

# 統計研究参考資料

No. 63

フィンランドにおける

レジスター・ベースの統計生産

— 翻 訳 —

2000年1月

法政大学日本統計研究所

The Japan Statistics Research Institute  
Hosei University





# 1 フィンランドにおけるレジスター・ベースの センサス統計の統計的特徴と品質

Riitta Harala、フィンランド統計局

(シンポジウム'97 サーベイとセンサスにおける新しい方向、1997年11月5-7日、オッタワ)

## 1. 統計的目的への行政データ使用の背景

過去20～25年の間にフィンランド、そしてまた北欧諸国においては、行政記録からの統計目的でのデータの利用が継続的に増加してきている。実際に今日、フィンランド統計局が収集しているデータの約94%は、行政的源泉からのものであり、データ収集の6%だけが、人口あるいは企業からの直接的データ収集に基づいている。

フィンランドにおける行政レジスターの利用を促進した要因は幾つかあるが、最も重要なものは次の要因である。

**唯一の識別番号の使用** フィンランドにおいては、唯一の個人識別番号が使われており、われわれはまた、あらゆる中央レジスター—そのうちのにはとりわけ、社会保障と納税の事項のものがある—に使われている、幾つかの他のほとんど一律の識別番号システム（企業、家、住居の識別）を持っている。

**行政は国家規模のデータベースを構築することに利益を認めている** 重要な要因はまた、社会保障と税についての大部分のシステムが国家レベルのものであることあるいは、行政自体が、それらシステムを一つの国家規模のデータベースに結びつけることに利益を持っていることである。さらに、最近年には、社会保障の基本的給付のちらばりを、ひとつの同じ部局に集中する努力が払われてきた。フィンランドの社会的給付の大半も課税されているが、このことは、税当局は、税のための社会保障機関からの情報を受け取ることを意味する。

**国民による受容** レジスターの使用を促進する非常に重要な要因はまた、行政諸目的のために収集されたデータを統計目的に利用するが、国民から広く受け入れられているという事実である。社会とその働きの合理性に対する国民の信頼はかなり強い。また、秘匿性の問題に関わってのフィンランド統計局への評判も非常に良い。これが受け入れられているのは、行政目的のために既に収集されたデータを統計目的にも使うことは、同じデータを調査票で収集するより一層安い点で、合理的

であると考えられている事実から来ている。また、人々や企業の回答負担を少なくするという点も、既にどこかに集められているデータを利用する非常に重要な理由である。

**法的基礎** 統計目的のために行政データを使用する法的根拠ももちろん、非常に重要な要因である。フィンランドでの行政レジスターを使用する法的根拠は、すでに幾つかの行政機関が保有しているデータを使用するのは合理的であるという広く行き渡った意見の反映である。これが、われわれの統計法が、行政データは利用可能なときには常に使われるべきであるという原理に基づいている理由である。統計法はまた、フィンランド統計局に行政データを獲得し、使用し、リンクさせる広い可能性を許している。法はまた、われわれが統計法規において非常に重要な問題であると考えているデータの秘匿性をも規定している。

## 2. 人口センサスにおける行政データの使用

行政目的で収集されたますます多くのデータをまた統計を編集するために使うという動きの最善の例は、フィンランドでの人口・住宅センサス制度である。フィンランドは行政記録を人口センサス統計に使うという長い伝統を持っている。レジスター・データの使用は、行政データがはじめて使用された 1970 年センサス以後、センサスごとに増大してきた。

1990 年人口・住宅センサスは、フィンランドではじめて、全人口向けの調査票なしに、すべてレジスター・ベースで行われた。その後、1995 年センサスは、同じシステムに基づいて行われ、われわれは、2001 年センサスもまた同じ方法を使用するつもりである。

レジスター・ベースのセンサスシステムの構築は、既に 1980 年代のはじめに、何よりも費用が理由になって開始された。行政レジスターは、そのときまでに十分に発展していて、レジスター・ベースのシステムを可能にしていた。統計の必要の拡大もまた、レジスター・ベースのシステムを発展させた主な理由のひとつであった。レジスター・ベースのシステムによる統計生産は、調査票によるよりも、はるかに経済的である。このことは、センサスを 5 年毎とあるいは 10 年毎よりもひんぱんに実施できることを意味する。

実際に、この作業はまた、センサス型の統計を毎年生産する結果になった。レジスター・ベースの建築物・住宅統計の毎年の生産と人口の経済的活動についてのレジスター・ベース統計（地域的就業統計）の生産の両方が 1987 年にスタートした。例えば、地方自治体あるいは地方自治体の下部地域ごとの就業統計の利用者たちは、それらデータを毎年持たない状況を考えることすらできなくなった。地域的就業統

計システムはまた、フィンランド統計局の人口統計部門の新統計システムの大部分において、ますます重要なバックグラウンドデータになった。例としては異なる種類のフロー統計、教育指標、外国人についての統計その他がある。

## 2.1 方法

統計を編集する際に、行政記録を使用する上で利用可能な手段は多様である。行政データを利用する方法は、原理的には、次の種類に区分できる。例えば、

- 1) 直接的製表
- 2) サンプリングや計算のフレームワーク
- 3) モデルに基づいた推定、そして
- 4) レジスター推定

フィンランド統計局は、レジスターの利用においてこれらすべての方法を使っている。レジスター推定は、フィンランドでレジスターベースのセンサスシステムと地域的就業統計の下部システムで使われる第一の方法である。これは、決定規則のモデルあるいは最適グループ—これによって、人が行政データの1つ以上の項目から、人口の各単位を示す統計的概念に適合する最大限すくれたパラメーター値をひきだすことができる—を構築するために、調査票ベースのデータを使用する。

この方法はこのように、幾つものレジスター・ベースを同時に使用する。それによって、例えば、個人の主な経済活動をひきだすためには、20以上のレジスター・データを使用することが必要である。レジスターのこの共同利用は、個々のレジスターが必ずしも完全に網羅的でないことということを意味する。それらは、互いに補うよう使うことができるのである。同じ個人についての情報は、このようにして異なるデータのセットから入手可能であり、それらの中から、材料という点で全体的に見て、最も信頼できるものを選出することができるのである。

個々の現象についてのデータは通常、幾つかの源泉から引きだされるので、この方法はレジスター・データの変化に対して特に対抗力があるわけではないことに注意するべきである。このことは、個人の就業データは、雇用年金制度、納税レジスター、個人自らあるいは使用主が提出する報告書、を通じて入手可能なことを意味する。使用主が、例えば、税務機関に報告する義務を怠うなら、データはともかくも、年金制度あるいは個人自らの納税申告書を通じて入手可能なのである。

## 2.2 レジスター

フィンランドの人口センサス制度は、人口センサスの対象単位とそれら相互のつながりについてのデータを与えるよう社会が維持している基本的レジスターによって、それらの単位は、個人、家族、世帯、建築物（とコテージ）、住居（と事業店舗）、企業、事業所からなる（付録 1）。

中央人口レジスター (*The Central Population Register*) は、レジスター・ベースのセンサス制度においては、対象となる人口を定義する基本的なレジスターである。就業統計にふくまれる人口は、12月31日にフィンランドでの法的住所を持つすべての人々である。人口のすべての基本的な人口特性（識別番号、性、年齢、婚姻上の地位、出生地その他）は、中央人口レジスターからひきだされる。

建築物・住居のレジスター (*The Register of Buildings and Dwellings*) は、すべての建築物と住居を定義する基本的レジスターである。それはまた建築物と住居の特性をふくむ。レジスターの基本的データは、1980年の人口・住宅センサスで収集され、以後、このレジスターがおかれている人口レジスター・センターによって維持されている。

フィンランド統計局企業・事業所レジスター (*The Statistics Finland's Register of Enterprises and Establishments*) は、行政レジスターではないが、同じ種類の特性を持っている。例えば、産業部門、所有のタイプ、仕事場の法律上の形態および所在は、企業・事業所レジスターと、公共団体・地方政府職務単位のレジスター (*Register of Public Associations and Local Government Functional Units*) を使って結びつけられる特性である。

これらの基本的レジスターの他に、センサス制度はまた、約 30 の他のレジスター・ファイル（付録 2）を使って、そのデータをセンサス単位と結びつけている。それらのファイルとして以下のものがある。

### 労働省求職者レジスター (*The Ministry of Labour's Register of Job Applicants*)

このレジスターは、職業紹介所を通して求職しているすべての人をふくんでいる。このレジスターからひきだされる最も重要なデータは、

- 失業者になったときと終わったとき（その年の各失業期間）
- 失業が終わった理由
- 雇用が始まったときと終わったとき。

兵籍と良心的兵役拒否者についての情報は、総司令部の兵籍レジスター (*the*

*General Head Quarters Register of Conscripts*) と労働省の良心的兵役拒否者レジスター(*the Ministry of Labour's Register of Conscientious Objectors*) から得られる。これらのレジスターから得られたデータは、個人識別番号と兵役/非兵役が始まったときと終わったとき、である。

就業関係についてのデータは、幾つかの異なった源泉から受けとる。これらの関係は、3つの異なるグループに分けることができる。すなわち、(自営業者をふくむ)民間部門の就業関係、政府部門内、(地方自治体の連合をふくむ)地方政府部門内、である。

中央年金保障機関 (*The Central Pension Security Institute*) は、民間部門の就業関係に関するほとんど総てのデータを提供する。それは、1年内の1カ月以上にわたる期間におよぶすべての就業関係を記録している。さらに、就業関係が始まった日と終わった日をふくむ。

中央年金保障機関が提供するデータのうち、以下のデータが最も重要である、すなわち、

- 企業名
- 雇用関係がはじまった日と終わった日
- 年金制度のタイプ
- 年金規制番号

民間部門の就業関係データの他の源泉は、全国教会理事会の従業者レジスター(*the National Church Board's Register of Employees*)と、(中央年金保障機関からの)船員年金制度レジスター(*the Register of Seaman's Pension Scheme*)である。

中央政府部門(*the Central Government sector*)の雇用関係データの主な部分は、中央政府の雇用者についての国家財務省レジスターから来る。例えば、それらデータとして次のものがある。

- 部署番号
- 事業所の数
- 雇用関係が始まった日と終わった日
- 職業

中央政府部門の雇用関係についての他のデータ源泉としては、社会保険機関 (*the Social Insurance Institution*)、フィンランド銀行(*the Bank of Finland*)、ヘルシンキ大学(*Helsinki University*)および郵政電信電話局(*the Post and Telecommunication Office*) のレジスターである。

地方政府部門の雇用関係データの大部分は、地方政府と地方自治体連合(*municipal federation*)の雇用関係のレジスターを管理している地方自治体年金機関(*the Municipal Pension Institution*) から来る。例えば、受け取るデータは



- 団体会員番号
  - 雇用関係がはじまった日と終わった日
  - 職業
- である。

オーランド政府の雇用関係レジスター (the Employment Relationship Register of the Government of Åland) は地方政府部門の雇用関係情報についてのもうひとつの源泉である。

異なる源泉からのすべての就業関係情報を収集する作業との関連で、独立した就業関係レジスター(employment relationship register)が編集される。このレジスターは、その年間の就業関係のすべてをふくむ。

すべての雇用関係の賃金/俸給についてのデータが、**国税理事会の所得・年金レジスター(the Income and Pension Register of the National Board of Taxes)** (賃金と税控除についての使用主申告書) から獲得される。これらのデータは、年末時点で有効な雇用関係を選ぶときに必要になる。

学生についてのデータは、また幾つかの源泉から獲得される。**フィンランド統計局大学生レジスター(Statistics Finland's Register of University Students)**は、秋学期に大学でんでいるか、大学に登録している者について情報を、大学での学習のタイプとともに提供する。

**学生奨学金レジスター(the Register of Study Aid to Students)**は、その年(秋学期と/あるいは春学期の期間)に与えられた奨学金のデータを提供する。

**共同選抜レジスター(the Joint Selection Register)**は、当該年あるいは先立つ2年間の間に、教育機関において最上級学年レベルでの学習をはじめた学生について情報を提供する。

**教育省高校・職業訓練学校の学生レジスター(the Ministry of Education's Register of Students in Senior Secondary Schools and Vocational Schools)** は、これらの学校の学生のコードと教育機関の所在地についてのデータを提供する。

**社会保険機関のレジスター(the Social Insurance Institution's Register)** は、年金受給者の情報を提供する。就業統計で使われる統計は年金の開始日と年金のタイプについてのものである。

**国税理事会のレジスター(the Register of the National Board of Taxation)**は、納税に関わる個人の諸税、資産、負債とともに、所得と給付に関する多くの異なるデータを提供する。

**フィンランド統計局の最終教育・学位レジスター (Register of Completed Education and Degrees)** は、フィンランド国民が海外で得た学位の大部分とフィンランドで得たすべての学位をふくんでいる。

最終教育・学位レジスターは、就業統計について最高度のデータを提供する。このレジスターは個人が持つすべての学位をふくんでいるので、それらを使うこともまた可能である。取得学位を叙述する情報は次のものを含んでいる。すなわち、

- 教育の6桁コード
- 学位取得の日付
- 教育機関の所在地
- 教育機関のタイプ
- 教育機関のコード

### 2.3 レジスター・ベースセンサス統計の品質

フィンランドのレジスター・データの信頼性は、レジスター・ベース・センサス制度の導入についてのあらゆる決定の前に既に検討された。1990年センサスとの関連でもまた包括的な評価調査が行われた。この評価研究においては、レジスター・ベースで生産されたと同じデータが、母集団の2%標本からの調査票ベースで収集された。この結果は別個の出版物として発行され、英語でも利用可能である。

年次的レジスター・ベースの就業統計の品質をモニターする最も重要な手段は、労働力調査を参照資料として使う継続的な年次的品質管理である。1995年センサスデータの信頼性もまた同じやり方で評価された。

労働力調査は、参照資料として2つの目的に仕える。すなわち、2つの方法で生産された結果のレベルをモニターすることと、この方法が同じ仕方で分類されたデータを生産する程度をモニターすること、である。

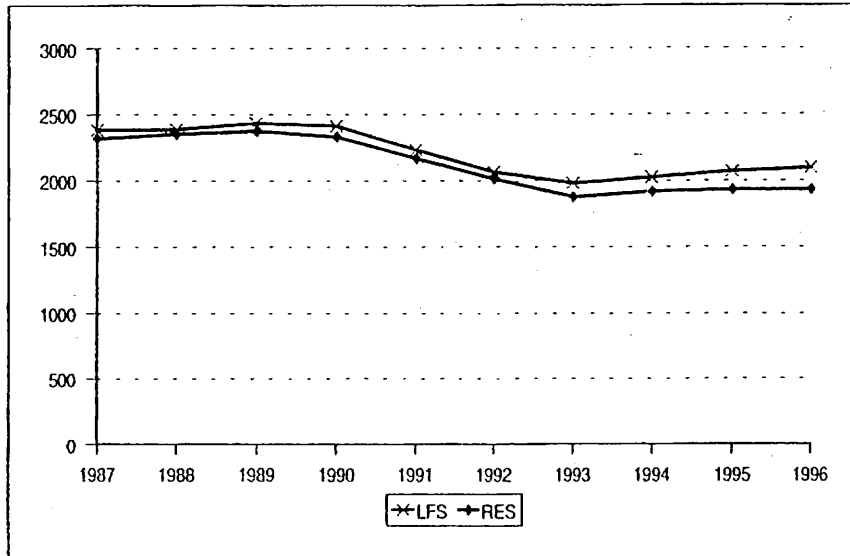
労働力調査とレジスター・ベースの就業統計は特徴的にはわずかながら異なる結果をもたらした。年次的には、結果のレベルあるいはそれらの間に大きな食い違いは観察されなかった。次の図(図1)からは、数字のレベルはわずかに違うが、動きは同じであることを見ることができる。

レジスター・ベースの就業統計の品質はまた、単位レベルのデータを労働力調査で収集されたデータと比較することによって、毎年モニターされる。労働力調査のデータは次の3つの仕方で使用される。すなわち、

- 1)データ処理における誤差を確認する
- 2)決定規則での変更を要求する状況を確認する、そして
- 3)結果のレベルをチェックする

ミクロレベルの品質管理は、実際には、単位レベルで労働力調査(LFS)の12月標本を、レジスター・ベースの地域就業統計(RES)の対応する標本(両方で同じ人物についての面接によるデータとレジスターベースのデータ)とともに縦横の表に

