純な解析的モデルが存在しないような，複雑な現象についてシミュレーションを用いることである。たとえば，結婚過程はそうである。なるほど，稲葉 (1997) が指摘するとおり，Fredrickson (1971) 以来，男女のペア形成の分布を考慮した非線形微分方程式モデルはいくつか提案されているが，たとえば，交叉イトコ婚を数学的に表現するには，まず交叉イトコという親族関係の変数を作り，年齢構造とは別にモデルに組み込んで，親族関係ごとに関数を定義しなければならないわけで，この手口ではきわめて難しい。その点，個人レベルのシミュレーションならばルールベースという形で結婚を表現することは容易である。他の方法ではまったく解が得られないなら，シミュレーションの利用は大いに価値があるのだ。

## 8.5 人口再生産シミュレーション

個人レベルで人口再生産を表現するにはどうしたらよいだろうか？ まず，シミュレーションの単位となる「個人」とはいかなるものか（どういう属性をもつか），決める必要がある。次に，人口再生産を表現するには，基本的に，出生と死亡の発生規則を決めなければならない。死亡は個人単位で決まるイベントと考えれば，死亡の発生確率は個人レベルで決めることができる（年齢の関数とすることが多い）。出生はペア単位で起こるイベントなので，ペア形成のルールを規定することと，ペアの出生力を決めることの両方が必要である。これら4段階を規定した後で，組み合わせることによって集団全体の人口再生産がおこなわれる。ただし，Goldman らの一連の研究が既婚未婚が死亡に影響を与えることを示唆しているので（たとえば Hu and Goldman, 1990），死亡を個人レベルのイベントとするこの枠組み自体も検討の余地を残している。

### 8.5.1 個人の定義

いいかえれば，ヒトの属性のうち，どの部分をピックアップするかということである。最低限必要なのは，個体識別子，親子関係，年齢（月齢，日齢），配偶関係であるが，居住地や疾病抵抗性遺伝子といった属性を扱うこともできる。どの属性を要素に含めるかは，死亡やペア形成のルールベースを決める際の必要性から決定する。C 言語では構造体として定義するのが自然である。

### 8.5.2 死亡発生確率

個人レベルでの死亡発生確率を決めるには，大きく分けて4通りの方法がある。モデル生命表を使う方法，死亡曲線のモデル（Siler モデルなど）を使う方法，Brass のロジット

システムを使う方法，及び，雪崩モデルである。個人単位の人口シミュレーションには雪崩モデルがもっとも適していると思う（ただし乳幼児死亡は別に考えなくてはならないが）。

### 8.5.3 ペア形成

これを扱えることがシミュレーションの最大の利点である。たとえば，年齢差何歳未満で，異なるクランに属していて，配偶者が何人以内といった条件を満たす配偶者を探して，候補がいればその中からある確率で結婚する，という複雑な配偶者選択ルールでも，シミュレーションプログラム上では，条件式のかたまりとして，わずか数行で記述できる。

### 8.5.4 ペアの出生力

まともに扱うことがきわめて難しく，まだ成功した研究はないと思う。Bongaarts の近成要因モデルに男性側の振る舞いを組み合わせることで，シミュレーションは可能かもしれないが，まだ行われていない。

### 8.5.5 擬似乱数列について

イベントをプログラム上で「確率的に」起こすために，乱数を用いる必要がある。つまり，確率 0.3 で起こるイベントであれば，区間 (0,1) の一様乱数を発生させて，その値が 0.3 未満だったらイベントを発生させ，そうでなければイベントを発生させない。数学的に厳密に乱数列を扱うことはきわめて難しいが，シミュレーションで用いるには，次の 4 つの性質を満たしていれば，擬似乱数列で十分である。(1) 連続する値の間に線形の関係がない，(2) 値の出現する期待値がとりうる値全部について等しい（と期待される），(3) 短い周期の繰り返しにならない，(4) 算法で生成される（初期値を決めれば再現できる）。R が使えるようになる前は，Lehmer の乗算合同法やメルセンヌツイスターを実装したライブラリをリンクする必要があったが，R ではさまざまな擬似乱数を組み込み関数として利用できるのも非常に便利である。

Pseudo random numbers ● In computer simulation, random (by-change) events are simulated by pseudo random numbers. ● Many kinds of pseudo-random number generators (PRNG) have been ever suggested. ? Traditionally, Lehmer's linear congruential method was widely used (eg. Park and Miller, 1988). ? Nowadays, Mersenne-Twister (Matsumoto and Nishimura, 1998) is de facto standard. ? In R, RNGkind() can specify various PRNG, but usually leave it as default Mersenne-Twister. ? In R, set.seed() can specify the initial seed of pseudo-random numbers. The same seed generate same pseudo-random numbers, so that it's

reproducible.

### 8.5.6 評価について

cflow や lint, GNU で公開されている Checker, あるいは小さなデータでのテスト実行を通してのプログラムの評価とは別に, モデルの評価 (妥当性検証と感度分析) を行う必要がある。結果をどの変数で評価するかということと同時に, 表現方法が問題となる。人口再生産シミュレーションで評価対象としてよく選ばれる変数は, 人口あるいはその増加率である。目的によってはもっと特異的な変数を選んでもよいが, 少なくとも妥当性検証については, 実測データが得やすいことが選択基準となる。表現方法としては, 平均と標準偏差であらわされることが多いが, 注意すべきなのは, 一例でも「偶然変わったことが起こりうる」なら重要だということである。つまり, たとえば, 同じパラメータで乱数列の初期値だけが異なるシミュレーションを 100 回実行した場合に, 100 回中 99 回は人口が単調増加したとしても, 残りの 1 回で集団が絶滅するならば, その 1 回の試行には大きな意味がある。それゆえ, 結果の表現としては, 全部のシミュレーション結果のグラフ表示を含むべきである。

妥当性検証については, 「1 度でも実測値と完全に一致するならば, 現実がそのシミュレーションで想定したメカニズムの通りに起こった可能性を否定できなくなる」という意味で, 完全一致するまでパラメータや乱数を変えて試行を繰り返すのが王道である。Dyke (1981) も MacCluer (1980) も, これを「シミュレーションならではの妥当性検証」として強く支持している。しかし現実には, 平均値の変化のグラフを書いて, 実測データの変化と重ねて「だいたい合っている」といって済ませている論文が多くみられるのは, 完全一致を達成するのが難しいからである (不可測なパラメータを増やせば容易だが, それでは意味がない)。次善の策としては, 平均値に標準偏差や標準誤差をつけたり, 中央値に上下 5 パーセントイルまでの幅をつけ, その範囲内に実測データが入ることを確認するか, カイ二乗適合度検定で「適合していないとはいえない」ことで消極的に妥当性検証を行うことである。

感度分析とは, モデルが妥当であるといえた上で, 個々のパラメータを変えたときの結果の変化を検討することである。「単純な解析モデルが存在しないような複雑な現象のシミュレーション」モデルを作る目的はここにある場合が多い。分布を求めることが目的なら必須ではないが, パラメータの性質を知るためにも有効なので, 常に感度分析はすることを奨める。

### 8.5.7 プロトタイプの作成

Definition of individual: Up to the purpose

- Here each individual has ID, SEX, AGE, MRD (initially 0, spouse's ID after marriage), NC (number of children ever born), LCB (year at latest childbirth) as variables in data.frame
- If we consider the postpartum amenorrhea, LCB should be recorded at month
- Initial population was calculated by stable population theory

● Mortality: Lifetable is needed.

Actual life table, model life table, or mechanism-based model (Avalanche model).