図1. 労働力調査(LFS) (12月標本) と (レジスター・ベースの) 地域就業統計 (RES) に示される就業人口



することで行われる。標本の大きさは約 12,000 であり、調査対象期間は LFS では 12 月の第 2 週、RES では 12 月の最終週である。この時間的差違は、比較からえられる結果を解釈するときには考慮されなければならない。

比較は、標本をなしている人口の主な経済活動や産業部門についてのレジスター・ベースと面接ベースの情報を縦横の表にすることで行われる。両システムにおいて同じ仕方で分類された人は対角線におかれ、異なって分類された者は対角線から離れた場所にくる。比較は 1987 年から行われ、このとき対角線からの偏差は正確に分析された。対角線からの偏差が前年のものと異なるなら、この場合は基本資料にあたってより詳細に検討される。偏りが、データ処理における誤差によるのであれば、この誤差は除去される。例えば、誤差がレジスター・データの変更あるいは関連する立法によって生み出しているなら、このデータの処理を規制する規則において必要な変更が加えられるかも知れない。

偏りは、しばしばレジスター制度と面接法の違いから生じているとみることができ、どちらの結果が正しいと明確に語ることはときとして不可能である。すなわち、面接によって得られた情報が全く正しいものとは必ずしも見なすことはできない。実際には、面接者あるいは被面接者が矛盾する情報に関わる状況の下で判断しなければならない場合がしばしばあり、他方でレジスター法は、矛盾を解決する論理的

規則を持っている。そういった場合のレジスター法の優位性は、その論理的な性格である。というのは、同じ情報を保有する人は異なる結論に到達するのに対して、機械は常に同じ仕方での異なる問題を解決するからである。

労働力調査を地域就業統計に対比して作る縦横の表に見いだされる対角線要素の時系列を付録3にふくめている。

## 2.4 レジスター・ベースのセンサスタイプの統計の統計的特徴

例えば、レジスター・ベースの就業統計システムは、人々が生涯にある状態から他の状態へ移動する（例えば、若者がどのように労働生活に入るか）過程を検討するすぐれた機会を提供することができる。それらは、経過の様々な時点での主な経済活動および就業を調べることを可能にするからである。就業統計システムは、叙述に関する限りで次の長所を持っているからである。

1. この資料が全体を数えることに基礎を置いている。このことは、正確な分類（例えば、教育の）、や地域から見ての分析をもまた可能にする。この状況は個人のレベルでもまた検討できる（例えば、 $y$ 年と $y+1$ 年の）。
2. この資料は毎年編集される、したがって年次間のフローもモニターできる。
3. この資料は、地図の座標に基礎をおく。したがって、統計単位は地理的に位置づけることができるし、その資料は座標に基礎を置く地域的単位のどのような集合についての統計を編集するためにも使うことができる。
4. この資料は識別コードをふくんでいる。例えば、ある与えられた時点での労働市場に対する関係での個人の状況を、データコード（社会保障番号）に基づいて検討することができる。その資料はまた、分析の目的に対して要求される他のレジスター・ファイルとリンクさせることができる。
5. この資料はあらゆる活動の形態を記録している。これは同時の学習と仕事と、同時に幾つかの仕事を持つこともふくんでいる。

## 文献

*Harala, Riitta*: Continuous quality check of register-based employment statistics, European Workshop on Using Administrative Data in Population and Housing Censuses, Statistics Finland and Eurostat, Helsinki 1995

*Harala, Riitta*, Evaluation of the results of the register-based population and housing census 1990 in Finland, Statistical Journal of the United Nations ECE(1996)63-72, IOS Press

Evaluation study of the 1990 Census, S Statistics Finland, SVT, Population Census, Volume 9B, Helsinki 1994

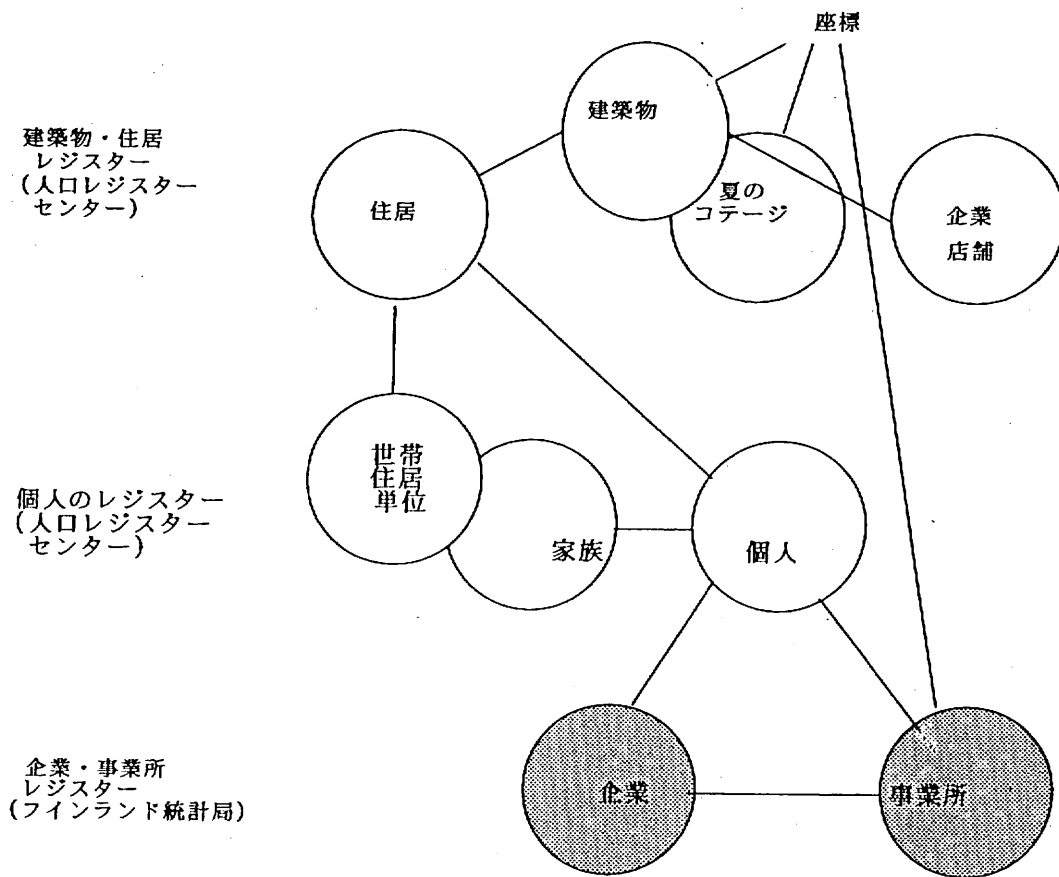
*Laihonen, Aarno*. Making gold from scrap, IAOS and Statistics Iceland, 5<sup>th</sup> Independent Conference Reykjavik, Iceland, July 1996

*Ruotsalainen, Kaija*: The register-based employment statistics in Denmark, Finland, Norway and Sweden, Statistics Finland and Eurostat, Helsinki 1996

*Vihavainen, Hikka*: Potential use of administrative registers to alleviate the burden on census or survey respondents, paper prepared for the third Mondorf Seminar 25-26 January 1996

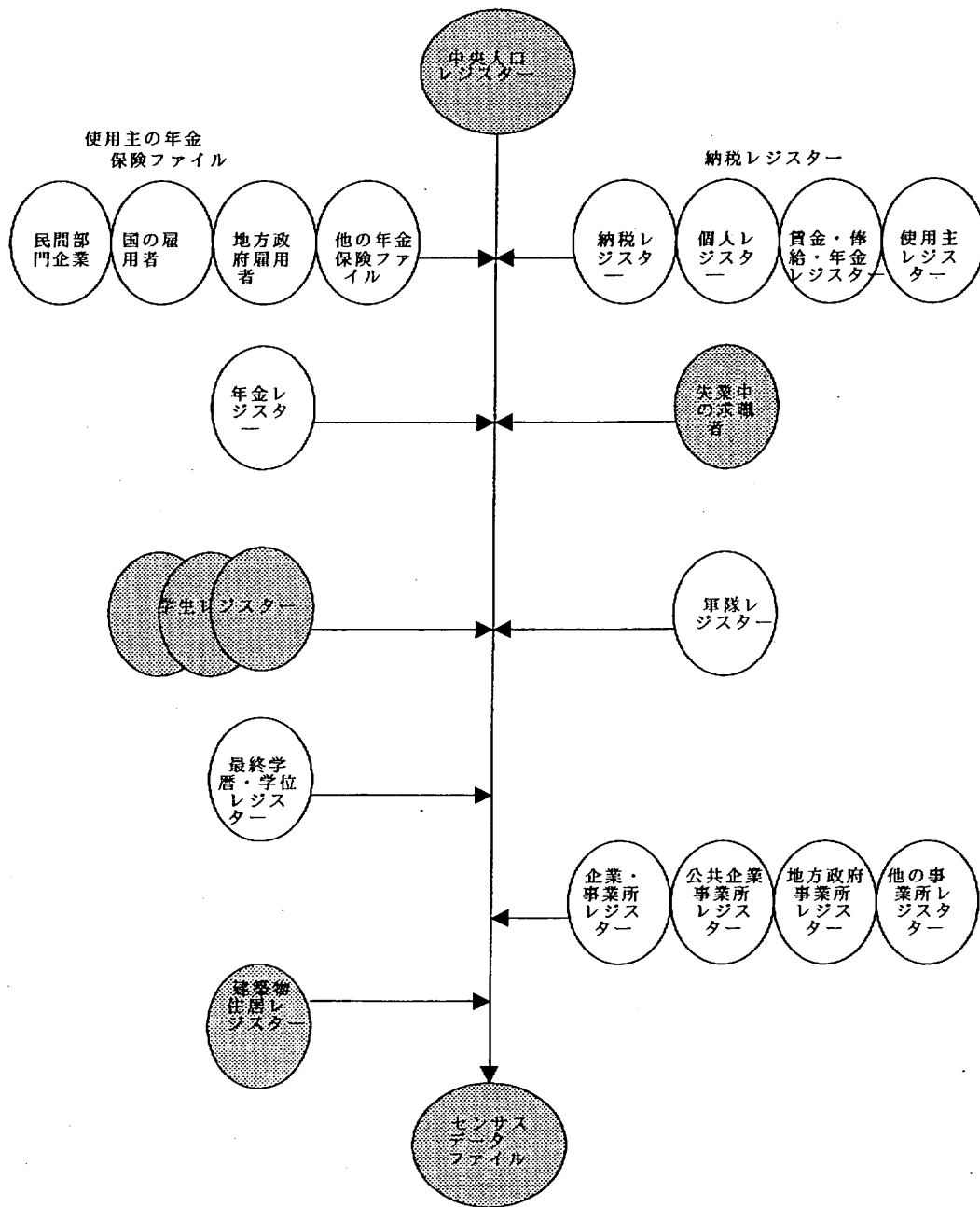
付録 1

レジスター・ベースのセンサス  
に属する単位とそれら相互間のリンク



付録2

レジスター・ベースのセンサス制度  
 におけるレジスターと行政記録の利用



### 付録3

#### 労働力調査と地域就業統計の組み合わせ表における対角線要素

	1995	1994	1993	1992	1991	1990	1989	1988
労働力	95.5	95.5	95.4	94.3	95.2	95	94.7	94.3
被用者	95.2	95.3	94.5	93.3	93.9	94.6	94.7	94.1
賃金・俸給稼得者	94	94.4	94.1	92.5	92.4	93.2	92.9	93
自営業者	83.6	80.5	79.9	78.3	81.4	80.9	82	80.3
失業者	81.7	80.6	80.3	77.2	75.2	60.7	61.2	64.3
学生	87.1	85.3	87.8	88	88.4	85.3	83.1	85.9
年金生活者	95.4	95.3	94.6	95	93.8	94.1	94.1	93.2
徴兵された者	97.6	95.2	94	96.5	96.5	56.7	97.2	98.6
その他	59.4	58.5	53.2	52.3	55.6	51.1	46	45.2



## 2 フィンランドにおけるレジスター・ベースの1990年人口・住宅センサス結果の評価について

Riitta Harala, フィンランド統計局人口統計部  
フィンランド統計局 FIN-00022、フィンランド  
Statistical Journal of the United Nations ECE  
11(1994), IOS Press

フィンランドの1990年の人口・住宅センサスは、調査票なしの、すべてレジスター・ベースで実施された最初のセンサスであった。フィンランドの人口センサス制度は、人口センサスの対象となる単位とこれらの単位間の結びつきに関するデータを提供する社会の基本的な行政レジスターに依拠している。それらの単位は、個人、家族、世帯、建物、住居、企業、事業所、から構成される。

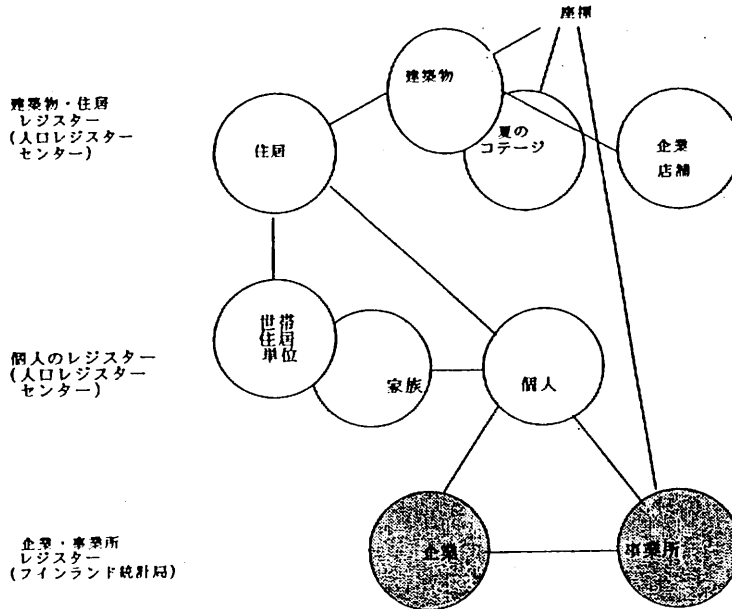
あるデータ収集の方法から他のそれへの転換が、時系列上にある種の断絶をもたらすのは避けられない。こうした断絶のために、調査票ベースのセンサスデータとレジスター・ベースのセンサスデータとの違いを分析する集中的な評価研究を行うことが、フィンランド統計局で決議された。

### 1. 序

フィンランドにおいて、1990年の人口センサスは、全人口に調査票を配布することなく、完全に行政レジスターに基づいて実施された最初のものであった。レジスター・ベースのセンサス制度を構築する作業は、主に予算上の理由から、すでに1980年代初頭に始まった。行政レジスターは、その時期までに、レジスター・ベースの制度を可能にするほど十分に発展していた。統計へのニーズの高まりも、レジスター・ベースの制度を発展させる主な理由の1つであった。レジスター・ベースの制度による統計の生産は、調査票ベースのそれよりもはるかに経済的である。これは、5年あるいは10年おきの周期でよりも頻繁にセンサスを実施できることを意味する。レジスター・ベースの建物・住宅統計の毎年の生産、および人口の経済活動に関するレジスター・ベースの統計（雇用統計）の生産は、どちらも1987年に始まった。

フィンランドの人口センサス制度は、人口センサスの調査対象単位とこれら単位間の結びつきに関するデータを提供する、社会の基本的なレジスターに依拠してい

図1. レジスター・ベースのセンサスに属する単位とそれら単位間の結びつき



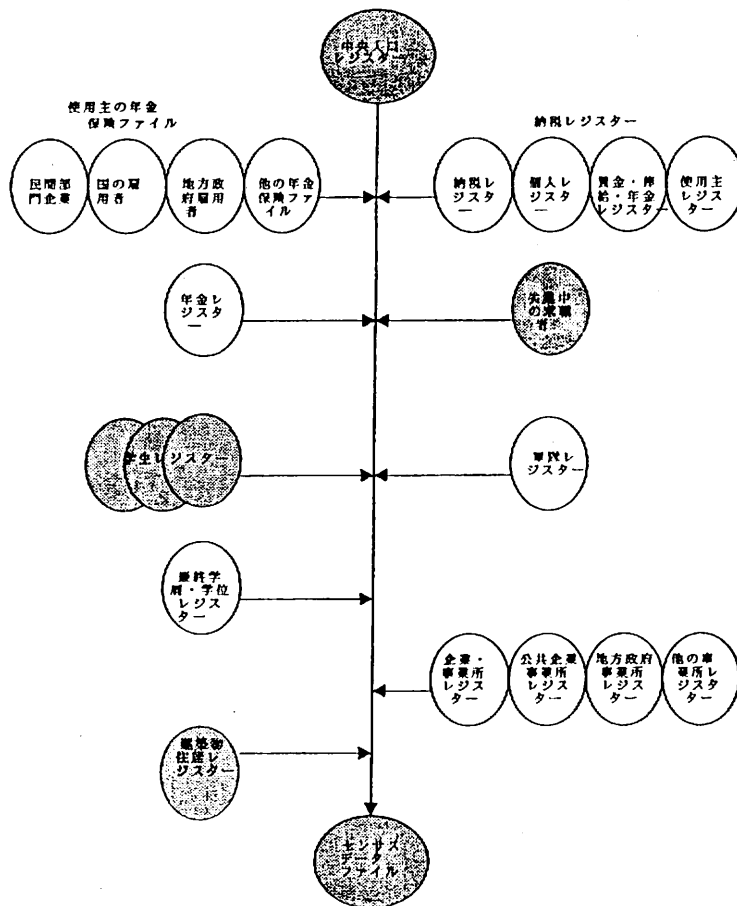
単位間の結びつきに関するデータを提供する、社会の基本的なレジスターに依拠している（図1）。それらの調査対象単位は、個人、家族、世帯、建物（夏の Cottage）、住居（と店舗）、企業、事業所、から構成される。