Here  $q_x$  from 2010 Japan's lifetable for males and females

● Mating: Coale-McNeil's model was applied

CM(scale=0.9, a0=16, k=1)

● Fertility: Coale-Trussell's model was applied

CT(M=0.9, m=1.0)

Stochastic occurrence of vital events

● For deterministic macro simulation, if  $N$  people have an event at probability  $p$ , the expected number of events is  $Np$ . If considering stochasticity in R, `rbinom(1, N, p)` generates expected number with binomial distribution.

● For microsimulation, if  $N$  people have an event at probability  $p$ , generate uniform random numbers in  $(0, 1)$   $N$  times, if the number is equal to or less than  $p$ , the event occur. It should be recorded for each individual. Otherwise, `rbinom(N, 1, p)` also simulate the event's occurrence for each individual (1=event, 0=no event).

他のコードから呼ばれる設定

```
# http://minato.sip21c.org/demography-special/MSIMsetting.R
# Microsimulation model for population growth by R
# common setting for 2 parts
# (C) Minato Nakazawa <minato-nakazawa@umin.net>
if (require(fmsb)==FALSE) { install.packages("fmsb"); library(fmsb) }
set.seed(1234567)
#
MA <- 100 # Maximum age
qxM <- Jlife$qx2000M
qxF <- Jlife$qx2000F
LxM <- clifetable(qxM)$Lx[1:(MA+1)]/100000
LxF <- clifetable(qxF)$Lx[1:(MA+1)]/100000
Marriage <- CM(0.9, 16, 1) # ages for 10:60
ASMFR <- c(rep(0, 12), CT(M=0.9, m=1.0), rep(0, 11), rep(0, MA-60))
ASFR <- c(rep(0, 10), Marriage$G, rep(1, MA-60))*ASMFR
cat(sprintf("Theoretical TFR = %5.3f\n", sum(ASFR, na.rm=TRUE)))
```

初期人口の生成

```
# http://minato.sip21c.org/demography-special/MSIMINIT.R
# Microsimulation model for population growth by R
# Part I: Making initial population
source("./MSIMsetting.R")
inysize <- 10000 # Initial population size
FASFR <- ASFR*(1/2.05) # Female only age-specific fertility rate
FBSP <- FASFR*LxF
NRR <- sum(FBSP, na.rm=TRUE)
mu <- sum((0:MA+0.5)*FBSP, na.rm=TRUE)/NRR
IRNI <- log(NRR)/mu
y <- 0:MA+0.5 # Mean ages
coefR <- exp(-IRNI*y)
stableM <- LxM*coefR
stableF <- LxF*coefR
Scale <- inysize/sum(stableF+stableM)
sMs <- as.integer(stableM*Scale + 0.5)
sFs <- as.integer(stableF*Scale + 0.5)
NMP <- sum(sMs)
MP <- data.frame(ID=1:NMP, SEX=rep(1, NMP), AGE=rep(0:MA, sMs),
  MRD=rep(0, NMP), NC=rep(0, NMP), LCB=rep(0, NMP))
NFP <- sum(sFs)
FP <- data.frame(ID=1:NFP, SEX=rep(2, NFP), AGE=rep(0:MA, sFs),
  MRD=rep(0, NFP), NC=rep(0, NFP), LCB=rep(0, NFP))
EMP <- c(rep(0, 10), Marriage$G, rep(1, MA-60)) # Proportion ever married
cumsFs <- 0
for (i in 1:(MA+1)) {
  NMRD <- EMP[i]*sFs[i]
  j <- 1
  while (j<=NMRD) {
    MAGE <- FP$AGE[cumsFs+j]
    SPAGES <- (MAGE-3):(MAGE+6)
    MPool <- subset(MP, (MRD==0))
    MPool <- subset(MPool, (AGE %in% SPAGES))
    if (NROW(MPool)>0) {
      SPID <- ifelse(NROW(MPool)==1, MPool$ID[1], sample(MPool$ID, 1))
      FP$MRD[cumsFs+j] <- MP$ID[SPID]
      MP$MRD[MP$ID[SPID]] <- FP$ID[cumsFs+j]
    }
    j <- j+1
  }
  cumsFs <- cumsFs + sFs[i]
}
TP <- rbind(MP, FP)
cat(sprintf("Sizes were %d/%d for Male/Female.\n", NMP, NFP))
write.table(TP, "./initialpop.txt", row.names=FALSE, sep="\t")
```