中央人口レジスターには、フィンランドに住むすべての人のデータが、その家族と世帯のメンバーに関するデータとともに含まれている。人口レジスター・センターの建築物とアパートのレジスターは、単位のデータとすべての建築物・住居・店舗の特徴をふくんでいる。フィンランド統計局の企業と職場のレジスターは、順に、企業と、企業が活動する場所に関するデータを含んでいる。個々のレジスターは、標準的な識別番号を用いており、単位をそれらの番号を使って相互に結びつけることができる。すなわち、人々は、自分たちが住む建物やアパート、自分たちが働く企業や職場と結びつけられるのである。すべての単位は、地図の座標によって位置づけることができ、これによって、地域統計の生産が可能になる。

人口センサスでは、これらの基本的なレジスターに加え、およそ 30 の他のレジスターファイルを利用した（図2）。

これらのレジスター・ファイルからのデータは、センサスの単位と結びつけられた。これらのファイルには、たとえば次のものが含まれている。

図2. 1990年センサスにおけるレジスターと行政記録の利用



- 特に、主な活動のタイプと仕事のタイプ（産業部門、使用主の部門、職場の所在）を判定するために、人々の所得、所得のタイプ、使用主に関するデータを提供する租税ファイル。
- 主な活動と就業のタイプを判定するために、就業の継続とタイプに関するデータを提供する、就業者の年金保険ファイル。
- どれだけ期間個人が退職していたかに関するデータを提供する、社会保険機関の年金レジスタ。
- 人々が失業していた時期に関するデータを提供する、労働省管轄の失業レジスタ。
- センサス実施の年と実施時期に、異なる教育機関での学生の学習に関するデータを提供する、様々な学生レジスタ。

- 徴兵隊員に関するデータを提供する、国防省の軍隊レジスター。
- 教育に関するデータを提供する、フィンランド統計局による最終学歴・学位レジスター。

レジスター・ベースの人口センサスは、調査票ベースのセンサスよりもはるかに少ない人的資源の利用で済むので、予算の点からは経済的である。この新たな制度は、1990年人口センサスの費用を、大規模な調査票ベースのセンサスの約1/10にまで削減した。

## 2. 評価研究

しかし、一方のデータ収集体系から他のそれへの転換は、時系列にある種の断絶をもたらすのは避けられない。2つの異なる方法によって生産された統計は、相互にズレを持っており、利用時に解釈の問題をもたらす。これらの違いは、レジスター・ベースのセンサスの導入を最終的に決定する前に既に、1980年と1985年の人口センサス資料を使って、レジスター・ベースの統計と調査票ベースの統計を比較することによって、検討された。これらの検討は、レジスター・ベースの統計が十分に高い品質をもつことを示した。

しかし、2つの方法によって生産された統計の間には一定の差異があるため、1990年センサスの実施にあたって、それらの差異を研究するための集中的な評価研究に着手することが決まった。その目的は、行政レジスター統計の誤差と短所をさらに深く分析すること、そして標本によって、レジスター制度を用いては判定できない一定のセンサスデータ（たとえば、パートタイム雇用、人々の通勤方法、家業従事者に関するデータなど）を生産すること、であった。この評価研究は、調査票を用いて（母集団の2%の）標本からセンサス・データを収集し、これらのデータとレジスターから得られたデータとを比較することによって、行われた。

フィンランド統計局は、そのレジスター・ベースのセンサス制度を発展させ際に、データが以前の諸センサスからのものと通約できることを保証するために、以前の諸センサスで使われた概念と定義にできる限り近いものを保持しようと努めた。それにもかかわらず、いくらかはレジスター・ベースの制度はある程度まで、現実とはわずかに異なるかも知れない行政的な観点から諸現象を観察している。他方で、レジスター制度は常に、同じ属性をもつ現象と人々を、同じ方法で扱っている。これは、人々自身が質問されるときに、いつも保たれる形ではない。その場合には、調査対象である人々の主観的な観念、つまり記憶違いや偏った回答によって、その結果が影響されるのである。たとえば、自分自身を学生とみなしている人は、自分

表1  
登録データと調査票データにおける主な活動のタイプ別にみた15～74才の人口

主な活動のタイプ	登録データ	調査票データ	差	差の割合 (%)	同じ方法で分類され た人々の割合 (%)
労働力人口	2473300	2464400	8900	0	95
就業者	2332300	2315700	16600	1	94
失業者	141000	148800	-7800	-6	71
非労働力人口	1277900	1286800	-8900	0	97
学生	329100	359900	-30800	-9	87
年金受給者	756900	762700	-5800	1	95
徴集兵	28000	24000	4000	14	76
その他	163900	140200	23700	14	55
合計	3751200	3751200	0	0	91

表2  
登録データと調査票データにおける主な活動のタイプ別の15～74才の人口

調査票データにおける 主な活動のタイプ	登録データにおける主な活動のタイプ						合計 (%)	N
	就業者	失業者	学生	年金受給者	徴集兵	その他		
就業者	93.7	23.1	9.9	3.5	14.6	20.8	61.7	231500
失業者	1.0	71.3	1.6	0.3	3.9	10.6	4.0	148800
学生	2.3	1.9	86.7	0.3	4.6	8.6	9.6	359900
年金受給者	1.3	1.1	0.6	95.1	0.3	4.7	20.3	762700
徴集兵	0.0	0.4	0.1	0.0	76.1	0.5	0.6	24000
主婦	1.3	1.8	0.7	0.6	0.0	42.8	2.9	110100
その他	0.3	0.3	0.4	0.2	0.4	12.0	0.8	30100
合計 (%)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	
N	2332300	141000	329100	757000	28000	163900		3751200

は身分上学生であると述べたとしても、調査時に一時的に働いていたのであれば、労働力の一部分として分類されるべきである。

言い換えれば、レジスターと調査票による方法が異なる結果を生産する場合には、どの方法が正確な結果を生産しているかを語ることは、しばしば不可能である。なぜなら、問題にしている現象は、ある特定の 카테고리での解釈をこえては分類できないことがしばしばであり、それらの現象を分類する方法は、それらの現象をどのように見そして重視するかに依存するからである。

この評価研究は、主な人口センサスデータのすべてを、選出された標本に関する調査票の結果とを比較した。ここでは、個人の経済活動に関連する最も重要な結果だけをとりあげる。

### 2.1. 主な活動のタイプ

レジスター・データと調査票データを比較する評価研究は、表1に示すような人口の主な活動のタイプに関する結果を生み出した。

この表が示すところでは、全人口の91%は、レジスター統計でも調査票統計でも同じカテゴリーに分類された。また、レジスター・ベースの労働力人口のうち、その95%は調査票に基づく労働力人口に属し、レジスター・センサスで非労働力人口に分類された人々の90%は、それと同じように調査票統計において分類された。

この分類における最大の数量的な差異は、就業者と学生の数にあった。レジスター統計は、調査票統計よりも就業者の数を17,000人多く示したが、学生数は30,000人以上少なく示した。

表2が示すところでは、調査票によれば、1990年の就業者のうち、2.3%は学生であり、1.3%は退職者、1.3%は家におり、1%は失業者であった。他方、レジスターによれば、調査票のなかで自分を就業者だと述べている人々のうち、1.4%は実際のところ失業者であり、1.4%は学生、1.1%は退職者、1.5%は非労働力人口であった。

レジスター統計によれば、失業者数は、調査票ベースの統計においてよりも5.8%少なかった。他方、同じ形で分類された人々の%は、失業者の場合においてのほうが、就業者の場合においてよりもはるかに小さかった。これらの違いはおそらく、公的制度を通じては求職しておらず、したがって公的失業レジスターには載っていない失業者によって生じているのだろう。他方、レジスター統計は、一度は失業したが、現在は職を見つけた一定数の人々を含んでいるだろうが、こうした人々の登録事項は、レジスター上では更新されていない。

表 3

登録データと調査票データにおける産業部門別にみた就業者人口

産業部門	登録データ	調査票データ	差	差の割合 (%)	同じ方法で分類され た人々の割合 (%)
農林業 (A-B)	197600	182400	15200	8	77
鉱業 (C)	5100	5100	0	0	75
製造業 (D)	478500	473300	5200	1	90
エネルギー供給 (E)	24500	26500	-2000	-8	85
建設業 (F)	168900	175100	-6200	-4	83
商業 (G-H)	363700	370400	-6700	-2	88
運輸業 (I-J)	164600	165600	-1000	-1	89
金融業 (K-M)	246000	242200	3800	2	83
サービス業 (N-T)	630600	658000	-27400	-4	91



表4  
登録データと調査票データにおける産業部門別の就業人口

調査票データにおける 産業部門	登録データにおける産業部門											分類不能	非就業者	合計 (%)	N
	(A-B)	(C)	(D)	(E)	(F)	(G-H)	(I-J)	(K-M)	(N-T)						
農林業 (A-B)	76.7	2.0	0.3	0.4	0.6	0.3	0.8	0.3	0.2	5.3	16.3	7.4	182400		
鉱業と採石業 (C)	0.1	74.5	0.1	0.0	0.2	0.0	0.1	0.0	0.0	0.0	0.2	0.2	5100		
製造業 (D)	2.0	5.9	89.6	5.3	2.6	1.9	1.8	1.5	0.5	4.0	12.1	19.2	473300		
電気・ガス・水供給 (E)	0.1	2.0	0.4	85.3	0.5	0.0	0.1	0.2	0.1	0.4	0.9	1.1	26500		
建設業 (F)	0.9	7.8	0.9	1.6	83.1	0.8	1.1	1.4	0.3	12.0	8.6	7.1	175100		
商業 (G-H)	0.9	2.0	2.7	0.4	1.7	87.6	1.8	2.3	0.8	6.1	13.2	15.1	370400		
運輸業 (I-J)	1.2	3.9	0.3	0.8	1.6	0.7	88.9	0.5	0.2	2.3	4.8	67.0	165600		
金融・保険業 (K-M)	0.5	2.0	1.3	1.6	1.4	1.2	1.0	83.4	1.1	8.2	8.0	9.8	242200		
公務 (N-T)	5.7	0.0	0.7	1.6	2.1	1.1	1.0	3.9	90.8	28.1	28.8	26.7	658000		
分類不能 (X)	0.1	0.0	0.1	0.0	0.2	0.1	0.1	0.1	0.1	12.0	7.0	0.7	17600		
非就業者	11.8	3.9	3.5	3.7	6.2	6.3	3.4	6.3	5.9	21.8	0.0	5.8	143800		
合計 (%)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0			
N	197600	5100	478500	24500	168900	363700	164600	246000	630600	52700	127800		2460000		

調査票統計は、レジスター・ベースの人口センサスよりも学生数が9%多いことを明らかにした。レジスター統計によれば、調査票統計における学生の15%は就業者であった。その理由は疑いなく、学生が勉学の傍らで働くのは一般的なことからである。それなのに、学生たちは、他にある程度の収入のある職に就いていても、自分たち自身をただ単に学生とみなしている。しかし、センサス時に収入のある仕事をしているなら、レジスター上は就業者として定義されてしまう。

## 2.2. 産業部門

レジスター・ベースの人口センサス統計はまた、評価研究において、就業部門についても比較された。この分析(表3)によれば、就業者として分類された人々の91%は、両センサスにおいて、同じ主な部門のカテゴリーに分類された。

最大の数量的な差異は、サービス業(27,000人)と農林業(15,000人)において見られた。この後者の区分では、レジスター・ベースの統計は、調査票ベースの統計よりも8%多い就業者の数を示した。またしてもこれは、おそらく農業労働の本質のせいであろう。農業労働は季節性があるため、農業に従事する多くの人々は、その年の間のある時期に、他の形の収入を伴う職を捜し求める。このことが特に該当するのは女性である。女性たちは自分たちの農作業のほか、例えばサービス業でしばしば雇用されるのである。レジスター・ベースの定義では、職業を2つ以上兼ねる場合には、仕事は一般に、支払労働に優先権があるので、これらの人々は、レジスター上は農業労働者として分類されるが、調査票のうえでは、収入のある職に就いていると述べるのである。

従って、レジスター・ベースの方法は、調査票ベースの方法よりもサービス業従事者の数が4%少ないという結果を生んだ。これは、いくらかは、すでに触れた農業労働者によって説明できる。また、他の部門で働いている人々は、サービス業において第2の職をもっているかもしれない。

異なる部門区分に分類されたグループの大きさは、表4でさらに詳細に示されている。

## 3. 結論

概していうなら、以上の比較の結果は、以前の比較の結果にかなり一致したと結論づけうるだろう。人口の経済活動の場合においてのように、単位レベルでのレジスター・データと、労働力調査で収集されたデータとを比較することによって、

毎年、レジスター・ベースの統計と調査票ベースの統計の品質と通約性に関するチェックが変わらず保持されている。もし、レジスター・ベースの定義が調査票ベースの定義から非常に大きく逸脱しているように見えるのなら、必要に応じて、得られた結果に基づいてレジスター・ベースの定義を訂正したり修正することができる。

## 文献

- [1] Evaluation study of the 1990 census (1994), Statistics Finland, Population Census, Volume 9B
- [2] Pekka Myrskylä, Census by Questionnaire - Census by Registers and Administrative Records : The Experience of Finland, *Journal of Official Statistics*, Vol.7. No.4, 1991, pp.457-474, Statistics Sweden.
- [2] A.Laihonen and P.Myrskylä (1989). The Population and Housing Census : General Plan. The Central Statistical Office of Finland Reports No.1989:17, Helsinki.
- [3] C.Starck (1989). The 1985 Population and Housing Census, Reliability Study, Central Statistical Office of Finland, Studies No.157.

バララ・リッタは、フィンランド統計局の特別な計画職員である。彼女は、1975年から、フィンランド統計局と国家統計協会に尽くしてきた。彼女は、フィンランドのレジスター・ベース人口センサス制度の建設に責任をもつ計画職員の1人である。彼女は現在、レジスター・ベースの地域就業統計の生産に責任を持っている。

### 3 地理情報システム(GIS)とレジスター・ベースの 人口センサス

Riitta Harala、フィンランド統計局

Marja Tammilehto-Luode、フィンランド統計局

#### 1. 序

5年前には、地理情報システム（GIS: Geographical Information System）を使用している統計機関はそう多くはなかった。今日、ヨーロッパの統計機関のほとんどとその他地域の多くの国は、種々のGISを使って統計を生産し、改善している。

フィンランド統計局は、統計生産にGISを導入した最初の統計機関のひとつである。1990年のフィンランド人口センサスは、レジスターに全面的に基づいた2回目のセンサスであった。これら2つの要素は、フィンランドでは1990年代に統計の生産と配布に強い影響を与えてきた。

地域統計は、フィンランドでは、16世紀以降生産され利用されてきている。統計生産に持ち込まれた地理情報システム（GIS）とレジスター・ベースの人口センサスへは何が新しいのだろうか？人口センサスがレジスター・データを利用して実施されているフィンランドで、地理データ管理での統計生産の応用が、特に強力に開発されてきたのは何故だろうか？以下の例は、この主題について簡単な案内として役立つだろう。

フィンランドの人口居住地域は継続的に減少しており、住居はますます減少している地域に収斂してきている。この10年間に、人口減少は東部および北部フィンランドで最も大きかったが、中部地域での減少も規模が大きくなってきている。地図1は、居住のある1平方キロメートルのグリッド（碁盤目）正方形の助けを借りて下部地域単位の人口減を示しており、1980年の人口居住のあった正方形のどれだけ割合が、1981年から1995年間に人口減になっているかを説明している。

規定された場所—この場合には、1km×1kmのグリッド正方形—を伴った小さな統計単位の選択は、われわれが全地域でのこの現象を研究することを可能にする。グリッド正方形をもって、人口減少を全地域の空間的人口に対比することができるのである。

統計数が地方自治体においてだけ利用可能だとすると、人口減を描くことは、不可能でないにしても難しいだろう。例えば、居住用建築物の数の変化を描く数字は、新しい建築物がこの変化に与える影響を明らかにしはしない。たとえ無くなった建

建築物と新しい建築物の総数がそれぞれ別個に分かったとしても、新しい建築物が、無くなった建築物の代わりになったかのはどこかは明らかにならないだろう。地方自治体別の人口変化を通して人口減を記述することは、さらに難しいことが証明されるだろう。平均値は地方自治体のどの部分が移民でを失い、どの部分が獲得したかを明らかにしない。同様に、地方自治体から去った人口の地方自治体ごとの数字は、地方自治体内部の異なる地域での移民の流れについて何ごともしない。

グリッド正方形の助けをえての人口減の研究は、地域の空間的人口構造における真の変化についてだけでなく、この変化がどこで起こっているかについての正確な情報を与える。地図2はまた、地方自治体の中での正方形の位置を与える。そこでは、人口減の脅威は、最年少の住民が49歳以上である正方形を危機正方形（地域）と指定することによって研究されるだろう。

もし、再度、地方自治体ごとの総数だけしか利用できないなら、総人口にしめる49歳以上の人口の関連する割合が、地方自治体の活性全体について何らかの示唆を与えることができようが、それら49歳以上の人々がどこに住んでいるかを明らかにしない。例えば、地方自治体の中心に住んでいる人が移住する確率は、遠隔地域に住んでいる人が移住する確率よりはかなり小さい。また、地方自治体の中心の住居が空き家になることは、遠隔地よりも、はるかに起こりそうにない。

個人とその活動に関して、その時間と場所と結びついたデータを活用すれば、行政的な地域単位に基づく分析では観察できない発見をもたらす。人口減についての研究以外にも、グリッド正方形データの応用は、フィンランドでは、例えば、疾病の地域的発生（Karvonen et al. 1997）、失業（Muilu et al. 1998）、地方の産業構造（Raisanen et al. 1998）、地域的隔離（Vaattovaara 1998）そして移民（Kauppinen et al. 1997）の研究がある。

## 2. レジスターベース制度と地理コードづきのデータへ向かって

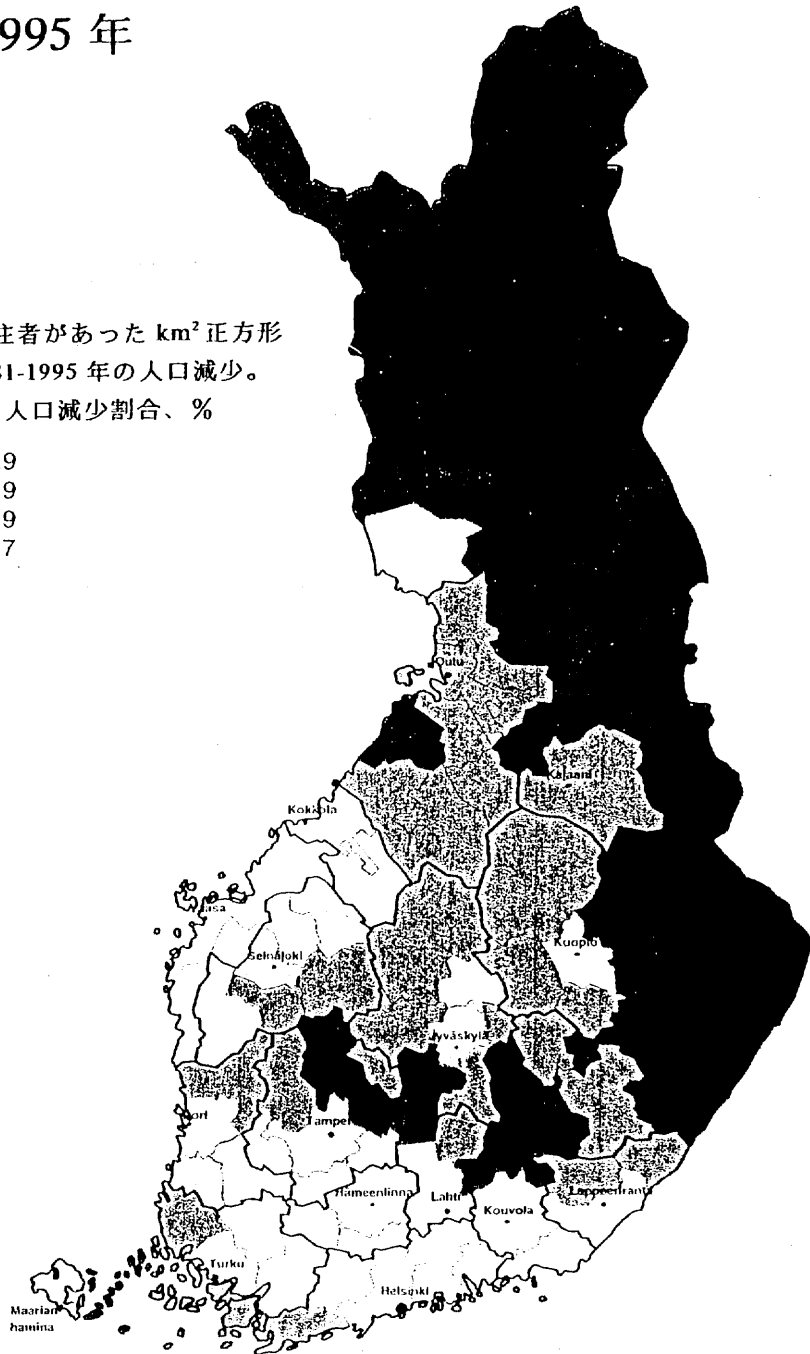
国際的にも、フィンランドでも、人口センサスは伝統的に、人々の活動を叙述する地域データの最も重要な—そして非常にひんぱんに唯一の—源泉になってきている。これが、社会・経済的地域データの生産と管理についてのどんな対話も、人口センサスの歴史とその前提の発展の歴史から提示すべき理由である。この前提は、今日、特にフィンランドでは、地域データの生産と多様なGISの方法論の利用に向けてユニークな機会を提供する。

人口センサスと人口統計の歴史はまた、われわれが今日行政レジスターのデータ

# 下部地域単位別人口減少 1981-1995年

1980年に居住者があった km<sup>2</sup> 正方形  
における 1981-1995年の人口減少。  
km<sup>2</sup> 正方形の人口減少割合、%

- 3.2 - 6.9
- 7.0 - 9.9
- 10.0 - 13.9
- 14.0 - 24.7



Source: Statistics Finland. Population statistics 1980 and 1995.

Statistics Finland

から毎年、レジスター・ベースのセンサス資料を編成するときに、われわれが、事実上、数世紀にわたる統計データの収集方法に戻る、言い換えれば、ひとまわりした、という事実を思い起こさせる。

## 2.1 人口センサスの歴史

フィンランドでは、人口センサスと人口レジスターは、数世紀にわたって互いに依存しあってきた。フィンランドがスウェーデン王国のひとつの県であった 16 世紀にさかのぼると、教区の当局は、租税と軍隊への徴兵目的のために、出生と死亡の記録を保有していた。例えば、土地レジスターや他の勘定書(accounts)もまた、この国の主要な産業である農業に従事している人についての詳細なデータをふくんでいた。16 世紀以降、地方行政当局の勘定書は、監査と全国的要約勘定書にふくめるため、ストックホルムに送られてきた。17 世紀には、これらのデータは中央政府の調査や計画作業で利用されていた (Luther 1993)。

1748 年には、「表局」(Tables Offices)がスウェーデン王国の最初のセンサスのために創立された。センサスのために、人口数、結婚、死亡、ジェンダー、婚姻上の地位や社会階級に関するデータが、教区や市長(city magistrate)から収集された。この収集は、1749 年、言い換えれば 2 世紀半前に行われた。

当初、フィンランドでは、センサスは不規則な間隔で行われたが、1880 年以降、10 年毎に行われた。1870 年からは、大都市の幾つかではまた、自身の手による補足的センサスを遂行した。

1917 年にフィンランドが独立した後、地方の人口レジスターに基づくセンサスが継続した。やがて、1938 年にセンサス方が成立したが、これによれば、「関連する建築物・住宅センサスを伴う人口センサスは、10 年毎に行われる」ものとされた。

フィンランドにおいて、センサス法に一般的に基づいた最初のセンサスは、1950 年に行われ、それ以来、10 年ごとに行われてきている。追加的な 2 つのセンサスが、1975 年と 1985 年に別個の法律に基づいて行われた。

## 2.2 人口センサスの方法の発展

1950 年と 1960 年のセンサスは、人口からのデータの直接的収集に依拠する伝統的な調査票センサスであった。行政データがはじめて利用されたのは 1938 年センサス法に対応して行われた 1970 年センサスであった (付録：フィンランドでの人口・住宅センサスでのレジスター・データの利用 1950-1995 年 参照) 中央人口レジスターは、宗教や市民権についての情報の源泉であったし、一方で所得データは税レジスターから得られた。1970 年センサスと関連して、最終学歴と学位のレジスターもまたフィンランドの中央統計局 (the Central Statistical Office of

# 1995年に最年少者が50歳以上である Pirkanmaaのkm<sup>2</sup>正方形



■ 最年少者が50歳以上である km<sup>2</sup>正方形

□ 最年少者が50歳未満である km<sup>2</sup>正方形

Source: Statistics Finland. Population statistics 1995.

Statistics Finland



Finland) (現在フィンランド統計局：Statistics Finland)に創設された。教育と学位についてセンサスで収集されたデータは、この新しいレジスターのための基本的データとなった。それ以降、このレジスターは、受験合格と獲得した学位についての教育機関からの年次的データによって更新されてきている。

そのデータ内容の点で、1975年人口センサス—これは住宅と経済活動サーベイとして知られている—は、完全な規模の人口センサスほどは包括的ではない直接質問サーベイであり、行政資料が広く利用された。年齢、ジェンダー、婚姻上の地位、ネイティブ言語と市民権についてのデータは、中央人口レジスターから獲得された。教育と学位についてのデータは、最終学歴と学位に関する中央統計局のレジスターから獲得された。

行政データ源泉として、そしてより優れた統計データの基礎としての両方から、建築物と住居についての全国レジスターの必要性が、1970年代に既に確認されていた。このレジスターは、1980年住宅・人口センサスと共同して構築されるべきことが決定された。このセンサスとともに、人口レジスター法に基づいて、中央人口レジスター調査票によって、住居と建築物についてのデータが収集された。建築物・住居レジスターは、人口レジスター・センターにおいて構築された。このセンターでは、それ以降、地方自治体の建築物担当部局と地方のレジスター部局からの最新のデータによって保持されてきた。後にこのレジスターは、その後の人口センサスのためと、その信頼性が1980年代半ばに改善されたうえで、年次レジスター・ベースの統計の生産に向けての、建築物・住宅データの基礎となった。

このレジスターは継続的に改善されており、1985年には、中間人口センサスが人口の経済活動（主な経済活動、仕事場、産業、所在地、職業）について、ただ1つの調査票を使うだけで実施できた。それにもかかわらず、なお調査票が全人口に対して送られざるをえなかった。

調査票ベースのセンサスが行われていた間に、行政データの利用だけに基づく人口センサスのモデルを開発する作業が1980年代に開始された。レジスター・ベースの人口センサスの、最も必要とされており、また技術的に最も難しい部分—言い換えると、人口の経済活動についてのレジスター・ベースの統計を生産するためのモデル—が完成して、最初のレジスター・ベースの地方就業統計が、1987年について生産することができた。この年以降、統計は同じシステムによって年次的に生産されてきた。このシステムを改訂するプロジェクトが現在、フィンランド統計局で始められている。この改訂は、労働市場の変化や統計的記述の変更の要求に応えるといった問題に関して、より深いものになるはずである。同時に、このシステムはまた、2001年の人口センサスでの必要という見地からも改訂される。

建築物・住居のストックについてのレジスター・ベースの統計もまた、1987年以

降毎年生産されてきた。

1990年人口・住宅センサスは、調査票を人口に送ることなしに、すべて行政レジスターを利用することだけで実施された。センサスにおいては、約30の異なる行政レジスター—その大部分は人口の経済的活動を述べている—からの資料が利用された。

1995年の人口センサスもまた、同じ生産モデルを使って行われた。このシステムは、2001年人口・住宅センサスでも使われる予定である。

### 2.3 レジスター資料

フィンランドのレジスター・ベースの人口センサス制度は、センサスが描くべき単位についての包括的データをふくむ基本的社会レジスターに基づいている。これらのレジスターには、人口レジスター、人口レジスター・センター人口情報システムのなかの建築物・住居レジスター、フィンランド統計局が保有するビジネス・レジスターがある。これらのレジスターは、すべての建築物と住居、フィンランドに所在する事業所とともに、フィンランドのすべての個人居住者を非常に包括的にカバーしている。すべての単位を、レジスターの識別制度を以て、互いにリンクさせることができる。すなわち、個人をその下で働いている使用主とともに、家族や世帯、住んでいる住居や建築物とリンクさせることができるのである。単位はまた、地図の座標を以て、地図上に位置づけることも可能である。

基本的レジスターの単位と特性データに加えて、約30の行政レジスターからの資料もまたフィンランドのレジスター・ベースの人口センサスには使われる。これらのレジスターは全人口あるいは人口の一部をカバーしている。使われている最も重要なレジスターには、納税レジスター（所得、所得のタイプ、使用主をふくむ）、雇用年金制度レジスター（雇用関係と使用主をふくむ）、求職レジスター（失業データをふくむ）、学生レジスターと国民年金庁レジスター、がある。利用されたより小さなレジスターは、例えば、非従軍奉仕活動を行った者、オーランド諸島の地方政府で働いた者の雇用関係、および海員の雇用関係についてのデータがある。

この制度の中心的な見地は、レジスターの識別制度が十分に優れていて、種々のデータの結合を促進できることである。レジスターは、重複したり、矛盾するデータをふくんでいる。レジスターが互いに補足しあい、そして重複が極端な弱点を少なくすること、が不可欠である。例えば、雇用関係に関するデータが納税レジスターから入手できないときには、それらは、雇用年金制度から獲得できる。

人口センサス・データは、レジスター推定を使いながら生産される。ここでは、変数の値は、幾つかのレジスターのデータを同時的に利用することによって、各単位について決定される。推論規則は、調査票を以て収集されたデータに可能な限り

近いデータを生産するように作りあげられている。この推論規則は、それ以前の人口センサスのデータを同じ時期のレジスター・ベースのデータとともに利用することによって作成された。この規則は、データが矛盾する場合に、種々の資料の優先順をもカバーする。調査票ベース制度との比較でのレジスター制度の長所の一つは、この制度が常に、同じ属性を持つ現象や人を同じ仕方で扱うことである。本人に送られる質問の場合には必ずしもそうではないからである。

### 3. 統計的地域とその境界

#### 3.1 地域ベースから、点ベースの統計制度へ

行政地域は伝統的に最も重要な統計的地域になってきている。行政地域に対する地域コードシステムは地域統計の基本であった。しかし、行政は動的であり、変化を続け、その地域の境界の変化をもたらししている。これらの変化に遅れないことは、最小の統計単位が時間の経過と共に変化する境界を持つ地域であるときには、難しい。

建築物はフィンランドでは、1970年人口センサスではじめて、地域統計の基本的単位として使われた。このことは、建築物の所在は、地図の座標を以て指定されなければならないことを意味した。地域と地域コードと並んで、地図座標、あるいは統計的単位の所在を正確に与える点が登場した。

したがって、フィンランドのレジスター・ベースの統計制度は、地域ベースではなく、点ベースである。フィンランドでは、建築物ベースの、座標を伴うコードシステムは、持続的に信頼をおけて柔軟な統計的地域のための優れた基礎を提供してきた。行政区域における大きな変化にもかかわらず、異なる地域ごとの時系列の計算もまた可能になる。