## シミュレーション (1)

```
# http://minato.sip21c.org/demography-special/MSIMEXEC.R
# Microsimulation model for population growth by R
# Part 2: Simulation
# (C) Minato Nakazawa <minato-nakazawa@umin.net>
# Stochastic random functions to simulate
# birth and death + calculate summary statistics
source("./MSIMsetting.R")
TP <- read.delim("./initialpop.txt")
FirstMarry <- c(rep(0, 10), Marriage$g, rep(0, MA-60))
#
simtimes <- 10
simyears <- 50
totalpop <- sexratio <- tfr <- depratio <- agingp <-
  matrix(rep(0, simtimes*simyears), simtimes, simyears)
#
# setting age-specific death rate
# use qxM and qxF defined in "MSIMsetting.R"
#
times <- 1
while (times <= simtimes) {
  years <- 1
  MP <- subset(TP, SEX==1)
  MAXIDM <- max(MP$ID)
  FP <- subset(TP, SEX==2)
  MAXIDF <- max(FP$ID)
  while (years <= simyears) {
    # for all people, check death. if die, clear spouse's spouse id & delete
    MAlive <- (runif(length(MP$ID), 0, 1) > qxM[MP$AGE])
    FP$MRD[FP$ID==MP$MRD[!MAlive]] <- 0
    FAlive <- (runif(length(FP$ID), 0, 1) > qxF[FP$AGE])
    MP$MRD[MP$ID==FP$MRD[!FAlive]] <- 0
    MP <- subset(MP, MAlive)
    MP$AGE <- ifelse(MP$AGE==MA, MA, MP$AGE+1) # increment ages
    FP <- subset(FP, FAlive)
    FP$AGE <- ifelse(FP$AGE==MA, MA, FP$AGE+1) # increment ages
    # for married females, calculate birth
    EMFP <- subset(FP, MRD>0)
    newborn <- (runif(length(EMFP$ID), 0, 1) <= ASMPFR[EMFP$AGE])
    WBFP <- subset(EMFP, newborn)
    males <- (runif(length(WBFP$ID), 0, 1) <= (1.05/2.05))
  }
  times <- times + 1
}
```

シミュレーション (2)