建築物の地図座標の採用は、統計的地域をより柔軟に決定することを可能にした。地図上の指定された地域内で、建築物座標とその中での現象を活用することで、「地域統計」を作り上げることが出来る。一般的な GIS パッケージの大部分を使えば、この多角形内の点の分析は、点（建築物）と多角形（統計的地域）に位置づけられるデータによって行うことが出来る。

#### 3.2 非行政的地域の発展

##### ローカリティ(*Localities*)と地方の結合

フィンランド全体を対象にした非常に早い 1950 年センサスでは、ローカリティは、行政地域とともに統計的地域として使われた。そのとき、少なくとも 500 人の住民の居る人口センターが自由な形の境界線で印をつけられた (*Taajamat/Localities*)

1995)。しかし、ローカリテイについての統計の歴史、そしてこれとともに、フィンランド統計局の地理情報の歴史は、1960年人口センサスにいたってはじめて開始した。このセンサスでは、同じ年に合意されたローカリテイーについての北歐共同定義が適用された。

北歐の定義によれば、少なくとも200人の住民を持つすべての建築物の集まりであり、建築物の間の距離が200メートルを越えないものが、ローカリテイーと定義される。当初の数例での狙いは、少なくとも可能なだけ、大きな地図上にそういった人口センターをおおよそ位置づけ、それらが地域のどれだけの大きさをカバーするのかを計算することであった。そのときはなお、ローカリテイーに関する実際の統計を生産することは可能ではなかった。このデータを積み上げることは、諸建築物の重点が座標を伴って決定された1970年になって可能になった。

ローカリテイーの定義は1960年以降変わらないままであったが、それらを境界づける方法は、データ技術発展のステップの中で変化した。地理情報ソフトウェアの発展が1990年に自動的なローカリテイーの境界付けを可能にした。これは、建築物群の端をつたう自由な形の線を作り出した。この線が創り出す多角形の出現は、建築物間の距離を計算する関数によって直接的に影響を受けることができ、与えられた指定に従って、人口の塊を形成した。

1998年には、主として地方研究の必要におされて、ローカリテイーを境界づける方法が200人以下の住民を持つ地方人口の塊をも対象とするものに、さらに発展した。1998年に行われた境界付けは、30人から500人の地方の塊に焦点をおき、これらについての統計の生産もまた現在進行中である。

#### 地方自治体の下部地域(Municipal sub-areas)

1970年人口センサス以降、地方自治体は、レジスターベースの統計がそのために生産される下部地域を定義する機会を得た。しかし、この地方自治体の下部地域の導入は、地方自治体の自由意思に任されてきており、その目的は標準化された大きさの地域を創り出そうとするものではなかった。この作業は個別の地方自治体みずからの必要から生まれた。

フィンランド統計局は地方自治体に対して、その地方自治体の下部地域区分の定義と/あるいは更新、および関連する名前のリストを問い合わせる手紙を毎年送っている。地方自治体は下部地域を地図上に描くか、その境界を数的な形で提出することを求められている。フィンランド統計局は、境界のデータを関連する地域コードで構成し、地理データ管理を使って要求された統計を生産している。

地方自治体の下部地域制度は、フィンランドのほぼすべての地方自治体で採用されている。幾つかの地方自治体では、異なる行政機関の統計的必要のために、幾つ

もの異なった下部地域の区分を適用している。

#### 郵便コード地域

コードと座標に加えて、フィンランド統計局ファイルのすべての建築物もまた住所を持っている。この住所は5桁の郵便コードである。1980年代以降、このコードは郵便コード地域ごとの統計の集計を促進してきた。郵便コード地域ごとの統計への需要は、1990年代に入るまで取り上げられなかった。主な顧客は、マーケティングに関わり、特に市場分析を行っている企業である。地理情報システムに対して郵便コード地域の境界を作ることもまた次第に必要なようになってきた。

郵便コードの地域境界は、建築物の位置データと郵便コードを同時に活用することによって形成された。同じ郵便コードを共有する建築物の周りに多角形が描かれ、この多角形は地域ネットワークに加えられた。郵便コードの境界は毎年更新されている。

#### グリッドと他の柔軟な地域

グリッド・データは通常、正規のグリッド正方形が統計的地域をなしている人口、労働力、就業についての統計に関わる。正方形の所在は地図の座標に点で打たれる。もっとも利用される正方形の大きさは1km×1kmである。

グリッド正方形によるデータは、1970年人口センサス以降、フィンランド統計局から入手可能である。指令にしたがって、建築物の中心座標を使った関連するレジスター・ベースの統計データが、座標に付け加えられてきている。建築物の座標によって定義可能な単位指定のデータが、グリッド正方形にまとめられている。顧客は、自分で自由に使える統計を受け取っているが、その統計は、グリッド正方形座標（底辺の左側）が地域コードをあらわしている。

グリッド正方形データは、もっとも柔軟な統計的地理データとみなされている。正方形は場所的には安定している。それは、行政地域のように、年によって移動することはない。したがって、それはどのような地域の変更とも無関係である。小地域によるデータは、より大きな地域に関連させてまとめることは容易である。

地方自治体の下部地域とともに、顧客が指定した他の地域についての統計を生産することも可能になっている。データの安全性にしかるべき配慮をしながら、フィンランド統計局は、たとえば、地図上に指定された計画地域あるいは自動車道からある距離に位置する地域についての人口データをまとめるのである。

## 4. GIS とフィンランド統計局の統計

1990年代初期に起こった地理情報技術の大きな前進によって、フィンランド統計

局は 1992 年に自身の地理情報プロジェクトをたてた。建築物を位置付けることができるレジスター・ベースのデータと、地図データの急速に進展しているデジタル化は、地理情報システムの一層の発展に対して非常に豊かな基礎であることが示された。このプロジェクトの目標は、フィンランド統計局の地理データ業務の合理化と地理データ生産物の開発であった。

プロジェクトが前進するとともに、顧客サービスにおいて、フィンランド統計局の膨大な地理情報データの貯えの、柔軟な利用を促進する体系的発展と保守が必要とされた。1996 年にこのプロジェクトは、人口統計課にその 1 セクションとして統合され、生産、保守、開発の課題がゆだねられた。今日、このセクションは 6 人のフルタイム職員を持ち、その課題は次のものである。

- 地理情報（統計と地図データ）の保守と開発
- 地域分類の保守と開発
- 主題別地図の生産
- 地理情報と地理情報技術を必要とする委託作業
- 生産物開発

顧客からの委託には、地理情報セクションでの労働時間の約 3 分の 1 があてられている。顧客は、標準化された生産物のほかに、個別的な注文に対応した編集物、地理データ管理システムを利用して改作されたレジスター・ベースの統計、数的統計地域境界の提供を受ける。顧客は、行政機関、科学研究者、そしてまた増加しつつある民間部門である。産業界の GIS は、フィンランドでは次第に地歩を築きつつある。マーケティングや市場分析に関わっている卸・小売業、銀行・保険機関、企業と特に電信電話会社が、自らの地理情報システムへの応用のために、フィンランド統計局のデータ資料について問い合わせをしている。

通常は、地理データセクションでの労働時間の 3 分の 1 が生産物開発にあてられている。開発された新しい生産物には、特に TeemaCD、地図上の地域統計がある。TeemaCD は、異なる地域ごとの統計のデータベース、それら地域を描写している地図データ、そして地域別データを速やか、かつ容易にブラウザーし、検索し、比較し、視覚化するアプリケーションをふくむものである。TeemaCD は革新的な統計的生産物であり、極端に膨大な統計を速やかに説明的に取り扱うことができる。

1999 年の生産物開発プロジェクトは、インターネット上での地図をベースにした選挙結果サービス（1999 年にはフィンランドでは 2 つの選挙が行われることになっている。すなわち、国会議員選挙と欧州議会へのフィンランド議員の選挙）と、地方研究に役立つフィンランド人口の統合に関する統計体系、である。

地理データを更新し保守することは当然にまた、地理データセクションの労働時間のある量をとる。同じように、地理データの貯えを柔軟に利用することも、継続

的な開発作業への投入を必要とする。統合的地理データベースに向けての発展は、グリッド正方形のデータベースを一層発展させることで開始された。その目的は最もひんぱんに求めがある社会・経済的地理データを、クロスセクションであれ、時系列であれ、フィンランドのどの部分についてでも、異なる大きさも正方形ごとに速やか、かつ容易に検索可能にすることである。

## 5. GIS とレジスターベースのセンサス—課題と機会

フィンランドでは、地理コードを付されたレジスター・ベースのデータが利用可能であるお陰で、社会・経済的地理データを活用する機会は、例外的に優れている。統計の処理過程での地理情報システム技術の利用は、したがって、統計生産と、そして特に統計を改善して市場化可能な生産物にする上での、基本的要素になる。しかし、地理情報システムの利用は、データの安全性、データの品質、問題別の地図化/視覚化のノウハウ、そして著作権といった問題に関連した幾つかの大きな問題、あるいは課題を生み出している。人口減の研究についてふれた序論での例で、われわれは結果を公表する以前にそれらすべての課題に直面した。

地理情報システムにおいて人間活動を描くために使われている地域単位はしばしば、可能なだけ小さくなっているために、単位データの識別が決定的問題になる。地図上に地域の所在を指定すること自体がまた識別を一層容易にする。標準化されたデータ保護の方法を使うことも問題がないわけでない。より大きな地域内でのデータの安全性が危険である地域が、あれこれの方法で処理されたとしても、小さな地域上の統計資料が何らかのより大きな地域レベルで統計的に信頼性をおけるなら、研究者にとっては十分であるかもしれない。しかし、配置計画をする者は、資料のどの部分においても、間違った。あるいは平均的数字を受け入れるわけにはいかない。彼/彼女は、現象の正確な場所についてのデータを必要とするからである。センシティブな地理データを、顧客がそこから必要な情報を獲得できるような形に処理することは、統計機関にとっては、大きな課題になる。

地理情報システムにおける小地域統計の利用はまた、より大きな地理的行政単位に基づいた伝統的統計が明らかにしないデータの品質問題を明るみに出す。小地域では示されていない誤差が、全資料に信頼できないものとの烙印をおす。しかし、地理情報システムは、地域統計—例えば、座標に従うと水域、あるいは、建築物コードが意味している異なる地方自治体に所在する建築物—の誤差をチェックする新しい道具を用意した。これまでは、それらの道具は、フィンランド統計局ではまだ十分活用されていない。

地図を伴った地理データの視覚化は、地理情報システムの重要な要素である。う

まく生産された問題別の統計的地図は、検討されている現象の地域的相違の主な特徴を速やかに示す。正しい表示の尺度を選ぶこともまた、データの安全問題に対するひとつの解決であることを示す。対照的に、貧弱な問題別地図は、現象についての全く誤った描写を与え、そして/あるいは、データの信頼性と安全性についての論争をさらに燃え立たせるだけになろう。フィンランド統計局は、統計的問題別地図の編集において基礎的指示を提供してきた。協力的な国際的プロジェクトは、また模範となる問題別地図（統計的問題別地図 1998）を編集する方法についての一連のガイドラインをまとめている最中である。

数的な地図データは、地理情報システムにおいて重要な役割を果たす。数的地図データは、地理データの視覚化にとって必要なことに加えて、例えば、道路からあるいはサービスからの距離といった新しい変数を形作るためと、バッファリング (buffering)、ならし、距離による分類、地域による指数といった資料の処理や分析のための新しい方法を導入するためにもまた利用できる。フィンランドと大部分のヨーロッパ諸国でもまた、数的地図データは高価であり、非常にきびしい著作権規制で保護されている。数的地図データの広い利用が、その価格によって制約されているのは明らかである。両方の当事者が利益を受ける可能性は、統計機関と地図化をする機関との間の協力において見出すことができよう。地図データはまた、多様な行政レジスターと比較しうる、基本的な統計資料（土地利用統計）とみなすことができる。これまでのところ、この目的のための利用は相対的には限られている。

社会・経済的研究での地理データの利用は継続的に増加してきている。地理データは行政地域単位ごとに生産される統計を補足し、行政的—そして通常は地理的に大きな—地域単位（参照 序論における人口減少の研究の例）を利用することでは見出されない隠れた情報を明らかにする。現在まで、研究は統計的方法に主として焦点をあてているが、地理情報システムの実際の分析方法は、わずかしか利用されていない。さらに、地理データ管理技術と場所的統計を結合することも、十分に活用されていない。例えば、標本を設計することと、標本ベースの統計処理である。

建築物の位置座標のうえに打ち立てられたレジスター・ベースの統計システムは、地理情報システムで統計利用に対して汲み尽くせない可能性を創り出す。これらの可能性を活用することは、実際にはフィンランドで今始まったばかりである。地理データの利用はまた将来において、大変有利であるとみなされうることなので、データの安全、著作権、価格、品質からくる課題に対する解答を見出すのは、困難でないといえることが、望まれる。



## 文献

- Guideline for statistical thematic mapping(1998), Working paper No.27, Work Session on Geographic Information Systems. Ottawa, Canada,5-7 October 1998, Statistical Commission and Economic Commission for Europe, Conference of European Statistician
- Harala, Riitta (1995). Continuous Quality check of register-based employment statistics. European Workshop on Using Administrative Data in Population and Housing Censuses, Statistics Finland and Eurostat, Helsinki 1995
- Harala, Riitta (1995). Evaluation of the results of the 1990 register-based population and housing census in Finland, Statistical Journal of the United Nations ECE(1995)63-72,IOS Presss
- Harala, Riitta (1998). Statistical Properties and Quality of Register-based Census Statistics in Finland, Symposium '97, New Directions in Surveys and Censuses, Proceedings, Statistics Canada, Ottawa 1998
- Karvonen, M.,Rusanen,J., Sundberg,M.,Colpaert, A., Naukkarinen, A.,Tuomilehto. J., for the DiMe Study Group(1997), Regional differences in the incidence of insuline-dependent diabetes mellitus (IDDM) in Finland during 1987-1991. *Annales Mediein*, 29,297-304
- Kauppinen, J., Rissanen, E., Rusanen,J., Naukkarinen, A., Muilu, T., Colpaert, A.(1997),Migration As a function of population, *Nordia Geographical Publications*,26,17-27
- Luther, Georg(1993), Suomen tilastotoimen historia vuoteen 1970( History of Finnish officail statistics up to 1970, in Finnish only), Statistics Finkland,Helsinki,1993
- Rusanen,J.,Naukkarinen,A.,Colpaert,A., Muilu,T.(1997),Differences in the spatial structure of the population between Finland and Sweden in 1995, Research reports 221, Statistics Finland
- Raisanen,S., Rusanen.,J., Naukkarinen,A.(1996) Socio-economic grid data and GIS for analysing the Changes in the Finnish countryside, In Second Joint European Conference & Exhibition on Geographical Information, March 27-29,1996, Barcelona, Spain. *Geographic Information From Research to Application through Co-operation*,Vol.1,651-660
- Suomi keskittyy ja autioituu (1997). Muuttoliike alueittain 1985-1996, Suomen Kuntaliitto. Helsinki ( Finland centralises and depopulates 1997. Migration by area 1985-1996. The Association of Finnish Local and Regional Authorities. Helsinki, in Finnish only)
- Taajamat 1995, Vaestolaskenta 1995 Osa 4.( Population census 1995, Localities 1995, Part 4, in Finnish only) Statistics Finland, Helsinki 1997)
- Tammilehto-Luode,M.(1993). Paikkatiedon hallinta Tilastokeskuksessa. Suomen tilastoseuran Vuosikirja ( Geographical information management at Statistics Finland, Yaebook of the Finnish Statistical Society, in Finnish only)

- Tammilehto-Luode, M. (1995). Geographical Information Systems. European Workshop on Census Geography, 26-27, June 1995, Vienna, Austria
- Tammilehto-Luode, M. (1998). Setting priorities for regional statistics in Finland. SCORUS'98, 8-11 June 1998, Belfast, Northern Ireland
- Tammilehto-Luode, M. (1998). Rural villages as statistical areas. ECE Work Session on Geographic Information Systems 5-7 October 1998, Ottawa, Canada
- Vaattovaara, M. (1998), Paakaupungin sosiaalinen erilaistuminen. ( Residential differentiation within The metropolitan area of Helsinki, in Finnish with abstract in English). Tutkimuksia ( Studies) 1998:7 Helsingin kaupungin tietokeskus (City of Helsinki Information Centre). Helsinki

付録 フィンランドの人口・住宅センサスにおけるレジスターデータの利用  
1950-1995年

項目	1950	1960	1970	1975	1980	1985	1990	1995
<u>人口データ</u>								
年齢(出生月日)	q	q	q	R	R	R	R	R
性	q	q	q	R	R	R	R	R
婚姻上の地位	q	q	q	R	R	R	R	R
言語	q	q	q	R	R	R	R	R
市民権	q	q	—	R	R	R	R	R
宗教	q	q	R	—	R	R	R	R
通常時の住居	q	q	q	q	q	R	R	R
<u>世帯・家族データ</u>								
世帯の大きさタイプ	q	q	q	q	R	R	R	R
家族の大きさタイプ	q	q	q	q	R	R	R	R
<u>経済的データ</u>								
主な活動	q	q	q	q	q	qr	R	R
従業上の地位	q	q	q	q	q	qr	R	R
産業	q	q	q	q	q	qr	R	R
職業	q	q	q	q	q	qr	R	R
使用主、仕事場	q	q	q	q	q	qr	R	R
社会・経済グループ	q	q	q	q	q	qr	R	R
所得	—	—	R	R	R	R	R	R
<u>最終学歴</u>								
最終学歴	q	q	q	R	R	R	R	R
<u>住居データ</u>								
住居単位の大きさ	—	q	q	q	q	R	R	R
貸家状況	q	q	q	q	q	R	R	R
<u>事業店舗データ</u>								
借り主の産業	—	q	q	—	q	—	R	R

## 4 標本調査における推定を改善するための

### レジスター・データの利用

－ケーススタディとしてのフィンランド労働力調査

Risto Lehtonen フィンランド統計局

Ari Veijanen フィンランド統計局

キーワードとフレーズ：行政レジスター、補助的情報、一般化回帰推定、労働力調査、モンテカルロ実験、社会統計

#### 1. 序

行政およびその他のレジスターから得られる補助的データは、標本調査における推定を改善するために国家統計機関によって頻繁に利用される。レジスター・ベースの補助的データの例としては、センサスからあるいは一国の全人口をカバーしている人口レジスターから得られる人口学的データ、税務当局のレジスターから得られる納税記録、あるいは職業紹介所のレジスターによって失業中の求職者としての地位を示す記録などがある。総人口についての推定手続きにおいてレジスター・データを利用することから期待される利点は、例えば、効率の改善、そして単位無回答による偏りの削減である (Särndal *et al.* 1992)。

推定手続きにおいて、レジスター・データを組み入れる代替的テクニックがある。例としては、事後階層化 (post-staratification) (Holt and Smith 1979)、比率と回帰推定 (Särndal *et al.* 1992)、そしてレイキング (raking) とキャリブレーション (calibration) 手続き (Deville and Särndal 1992 ; Deville *et al.* 1993) である。これら推定手続きの多くは、モデルに支援された推定の傘下であり、そしてより深く Särndal *et al.* (1992) で議論されている一般化回帰 (GREG) 推定量の特殊ケースと考えることができる。この手続きは、カナダ統計局の GES (Estevao *et al.* 1995)、スウェーデン統計局の CLAN (Anderson and nordberg 1998)、オランダ統計局の Bascula (Renssen *et al.* 1997)、そして INSEE によって発展された Calmar など、いくつかのソフトウェアによって、利用者が利用可能になってきている。

2 項(binary)もしくは多項(polytomous)調査変数の階級度数の推定は、調査の実践において共通して突き当たる推定問題である。被雇用者、失業者、労働力外にある者の総数の推定は、労働力調査における典型的な例である。この論文では、推定問

題は新たな推定手続き、Lehtonen and Veijanen(1998a)が導入した多項ロジスティック一般化回帰推定[multinomial logistic generalized regression estimation](LGREG)、によって考察される。職業紹介所の行政レジスターからの補助的データが、LGREG推定の利点を示すのに効果的に利用される。

この論文は以下のように構成される。統計目的、とりわけ社会統計のためのレジスター・データの利用の概観が第2章に示される。レジスター・データを用いるある推定手続きの概要が第3章に示される。フィンランドにおける労働力調査からとったケーススタディが第4章で示され、議論されている。要約は5章に示される。

## 2. 統計目的のためのレジスター・データ

行政レジスターのカバレッジと品質、そしてそれらの利用可能性と統計目的のための利用は、異なる国の統計機関の間で非常に異なる。これは政府統計の生産における伝統の違いによることがある。ひとつの重要な理由は、利用可能性のための法的基礎や、統計目的のための行政記録の利用は異なる国々において異なっているということであろう。また、政府統計のために行政的データ源を使うことに対して、国家統計機関の間での文化的やその他の違いもありうるだろう。EU 諸国についての現在の状況や異なるアプローチについての概観は Grais (1996) と欧州委員会 (1997) によって与えられている。

上で述べたような違いにもかかわらず、行政およびその他のレジスターの統計目的での利用は、将来的に増加していくだろうと考えることができる。行政レジスターの重要性と潜在的な有用性は、例えば Franchet(1997)と Defays(1997)が示したように、EU 内において大きく強調されている。このトピックに関して最近いくつかの国際会議が Eurostat-EU 統計局によって組織された (Grais 1996、欧州委員会 1997)。

フィンランド統計局は、行政およびその他の源泉に基づいて、包括的で高い品質の統計レジスター制度を維持している (Myrskylä 1991、Hahara 1995 and 1997、Vihavainen 1996)。そこには政府統計の編集のためにレジスター・データを利用してきた長い歴史がある。例えば、フィンランド統計局が現在生産している政府統計の 90 パーセント以上はレジスターの情報源に基づいており、レジスター・ベースのセンサスが 1990 年、1995 年に行われた (次の調査は 2001 年に実施される)。1994 年に制定された統計法は異なる省庁が保有している行政レジスターを、統計目的のために利用することを可能にした。個人のプライバシー保護に関する法規、マイクロデータの収集と利用に関する倫理的規則と機関の標準、データの配布と政府統計の出版において用いられるデータ開示の技術についての法規に加えて、この法律は、

重要なデータの秘匿性問題を効果的に規制している。統計目的にレジスター源を有効に使うための文化的、法律的、技術的基礎は、このようにうまく設置されているのである。

フィンランドの政府統計システムの多くの特性はまた、その他の北欧の統計局にもあてはまる。スウェーデンについては Wallgren and Wallgren(1997)、デンマークについては Spieker(1997)、ノルウェーについては Gåsemyr(1996)と Thomsen and Kleive Holmøi(1998)を参照せよ。

北欧諸国の国家統計機関のもうひとつの特性は、サーヴェイとレジスターの両方で唯一の個人識別番号を用いて、マイクロレベルでレジスター・データとサーヴェイデータとを合併させることが技術的に容易なことである。このことは、例えば社会統計の生産のために、レジスター・データとサーヴェイ・データを結合して利用することに十分可能にする。

さらに加えて、マイクロレベルでレジスター・データとサーヴェイ・データを合併させることが容易なので、一般化回帰推定のような、進んだ推定手続きの効果的な利用が可能になる。

### 3. 一般化回帰推定

#### 3.1 背景

一般化回帰推定は、推定手続きにレジスター・データを組み入れるための強力な手段を与える。フィンランド労働力調査との関連で、雇用者と失業者の総数の推定を考えてみよう。Djerf(1997)と Lehtonen and Veijanen(1998a,b)は、もし補助変数が現在検討中の調査の変数との間に強い相関があるなら、その利得は大きいものでありうることを示した。これは例えば、ある者の失業に関する2項調査変数と、あるものが失業中の求職者として登録されているかどうかを示す2項補助変数とに適用できる。失業に関する調査変数は、国際労働機構(ILO)による標準的定義に基づいており、補助変数は、労働省が保有している失業中の求職者の記録(雇用局統計)からの行政記録を用いて構成されている。

労働市場での個人の地位(就業者、失業者、あるいは非労働力)の様な多項調査変数の階級別度数の推定について、推定問題は以下のように定式化される(Lehtonen and Veijanen 1998 a)。1つの階級における個人の数は、母集団全体にわたる階級の指標の合計、すなわち指標の総計に等しい。したがって、この問題は、母集団総数の推定のために作成されている方法を用いて解決できる。

母集団総数の推定において利用可能な補助的データを利用するためには、標準化一般化回帰推定量(Särndai *et al.* 1992)が連続的調査変数に対して共通して用いら

れる。連続変数については、調査変数の期待値は補助変数に線形的に依存すると想定することができる。

線形モデルは、2項変数、もしくはより一般的に多項変数にとっては、最善の選択ではないかもしれない。そういった変数として、Lehtonen and Veijanen(1998 a)は、多項ロジステックモデルに基づくロジステック一般化回帰推定量を導入した。この特殊なモデルのタイプを選択する動機は、このように一般化線形モデルとの関連で用いられているものに類似している (McCullagh and Nelder 1989)。

Lehtonen and Veijanen(1998 a)は、労働力調査データから作成された人口から標本が繰り返し抽出されるモンテ・カルロ・シミュレーション法によって、一般化回帰推定量の特性を研究した。以下では、Lehtonen and Veijanen(1998 a)の主要な研究結果を要約する。

### 3.2 多項調査変数についての一般化回帰推定

有限母集団  $U$  における  $N$  個の要素  $k$  と関連した多項  $m$  値ランダム変数  $Y_k$ 、例としては、労働市場状態 (就業者、失業者、非労働力) を考えてみよう。それらの実現された変数  $y_k$  は、大きさ  $n$  の標本  $s \subset U$  において観察される。目的は母集団における  $y_k$  の度数分布の推定である。このために、階級の指標  $Z_{ki} = I\{Y_k = i\}$ 、その実現値  $z_{ki}$  と定義しよう。したがって、ある与えられた人について、雇用されている場合には第一の指標について、失業している場合には第二の指標に、非労働力の場合には第三の指標に値 1 (他の場合にはゼロ) が割り当てられる。

Horvitz-Thompson (HT) の推定量  $\hat{t}_i^{HT} = \sum_{k \in s} a_k z_{ki}$  は、ここでサンプリングウェイトは  $a_k = 1/\pi_k$  すなわち包含確率 (inclusion probability) の逆数であるが、階級の総計  $t_i = \sum_{k \in U} z_{ki}$  についての最も単純な推定量を与える。明らかに、おそらくウェイトを除いて、HT 推定量はいかなる補助的情報も組み込んでいない。他方で、一般化回帰推定量は、望まれる補助的データを明らかに組み込んでいる。これは、適切なモデルを利用することによって成し遂げられる。

文献に示されているモデルの選定には 2 つの主要な選択がある。すなわち線形モデル (Sändal et al. 1992、「標準」GREG 推定量を提出している)、そして多項ロジステックモデルのような非線形モデル (Lehtonen and Veijanen 1998 a、ロジステック LGREG 推定量を示している) である。

### 3.2.1 GREG 推定手続き

補助変数  $x_k$  のベクトルが母集団のすべての要素  $k$  について既知のものとする。補助変数の例は、性別、年齢階級、居住地域であり、2項変数は人物が失業中の求職者として登録されているかどうかである。標準一般化回帰 (GREG) 推定量は、

$Var(\varepsilon_{ki}) = \sigma_{ki}^2$  を持つ線形回帰モデル  $Z_{ki} = x_k' \beta_i^G + \varepsilon_{ki}$  によって支援される (Särndal *et al.* 1992、Estevao *et al.* 1995)。パラメータ  $\beta_i^G$  は、等式

$$\hat{\beta}_i^G = \left( \sum_{k \in U} a_k \frac{x_k x_k'}{\sigma_{ki}^2} \right)^{-1} \left( \sum_{k \in U} a_k \frac{x_k z_{ki}}{\sigma_{ki}^2} \right) \quad (i=1, 2, \dots, m), \quad (1)$$

を用いた加重最小二乗法によって推定されている。ここで、ウェイト  $a_k = 1/(\pi_k \theta_k)$  は、無回答について調整されたサンプリングウェイトである。実際には、推定された回答確率  $\hat{\theta}_k$  が用いられている。ここで考慮される GREG 推定量について、(1)

における分散は定数である、すなわち  $\sigma_{ki}^2 = \sigma^2$  である、と想定される。

GREG 推定量を得るために、適応する値  $\hat{z}_{ki} = x_k' \hat{\beta}_i^G$  が

$$\hat{z}_{ki}^G = \sum_{k \in U} \hat{z}_{ki} + \sum_{k \in U} a_k (z_{ki} - \hat{z}_{ki}) \quad (i=1, 2, \dots, m), \quad (2)$$

で与えられる階級の総計  $t_i$  の推定量に取り入れられる。

GREG 推定量(2)についての線形モデルの選定は、連続的な調査変数については十分正当化される。2項もしくは多項変数については、線形モデルは非現実的なものであるかもしれない。通常は、線形モデルよりもロジステックモデルが好まれるだろう。

### 3.2.2 LGREG 推定手続き

LGREG 推定手続きにおいて、期待値  $\hat{\mu}_{ki} = E(Z_{ki} | x_k; \beta) = P\{Y_k = i | x_k; \beta\}$  の推定量  $\hat{\mu}_{ki}$  は補助変数

$$\hat{\mu}_{ki} = \frac{\exp\{x_k' \hat{\beta}_i\}}{1 + \sum_{r=2}^m \exp \hat{\beta}_r} \quad (i=1, 2, \dots, m), \quad (3)$$

に非線形的に依存する。

ロジステックモデルのパラメータは、擬尤度 (pseudolikelihood) 法 (Godambe and



Thompson 1986; Nordberg 1986; Särndal *et al.* 1992 p. 517) によって推定されている。一般的に、尤度関数は Newton-Raphson 演算法のような適切な数的方法によって、数的に最大化されている。代わりに、一般化（加重）最小二乗（Lehtonen and Pahkinen 1996）法が、あるロジステック ANOVA タイプモデルにふさわしい計算的にはわずかしが要求しないテクニックを提供しながら利用することができた。

ロジステックモデルの定式化においては、予期される値はいつも「0、1」に存在するという事に注意するべきである。これに対して、線形モデルの定式化では、予期される値がこれらの自然な境界を越えることがあり得る。 $Z_{ki} = 1$ の確率が0もしくは1に極めて近い場合、この2つのモデルは異なる結果をもたらす。そしてさらに、多項調査変数について2つ以上の階級が存在する場合には、多項ロジステックモデルによって指標変数の結合分布を叙述するのはわかりやすい。

階級総計のロジステック一般化回帰（LGREG）推定量は、

$$\hat{i}_i = \sum_{k \in U} \hat{\mu}_k + \sum_{k \in S} a_k (z_{ki} - \hat{\mu}_k) \quad (i=1, 2, \dots, m), \quad (4)$$

によって定義される。

飽和した ANOVA モデルに対応する完全な事後階層化については、GREG 推定量 (2) に適合した変数  $\hat{z}_k$  は LGREG 推定量 (4) における推定値  $\hat{\mu}_k$  に等しいことを示すことができる。このようにして、完全な事後階層化において欠けたセルがない場合、GREG 推定量と LGREG 推定量は同一である (Lehtonen and Veijanen(1998 b)。これは、不完全な事後階層化、あるいは交互作用項のいくつかもしくはすべてが取り除かれる ANOVA モデルに対応するレイキングといった他のモデルに対しては、あてはまらない。

GREG 推定量(2)と LGREG 推定量値(4) は、母集団にわたっての予期値の合計を含んでいる。しかし、母集団の各要素についての補助変数に関する情報を持つことは、実際には必要ではない。GREG においては、補助の総数  $\sum_{k \in U} x_k$  を知ることで

十分である。なぜなら、(2) は  $\hat{i}_i^G = \hat{i}_i^{HT} + \left( \sum_{k \in U} x_k - \sum_{k \in S} a_k x_k \right) \hat{\beta}_i^G$  という形でも

また表し得るからである。完全な事後階層化という特殊な場合に関しては、LGREG において必要とされる情報は GREG で必要とされるものと類似している。不完全な事後階層化のようなその他の場合については、(4)の予期値の総計  $\sum_{k \in U} \hat{\mu}_k$  は、母集団における  $x_k$  のそれぞれの値の度数を知ることなしには計算できない。例えば、2つの離散型補助変数について、GREG では周辺度数が必要とされ、LGREG ではセ

ル度数が必要とされる。したがって、この点において、LGREG は GREG よりもより必要なのである。

補助的指標変数が LGREG 推定手続きにふくまれているなら、母集団にわたってのその総計は、LGREG 推定量によって正確に推定される。標準 GREG 推定量と共有の、このキャリブレーション特性は、多くの応用において望ましいものである。