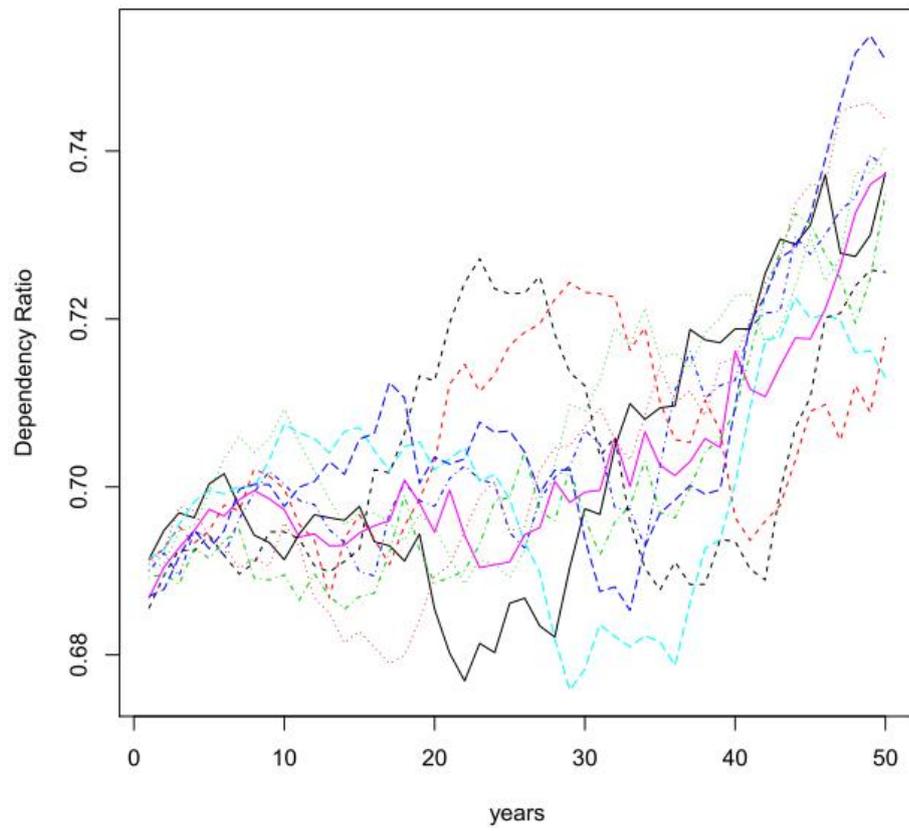
```

for (j in 1:length(newborn)) {
  if (newborn[j]) {
    if (males[WBFP$ID==EMFP$ID[j]]) {
      MAXIDM <- MAXIDM+1
      MP <- rbind(MP, c(MAXIDM, 1, 0, 0, 0, 0))
    } else {
      MAXIDF <- MAXIDF+1
      FP <- rbind(FP, c(MAXIDF, 2, 0, 0, 0, 0))
    }
    FP$NC[FP$ID==EMFP$ID[j]] <- FP$NC[FP$ID==EMFP$ID[j]] + 1
    FP$LCB[FP$ID==EMFP$ID[j]] <- years
    MP$NC[MP$ID==FP$MRD[FP$ID==EMFP$ID[j]]] <-
      MP$NC[MP$ID==FP$MRD[FP$ID==EMFP$ID[j]]] + 1
    MP$LCB[MP$ID==FP$MRD[FP$ID==EMFP$ID[j]]] <- years
  }
}
# make new couples for unmarried females (use Marriage$g)
UMFP <- subset(FP, MRD==0)
newmarry <- (runif(length(UMFP$ID), 0, 1) <= FirstMarry[UMFP$AGE])
for (k in 1:length(newmarry)) {
  if (newmarry[k]) {
    MAGE <- FP$AGE[FP$ID==UMFP$ID[k]]
    SPAGES <- (MAGE-3):(MAGE+6)
    MPool <- subset(MP, (MRD==0))
    MPool <- subset(MPool, (AGE %in% SPAGES))
    if (NROW(MPool)>0) {
      SPID <- ifelse(NROW(MPool)==1, MPool$ID[1], sample(MPool$ID, 1))
      FP$MRD[FP$ID==UMFP$ID[k]] <- SPID
      MP$MRD[MP$ID[SPID]] <- UMFP$ID[k]
    }
  }
}

```

シミュレーション (3)

```
# data recording
totalpop[times, years] <- length(FP$ID)+length(MP$ID)
sexratio[times, years] <- length(MP$ID)/length(FP$ID)
FPWC <- subset(FP, LCB==years)
tfr[times, years] <-
  sum(table(FPWC$AGE)/table(FP$AGE)[names(table(FPWC$AGE))])
TPP <- rbind(FP, MP)
depratio[times, years] <-
  (sum(TPP$AGE<15)+sum(TPP$AGE>=65))/sum(TPP$AGE>=15&TPP$AGE<65)
agingp[times, years] <- sum(TPP$AGE>=65)/length(TPP$AGE)
# proceed to next year
years <- years+1
}
# proceed to next simulation
times <- times+1
}
# output data
pdf("MSIMRES.pdf")
matplot(t(totalpop), xlab="years", ylab="total population", type="l")
matplot(t(sexratio), xlab="years", ylab="males / females", type="l")
matplot(t(tfr), xlab="years", ylab="Total Fertility Rates", type="l")
matplot(t(depratio), xlab="years", ylab="Dependency Ratio", type="l")
matplot(t(agingp), xlab="years", ylab="Aged proportion", type="l")
dev.off()
write.table(totalpop, "./totalpop.txt", row.names=FALSE, sep="\t")
write.table(sexratio, "./sexratio.txt", row.names=FALSE, sep="\t")
write.table(tfr, "./tfr.txt", row.names=FALSE, sep="\t")
write.table(depratio, "./depratio.txt", row.names=FALSE, sep="\t")
write.table(agingp, "./agingp.txt", row.names=FALSE, sep="\t")
write.table(TPP, "./aftersimpop.txt", row.names=FALSE, sep="\t")
```