失業率の合成比率推定量  $\hat{R} = \hat{i}_2 / (\hat{i}_1 + \hat{i}_2)$  は第 4 章の実験において応用されている。ここでは、 $\hat{i}_1$  は就業者の推定総数、 $\hat{i}_2$  は失業者の推定総数をあらわしている。 $\hat{R}$  の分散はテイラーの線形化技法 (Lehtonen and Veijanen 1998 a 参照) によって推定される。この第 1 次のテイラーの拡張は、余りにも小さな分散推定値を提供する傾向をもつように見える。この問題は、ブートストラップ (bootstrap) のような再サンプリング法を利用することで、おそらくは緩和できるだろう。

### 3.2.3 領域推定

母集団全体の推定値に加えて、推定値はたいてい部分母集団あるいは領域についても計算される。このために、母集団  $U$  はサイズ  $N_{(d)}$  の領域  $U_{(d)} \subset U$  に分割される。回答者の集合  $s$  は、 $n_{(d)}$  の要素をもつ対応する部分集合  $s_{(d)} = s \cap U_{(d)}$  から構成されている。GREG 推定におけるように、LGREG 推定量は

$$\hat{i}_{(d)i} = \sum_{k \in U_{(d)}} \hat{\mu}_{ki} + \sum_{k \in s_{(d)}} a_k (z_{ki} - \hat{\mu}_{ki}) \quad (5)$$

が与えられることで適用される。

これら推定量の望ましい特性は加法性である。推定量(5)は第 4 章の実験において、領域推定のために使われている。

この代わりに、より有効的な LGREG 推定量を、GREG 推定量について Särndal *et al.* (1992) が示唆したように、領域推定のために利用できる。すなわち、

$$\hat{i}_{(d)i}^a = \sum_{k \in U_{(d)}} \hat{\mu}_{ki} + \left( N_{(d)} / \hat{N}_{(d)} \right) \sum_{k \in s_{(d)}} a_k (z_{ki} \hat{\mu}_{ki}) \quad (6)$$

である。ここで、 $\hat{N}_{(d)} = \sum_{k \in s_{(d)}} a_k$  であり、 $N_{(d)}$  は領域の既知の大きさである。推定量(6)の非加法性は欠点のように見える。加法性は、式(3)において固定的な効果とされる領域の既知の大きさに関する補足的情報の組み込むことによって得ることができる。それとは別に、多水準モデル (Goldstein 1992, Goldstein and Rasbasch 1992) は、推定手続きへの組み込みについては強いられかもしれない。

## 4 ケーススタディ： フィンランド労働力調査

### 4.1 フィンランド労働力調査の補足的データ

フィンランド統計局が毎月実施している労働力調査における雇用と失業の統計についての推定手続きに対して、利用可能なレジスター・ベースの補足的データは多様にある。Djerf(1997)の実証的研究においては、ある者が労働省の雇用業務統計によって求職者として登録されているかどうかを示す行政記録が、最も重要なレジスター・ベースの補足的データとして利用されていた。これら記録は、個人ベースで調査データと併合された。これは両データ源において唯一の個人識別番号を利用して行われた。追加された補助的データ（性別、年齢、居住地域）は中央人口レジスター（Central Population Register）から集められた。これらの補助的データもまたミクロレベルで調査データと併合された。

雇用と失業に関する月次の政府統計の推定手続きにおいて現在用いられている補助的変数は、性別、年齢階級、居住地域、そして雇用業務統計にしたがって失業期間の長さを表すカテゴリカル変数である。雇用業務統計のレジスターにおけるタイムラグは約2週間である。したがって、実際の労働市場での地位に変化のあった人の割合は、この短い間隔の中では小さいと期待できる。レジスター・ベースでの求職者地位が、労働力調査において測定される就業上の地位とは異なって定義されていることに注意すべきである。調査での測定は標準的なILOの定義に基づいている。

レジスター・ベースの求職者の地位は、個人の失業についての調査での測定と強く相関している ( $r=0.8$ )。したがって、効率の改善と単位無回答による偏りの減少は、推定手続きにこれらの補助的データを利用することによって獲得される。雇用と失業に関する統計の推定手続きは、不完全な事後階層化あるいはレイキング・モデルに対応する標準線形 GREG 推定量に基づいている（詳細は Djef 1997 を参照）。

### 4.2 フィンランド労働力調査データでの実験

#### 4.2.1 作成された母集団

Lehtonen and Veijanen(1998 a)では、失業率の推定量の動きが詳細に研究された。1994年の連続3ヶ月についてのフィンランド労働力調査データがシミュレーション研究に用いられた。作成された母集団は 33,329 人から構成された。年齢階級（15歳-24歳、25歳-34歳、35歳-44歳、45歳-54歳、そして55歳-64歳）、性別、そして地域（3地域）は、それぞれの母集団のメンバーについての人口レジスターから得られた。求職者指標は、どの個人が失業中の求職者として登録されているかを示す雇用業務統計レジスターから得られた。これらすべての補助的データはミク

ロレベルで調査データと併合された。

無回答率は求職者の地位によって異なり、登録された求職者においてはその率は11.4%であったのに対して、その他の者についての率は7.6%であった。無回答の確率はロジスチック ANOVA モデルによってモデル化され、無回答率の最大尤度推定値(2.9%から22.8%の範囲におよぶ)はシミュレーションにおける無回答モデルとして利用された。

$N=30,835$  人で構成される人工的な母集団が、シミュレーション実験のために作成された。労働市場での地位は3つの階級、すなわち「雇用者」、「失業者」と「非労働力」として定義され、その母集団での度数はそれぞれ  $t_1=17,373$ 、 $t_2=4,433$ 、そして  $t_3=9,029$  であった。失業率は  $R=t_2/(t_1+t_2)=20.33\%$  で定義された。年齢階級、性別、そしてレジスター・ベースの失業の地位とのクロス表におけるセルは領域として用いられた。

人工的母集団から、 $n=1000$  人の大きさの独立ランダムサンプルが置き戻しなしの単純ランダムサンプリングで抽出された。それぞれのサンプルにおいて、無回答は元の母集団に合わせた無回答モデルによってシミュレートされた。回答確率は、無回答モデルにおいてと同じ ANOVA モデルを伴うロジスチック回帰によって、それぞれのサンプルから推定された。それぞれの算入確率  $\pi_k$  は回答確率推定値  $\hat{\theta}_k$  で乗じられる。

#### 4.2.2 シミュレーション研究の詳細

推定値のモンテカルロ平均値と標準誤差は、シミュレートされた標本から計算された。推定量  $\hat{t}_i$  についての設計効果は推定された分散の率、すなわち、

$Deff(\hat{t}_i) = \hat{V}_{mc}(\hat{t}_i) / \hat{V}_{mc}(\hat{t}_i^{HT})$  として計算された。ここで、 $\hat{V}_{mc}(\hat{t}_i^{HT})$  は HT 推定量のモンテカルロ分散推定値を表している (Lehtonen and Pahkinen 1996)。

階級別度数の信頼区間は、階級指標が独立であるかのように計算される。 $K=1,000$  のシミュレートされた標本について、名目上の有意水準 95%で受け入れられるカバレッジ率は [93.65%、96.35%] の間にある。

領域推定において、領域推定値の全体的な正確さは、 $D$  領域と  $K$  個の標本  $s_j$  にわたる相対的領域誤差の絶対値の平均によって計測される。すなわち、

$$MARD(i) = \frac{1}{D} \sum_{p=1}^D \frac{1}{K} \sum_{j=1}^K \frac{100 \left| \hat{t}_{(d_p)i}(s_j) - t_{(d_p)i} \right|}{t_{(d_p)i}}$$

#### 4.2.3 LGREG 推定についての結果

LGREG 推定量を構成するとき、補助変数  $x_k$  の構成要素は、年齢（5 階級）、性別、地域（3 地域）、そして求職者の地位に対応するダミーである。不完全な事後階層化もしくはレイキングにおいて、主な効果 ANOVA モデルは分類された補助変数に基づいている。求職者指標を伴うモデルと伴わないモデルとが比較された。ここでの関心は主にこの推定の効率にあるので、ウェイト  $a_k = 1/(\pi_k \hat{\theta}_k)$  での調整 HT 推定量は、未調整の HT 推定量の代わりに参考推定量として使われた。このようにして、HT 推定量の偏りは単位無回答によって、無視できるものであった（表 1）。これは LGREG 推定量によく当てはまる。

調整 HT 推定量は、予定効果 (*Deff*) 統計（表 1）で示されるように、LGREG 推定量よりも大きな分散をもつ。補助変数として年齢、性別、地域を伴うレイキング・モデルに基づく LGREG 推定量は、HT 推定量にいくらかの改善をもたらした。はるかによい結果が求職者指標を含む LGREG 推定量によって獲得された。この求職者指標は、他の補助変数よりも ILO の失業指標との間により強い相関 ( $r=0.83$ ) がある。したがって、これら行政的補助データは推定の効率を大きく改善する。カバレッジ率はすべての推定量について受容できるものだった（表 1）。

表 1 求職指標を伴う場合 (Y) と伴わない場合 (N) のレイキング・モデルについての失業率推定量  $\hat{R}\%$  の特性。SD は標準偏差を表し、CR (%) は 95%信頼区間のカバレッジ率を表す。

モデル	方法	$\hat{R}$	バイアス	SD	Deff	CR
	HT	20.32	- 0.0081	1.461	1	95.7
レイキング -N	LGRE G	20.31	- 0.0229	1.454	0.995	95.3
レイキング -Y	LGRE G	20.29	- 0.0419	0.901	0.617	94.8

#### 4.2.4 領域推定についての LGREG と GREG との比較

領域推定に関する LGREG、GREG 推定量の性質の比較について、領域はその領域の大きさに対応して 2 つのグループに分けられる。雇用者、失業者、非労働力の領域総計は、LGREG と GREG で推定された。性別、地域、求職者指標の補助変数に加えて、4 次の年齢の多項が含まれるモデルが用いられた。

領域総計は、とりわけ小さな領域において、GREG よりも LGREG を用いたほうがより正確に推定される（表 2）。モデルに連続的補助変数として年齢が含まれる場

合、失業率推定値の標準偏差は、20のうち19の領域において、GREGよりもLGREGについての方がより小さい。残念にも、LGREGによって得られる信頼区間がしばしば、小さな分散推定値によって、特に100よりも小さい観測値の領域のグループにおいて、狭すぎる。より大きな領域については、LGREGとGREG推定量の性質は近似している。

表2 相対領域誤差の絶対値平均 (MARDE) と信頼区間 95%の平均カバレッジ率 (CR)。真の度数  $t_{(d)i}$  ( $i=1,2,3$ ) を持つ領域での推定階級度数について。

(a) は 100 未満、(b) は 100 以上の観測値。モデルは年齢の多項を含む。

	方法	MARDE			CR		
		$\hat{t}_{(d)1}$	$\hat{t}_{(d)2}$	$\hat{t}_{(d)3}$	$\hat{t}_{(d)1}$	$\hat{t}_{(d)2}$	$\hat{t}_{(d)3}$
(a)	GREG	96.92	67.36	121.95	88.2	77.8	84.6
	LGREG	80.28	67.20	104.05	83.9	76.5	51.7
(b)	GREG	6.95	12.31	14.35	94.1	85.9	93.7
	LGREG	6.88	12.34	14.29	93.9	85.4	93.3

## 5. 要約

推定手続きにおいてレジスター・ベースの行政データを用いることは、標本調査に基づく政府統計の生産にとって大変有益となり得る。その利点は、効率の改善、無回答による偏りの縮小にある。フィンランドの労働力調査からの例は、その利点が対象となっている補助変数と調査変数との間の相関の強さに依存するというを示した。相関が強い場合には、適切な補助変数を組み入れていない推定手続きと比較すると、例えば、失業率の推定について分散を大きく縮小できる。

多項調査変数は、調査の実施において一般的に遭遇する。そのような変数の階級総計の推定値に関して、新たな推定手続きが Lehtonen and Veijanen(1998 a)によって導入された。ロジステック一般化回帰推定値 (LGREG) は多項ロジステックモデルに基づいている。後者は、階級指標については、一般化回帰推定 (GREG) において通常用いられている線形モデルよりも現実的かもしれない。完全な事後階層化については、LGREGとGREGの両推定量は同じ結果をもたらすが、レイキングのようなその他のモデルでは、その結果は異なる。

ロジステック一般化回帰推定は標準一般化回帰推定と密接に関係している。標準GREG推定では、利用可能なレジスター・データは、その推定手続きに、適合する

線形モデルによって組み込まれる。LGREG 推定では、階級指標の結合分布は多項ロジステックモデルで描かれる。モデル作成者の視点からすると、このモデルの定式化は、特に領域推定について、現実的でよく機能しているように見受けられる。Lehtonen and Veijanen(1998 a) は、ロジステックの定式化は線形の定式化よりも、特に比較的小さな領域において良い結果を生み出すということを示した。この結果は、フィンランド統計局が毎月実施している労働力調査から得られた実際のデータについて行われたモンテカルロ実験に基づいたものであった。

## 文献

- Andersson C. and Nordberg L. (1998). A User's Guide to CLAN 97. Örebro: Statistic Sweden.
- Defays D. (1997). Use of administrative sources for statistical purposes: threats and opportunities. In European Commission (1997). *Proceedings of the Seminar on the Use of Administrative Sources for Statistical Purposes*, Luxembourg, 15-16 January 1997. Luxembourg: Office for Official Publications of the European Communities, 121-130.
- Deville J.-C. and Särndal C.-E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association* 87, 376-382.
- Deville J.-C. and Särndal C.-E. and Sautory O. (1993). Generalized raking procedures in survey sampling. *Journal of the American Statistical Association* 88, 1013-1020.
- Djerf K. (1997). Effects of post-stratification on the estimates of the Finnish Labour Force Survey. *Journal of Official Statistics*, 13, 29-39.
- Estevao V., Hidioglou M.A. and Särndal C.-E. (1995). Methodological principles for a Generalized Estimation System at Statistics Canada. *Journal of Official Statistics*, 11, 181-204.
- European Commission (1997). *Proceedings of the Seminar on the Use of Administrative Sources for Statistical Purposes*, Luxembourg, 15-16 January 1997. Luxembourg: Office for Official Publications of the European Communities.

- Franchet Y. (1997). Opening Address for the Seminar on the Use of Administrative Sources for Statistical Purposes. In European Commission (1997). *Proceedings of the Seminar on the Use of Administrative Sources for Statistical Purposes*, Luxembourg, 15-16 January 1997. Luxembourg: Office for Official Publications of the European Communities, 1-6.
- Godambe V.P. and Thompson M.E. (1986). Parameters of superpopulation and survey population: their relationships and estimation. *International Statistical Review*, 54, 127-138.
- Goldstein H. (1995). *Multilevel Statistical Models*. London: Edward Arnold, New York: Halsted.
- Goldstein H. and Rasbasch J. (1992). Efficient computational procedures for the estimation of parameters in multilevel models based on iterative generalized least squares. *Computational Statistics and Data Analysis* 13, 63-71.
- Gåsemyr S. (1996). The quality of an integrated register-based statistical system. In proceedings of the 5th Independent Conference of IAOS, 2-5 July 1996, Reykjavik (CD Rom).
- Grais B. (ed.) (1996). *The Future of European Social Statistics. Use of administrative Registers and Dissemination Strategies*. Proceedings of the Mondorf Seminar, Third Session, 25 and 26 January 1996. Luxembourg: Office for Official Publications of the European Communities.
- Harala R. (1995). Evaluation of the results of the register-based population and housing census in Finland. *Statistical Journal of the United Nations ECE*, 12, 63-72.
- Harala R. (1997). Statistical properties and quality of register-based census statistics in Finland. Ottawa: *Proceedings of the Symposium 97, New Directions in Surveys and Census*, 45-50.
- Holt D. and Smith T.M.F. (1979). Post stratification. *Journal of the Royal Statistical Society A* 142, 33-46.
- Lehtonen R. and Pahkinen E.J. (1996). *Practical Methods for Design and Analysis of Complex Surveys*. Revised Edition. Chichester: Wiley.
- Lehtonen R. and Veijanen A. (1998a). Logistic generalized regression estimators. *Survey Methodology* 24, 51-55.
- Lehtonen R. and Veijanen A. (1998b). *On multinomial Logistic Generalized Regression Estimators*. Jyväskylä: Preprints from the Department of Statistics, University of Jyväskylä, No.22.