## 第 9 章

# 二次資料からマクロな関連をみる

従来、人口分析のテキストではあまり取り上げられていなかったが、R の得意なところなので説明しておく。

もっとも有名な例は、国を単位とした一人当たり GDP と平均寿命の関連性であろう。これは、R でなくても Gapminder というソフトを使うと、長期間の変化をグラフィカルに簡単に見ることができる。

平均寿命には、一人当たり GDP よりも所得格差が小さいことの方が関連が大きい。USA は一人当たり GDP が 1993 年に 24,680 USD で平均寿命が 1993 年に 76.1 歳だが、より平等主義な国であるオランダ、イスラエル、スペインは一人当たり GDP が USA より低いけれども（それぞれ 17,340 USD, 15,130 USD, 13,660 USD）平均寿命は長いし（それぞれ 77.5 歳, 76.6 歳, 77.7 歳）、所得格差が世界最小のスウェーデン（78.3 歳）と日本（79.6 歳）は世界最高の平均寿命を享受している。

ただし、日本は現在では所得格差が広がってきているにもかかわらず平均寿命は世界最高のままなので例外かもしれない。

“The World Factbook 2007” (CIA) から、これらの国の一人当たり GDP（購買力平価ベース）、Gini index（所得格差が小さいほど小さな値になる）、平均寿命を下の表に示す。ちなみに、一人当たり GDP はルクセンブルクが 71,400 USD で世界最高である。ソロモン諸島は 600 USD に過ぎないが平均寿命は 73.16 歳である。なお、一人当たり GDP を市場為替ベースで計算すると、スウェーデンの方が USA より多くなる（IMF のサイトを参照）。

国名	一人当たり GDP (USD)	Gini index	平均寿命
USA	43,800	45	78
オランダ	32,100	30.9	79.11
イスラエル	26,800	38.6	79.59
スペイン	27,400	32.5	79.78
スウェーデン	32,200	25	80.63
日本	33,100	38.1	82.02

OECD のデータはオンラインで公開されている。

# 文献

1. Chamberlain A (2006) *Demography in Archaeology*. Cambridge University Press.
2. Dyke B (1981) Computer simulation in anthropology. *Annual Reviews of Anthropology*, 10: 193-207.
3. Rowland DT (2003) *Demographic methods and concepts*. Oxford Univ. Press.
4. Mori Y, Nakazawa M (2003) A new simple etiological model of human death. *The Journal of Population Studies (Jinko-Gaku-Kenkyu)*, 33: 27-39.
5. Namboodiri K, Suchindran CM (1987) *Life table techniques and their applications*. Academic Press.
6. Newell C (1988) *Methods and models in demography*. The Guilford Press.
7. Preston SH, Heuveline P, Guillot M (2001) *Demography: Measuring and Modeling Population Processes*. Blackwell Publishing.
8. Smith DP (1992) *Formal Demography*, Plenum Press
9. Wood JW (1994) *Dynamics of Human Reproduction*, Aldine de Gruyter, New York
10. Wood JW, Holman DJ, Weiss KM, Bucanan AV, LeFor B (1992) Hazards Models for Human Population Biology. *Yearbook of Physical Anthropology*, 35: 43-87.
11. 稲葉 寿 (2002) 『数理人口学』, 東京大学出版会
12. 稲葉 寿 [編] (2007) 『現代人口学の射程』, ミネルヴァ書房
13. 大塚柳太郎・鬼頭宏 (1999) 『地球人口 100 億の世紀：人類はなぜ増え続けるのか』ウェッジ選書.
14. 岡田 實・大淵 寛 [編] (1996) 『人口学の現状とフロンティア』, 大明堂
15. 河野稠果 (2007) 『人口学への招待 少子・高齢化はどこまで解明されたか』中公新書
16. クヌース (1981) 『THE ART OF COMPUTER PROGRAMMING 3. 準数値算法/乱数』サイエンス社.
17. 黒田俊夫 (1976) 『日本人口の転換構造』, 古今書院
18. 国際人口学会 [編][日本人口学会 訳](1994) 人口学用語辞典, 厚生統計協会

19. ジョエル・E・コーエン著，重定南奈子・瀬野裕美・高須夫悟訳 (1998) 『新「人口論」：生態学的アプローチ』，農文協.
20. 舘 稔 (1968) 『人口分析の方法』，古今書院
21. 中澤 港，大塚柳太郎 (1997) 出生力の指標としての TFR と完結パリティの評価. 人口学研究, 21: 61-63.
22. 日本人口学会 [編] (2002) 『人口大事典』，培風館
23. マルサス (1973) 『人口論』中公文庫. (生活財が等差数列的にしか増加しないのに人口が等比数列的に増加するから生活水準を維持するには人口抑制が必要だと論じた古典。原著初版は 1798 年刊)
24. 和田光平 (2006) 『Excel で学ぶ人口統計学』，オーム社
25. 和田光平 (2015) 『人口統計学の理論と推計への応用』，オーム社

# 付録：R について

## 9.1 R の基本情報とインストールの方法

R は、R Development Core Team が開発し全世界の有志によってメンテされているフリーなデータ解析環境である。2017 年 5 月現在の最新版は 3.4.0 である。

## 9.2 人口分析に関連したパッケージ

標準的なインストーラでインストールすると `survival` パッケージが入っているので、ハザード解析などは `survival` パッケージをメモリにロードするだけで使える。

日本のマクロなデータ、生命表関連の関数、いくつかの人口モデル関数が拙作 `fmsb` パッケージに入っている。

人口構造を見るために欠かせない人口ピラミッドは拙作 `pyramid` パッケージの他、`plotrix` パッケージや `Pyramid` パッケージ\*1 が発表されている。

Rob Hyndman が中心になって開発し、コアチームも含めて大勢の協力者を得てメンテナンスしている `demography` パッケージには、HMD (human mortality database)\*2 を簡単に扱うための関数 `hmd()` や、Lee-Carter モデルを使うための関数 `lca()` など、有用な関数が多数入っている。

ベイズ流の人口推計をするには、`bayesPop` パッケージが使える。`bayesPop` をインストールすると、2008 年の国連の世界人口推計データをパッケージ化した `wpp2008` も依存パッケージとしてダウンロードされる。

国連人口推計データは、<http://esa.un.org/wpp/index.htm> から、ワシントン大学の Hana Sevcikova がパッケージ化しメンテナンスしているもので、`wpp2010`、`wpp2012`、`wpp2015` も既に発表されている。

---

\*1 <http://sites.google.com/site/timriffepersonal/r-code/pyramid>。Tim Riffe 氏が開発している

\*2 <http://www.mortality.org/>

# 索引

安定人口モデル, 22  
簡易生命表, 7  
合計出生率, 7  
死亡, 7  
出生, 7  
人口構造指標, 22  
人口動態, 7  
人口動態統計, 7  
ゼロ歳平均余命, 7, 53  
平均寿命, 7, 53

---