- McCullagh P. and Nelder J.A. (1989). *Generalised Linear Models*, Second Edition. London: Chapman and Hall.
- Myrskylä P. (1991). Census by questionnaire – census by registers and administrative records: the experience of Finland. *Journal of Official Statistics*, 7, 457-474.
- Nordberg L. (1989). Generalized linear modeling of sample survey data. *Journal of Official Statistics*, 5, 223-239.
- Renssen R. H., Nieuwenbroek N.J. and Slootbeek G.T. (1997). Variance module in Bascula 3.0: Theoretical background. Heerlen: Statistics Netherlands, Research Paper no. 9712.
- Sautory O. (1993). La macro CALMAR. Redressement d'un échantillon par calage sur marges. I.N.S.E.E. Série des documents de travail n°F 9310. Paris: I.N.S.E.E. (In French).
- Skinner C.J., Holt D. and Smith T.M.F. (eds.) (1989). *Analysis of Complex Survey*. New York: John Wiley.
- Spiker F. (1996). Data available for statistics from the public administration. In European Commission (1997). *Proceedings of the Seminar on the Use of Administrative Source for Statistical Purposes*, Luxembourg, 15-16 January 1997. Luxembourg: Office for Official Publications of the European Communities, 239-246.
- Särndal C.-E., Swensson B. and Wretman J.H. (1992). *Model Assisted Survey Sampling*. New York: Springer-Verlag.
- Thomsen I. And Kleive Holmøy A.M. (1998). Combining data from surveys and administrative record systems. The Norwegian experience. *International Statistical Review* 66, 201-221.
- Vihavainen H. (1996). Potential use of administrative registers to alleviate the burden on census or survey respondents. In Grais B. (ed.) (1996) *The Future of European Social Statistics. Use of Administrative Registers and Dissemination Strategies*. Proceedings of the Mondorf Seminar, Third Session, 25 and 26 January 1996. Luxembourg: Office for Official Publications of the European Communities, 41-48.
- Wallgren A. and Wallgren B. (1997). The role of administrative registers in an efficient statistical system: methodological problems and quality issues. In European Commission (1997). *Proceedings of the Seminar on the Use of Administrative Sources for Statistical Purposes*, Luxembourg, 15-16 January 1997. Luxembourg: Office for Official Publications of the European Communities, 177-194.

## 人口統計の250年

## 人口統計の歴史

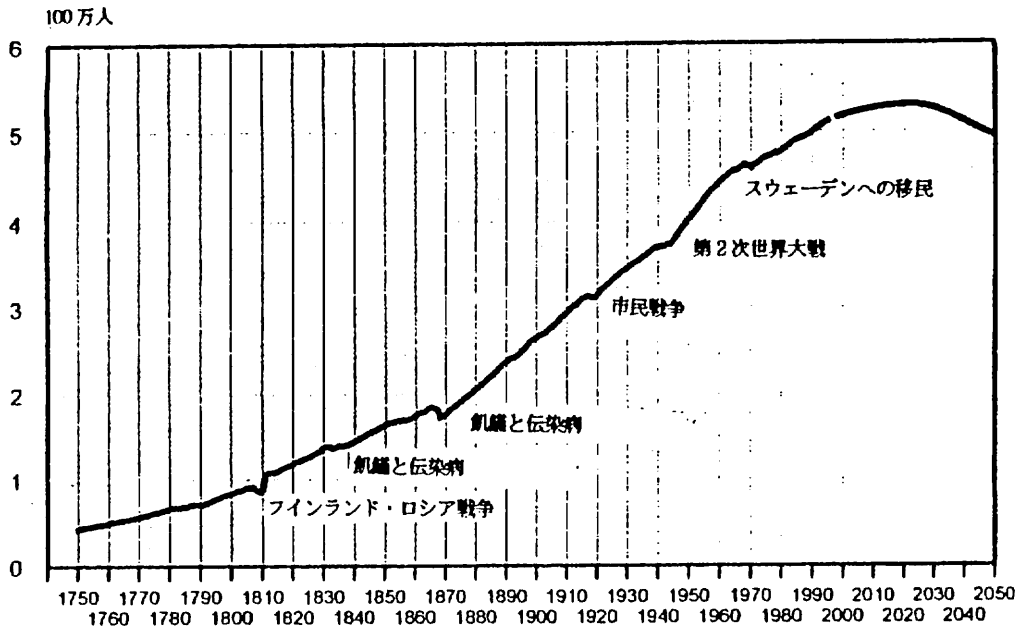


図1 人口 1749-2050年

**16世紀** 最初の人口表 (population roles) が作成される

16世紀のはじめに、スウェーデン王グスタフ・ヴァーサ (Gustavus Wasa)の布告によって、税の収集がはじまり、最初の人口表が作成される。

**17世紀初期** 教会が洗礼、埋葬、結婚の数を保管しはじめる

最初のデータは、1622年にヴェステラス (Vesterås) の司教区において、そして1928年に、現在フィンランドの地域であるオーボ (Åbo)の教区地域について集められている。当局は、軍隊に適格な男性のレジスターを作成しはじめる。

**1686 年**

### **教会法**

すべての教会は、法律によって、洗礼、埋葬、結婚、教区間の移住をふくむレジスターを保持することを義務とされる。

**18 世紀初期**

### **洗礼と埋葬についてデータが収集される。**

幾人かの書記が、その教区であった洗礼と埋葬の要約をまとめる。後に、統計家は、洗礼と埋葬についての総ての利用可能なデータを 1722 年以降について集めている。人口統計の定期的な作成を開始するよう様々な提案が行われる。例えば、スウェーデン科学アカデミーは、人口動態統計に関する計画を作成している。

**1748 年**

### **人口動態統計の作成をはじめることが決定される。**

1748 年 2 月 3 日、スウェーデン王、フレデリック I 世は、人口動態統計の作成の開始について、データを収集する書式とともに承認を与える。

**1749 年**

### **世界で最初の年次人口統計**

データは、3つの書式を使って収集されている。1つは、人口の大きさを、性別、社会階級に対応して、年齢の5歳区分の年齢グループ別に与えるものである。他の2つの書式は、社会階級別の洗礼、結婚、死因別の年齢5歳区分別の埋葬に関するレジスター・データについてのものである。人口変動に関する統計的書式は、毎年記録されているが、センサスの書式は、1751 年からは、2年毎に、1775 年から 1880 年までは5年毎に、それ以後、1940 年までは10年毎に作成されている。

**1756 年**

### **世界で最初の統計局**

人口学的な要約を作るために、ストックホルムで、特別な統計委員会が創設される。

**1773 年**

### **出生統計の改革**

出産した女性についてのデータが、母親の年齢についての5歳グループ別に収集されるようになる。

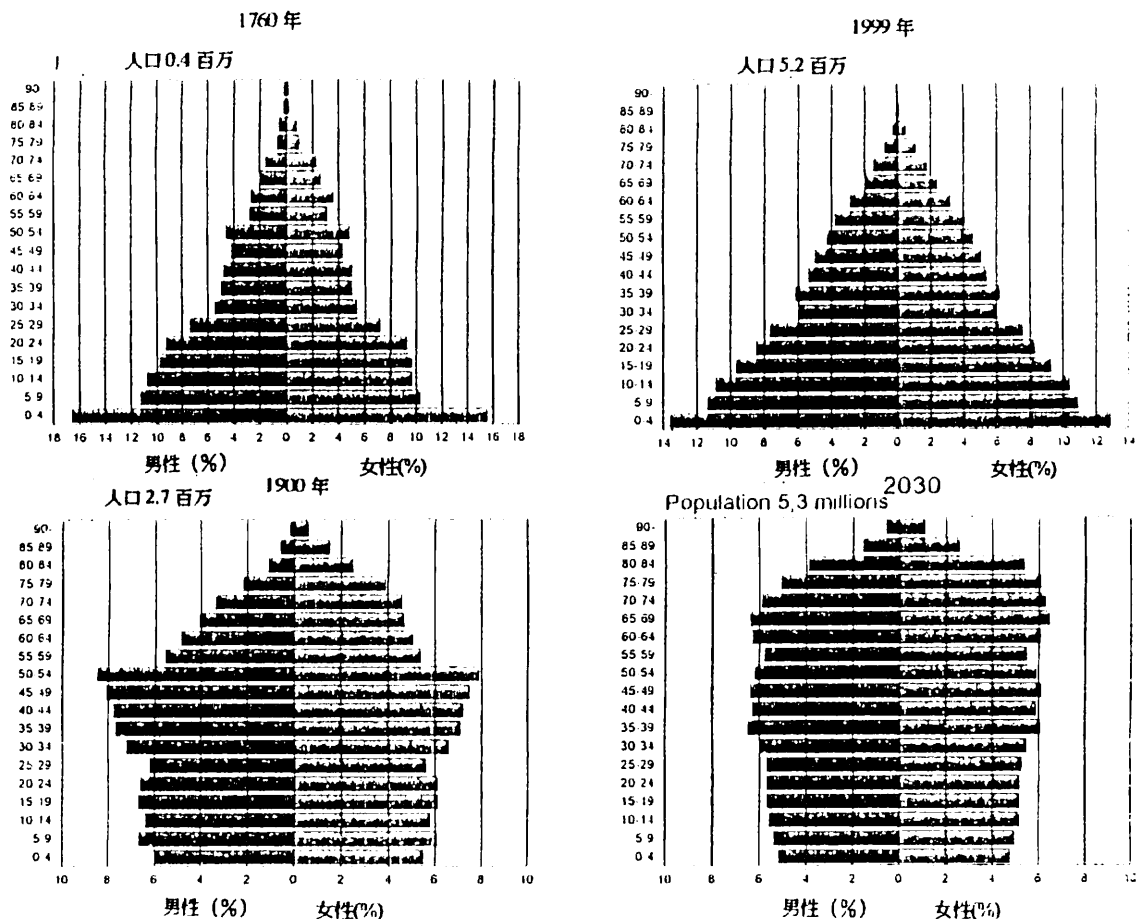


図2 人口年齢ピラミッド 1750-2030年

**1775年** 社会階級に関する書式の改革

2つにの職業を持つ者が2回書式に登録されるよう社会階級についてのレジストレーションが変えられる。これは、明らかにこれまでのものより正確度が低い。

**1802年** 社会階級に関する書式の改革

職業を二重に登録することは停止され、別個の書式が、町や地方に導入される。この書式での職業のリストが拡大される。

**1809年** フィンランド大公国

1808~1809年のフィンランド戦争の結果、フィンランドは大公国としてロシア帝国に属することになる。これは、スウェーデンの支配下にあったフィンランドの数世紀の終わりを印すものである。

**1812 年**

**古いフィンランド地域がフィンランド大公国に  
繰り入れられる**

結果として、フィンランドの人口は、ほぼ 190,000 人増加する。人口の一部は正教徒であり、人口登録からは除かれている。フィンランドに移動してきたロシアの役人と兵士はもまた動態統計から除外されている。

**1859 年**

**人口研究に関する最初の博士論文**

パーヴォ・ティッカネン(Paavo Tikkanen)の論文「フィンランドにおける人口の大きさと人口段階の関係」は、フィンランド語で書かれた最初の博士論文の 1 つである。

**1865 年**

**中央統計局が創設される**

中央統計局が、最初は臨時的なものとして、5 年後に恒久的なものとして創設される。

**1870 年**

**最大の町での最初のセンサス**

データの直接的収集に基づく最初のセンサスが、ヘルシンキ(Helsinki)、トゥルク(Turku)、ヴィープリ(Viipuri)とオウル(Oulu)で行われる。それらのデータは世帯ごとに収集される。人口学的データに加えて、家族、世帯、住居に関するデータもまた収集される。1870 年から 1930 年まで、センサスは、他の大きな町でも 10 年毎に行われている。

**最初の人口動態統計が公表される。**

最初の人口動態統計出版物『1865 年 12 月 31 日のフィンランドの人口』がフィンランド政府統計シリーズ VI で公表される。この後、年次統計が定期的に発行される。

**1877 年**

**人口統計の改革**

人口と死亡に関するデータが 1 歳区分に拡大される。移住データが、それまで町についてだけであったものが、すべての自治体から収集されるようになり、統計の書式は広範なものになる。最初の離婚統計が作成される。

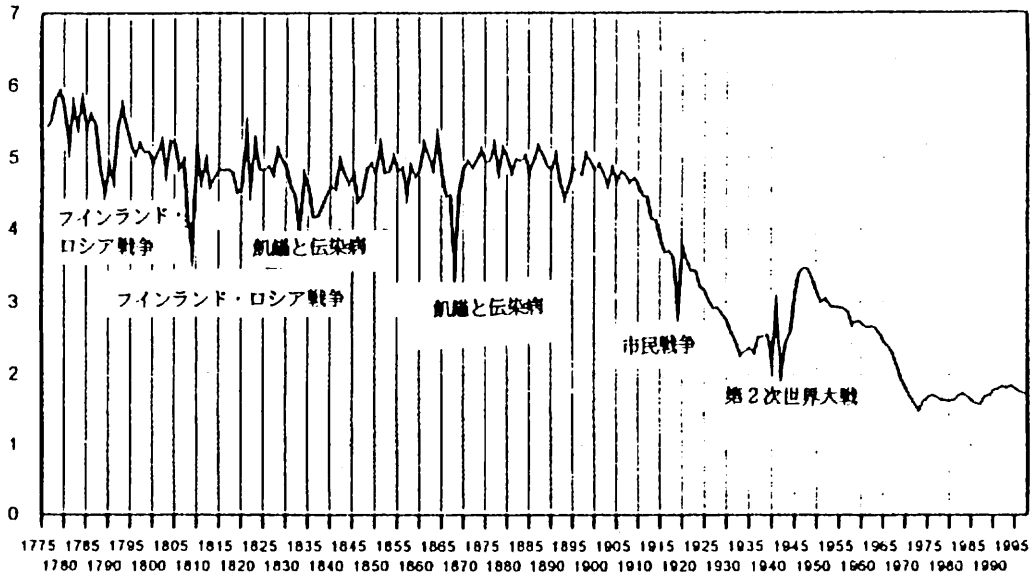


図3 合計特殊出生率 1776-1998年

### 1881-1890年 最初の生命表が公表される。

最初の生命表が公表される。後に、生命表は、1751年から1880年まで、10年期間毎に計算された。

### 1883年 最初の移民データが収集される。

オウル(Oulu)とヴァーサ(Vaasa)県からの海外への移民は、県の行政委員会に登録する義務を負うようになり、この行政委員会がその後、人口統計部局にこのデータを転送する。1893年に、移民統計がはじめて全国について作成される。この統計は合衆国へ移住する者だけをふくんでいる。

### 1900年 パスポート一覧表に基づく移民統計

最初の移民統計がパスポートの一覧表に基づいて作成される。最初、この統計は合衆国への移出民だけを含めていたが、1924年以降、他の諸国への移出民も登録されるようになる。移民統計は、1984年迄、パスポート一覧表に基づく。移民統計は、1945年まで、フィンランド政府統計シリーズ XXVIII で公表された。

## 1910年 現存人口と不在の人口

移出民はしばしば、教区のレジスターからは除かれず、そのデータは人口統計に登録されたままになっている。地方の人口レジスターの管理者は、移出して、もはや人口レジスター地域には住んでいない人口、すなわち、いわゆる不在人口とは反対の、人口レジスター上の人口を当局に通知する義務を負うことになる。その狙いは、人口の実際の大きさに関するより正確な情報をうることにある。

## 1917年 フィンランドが独立する。

フィンランドは、1917年12月6日に独立したフィンランド共和国になる。

### 住民登録が創設される。

どの合法的な教会にも属していないすべての人が、住民登録に入ることになる。人口統計は、このようにして今やフィンランドに住んでいる総ての人をふくむようになる。

## 1923年 人口登録システムを改革する試み

人口統計の改革について作業した委員会は、人口レジストレーションは、今後、国家当局のみが遂行することを提案している。同時に、人口センサスのためと人口変動の登録のためのデータもまた、人口の目録に登録されるべきことも提案した。この時点まで、人口レジスターの管理者は、このデータを統計家に完全な統計表として配布していた。教会は、この改革提案を拒否している。

## 1930年 離婚統計の改革

1930年結婚法の結果、今や離婚と別居に関するデータが、識別可能な書式で裁判所から受け取れるようになる。

## 1934年 最初の人口予測

中央統計局は国全体についての最初の人口予測を公表している。それは、フィンランドの人口が、400万人を越えることは決してないと予測している—現在の人口は500万人を越えている。

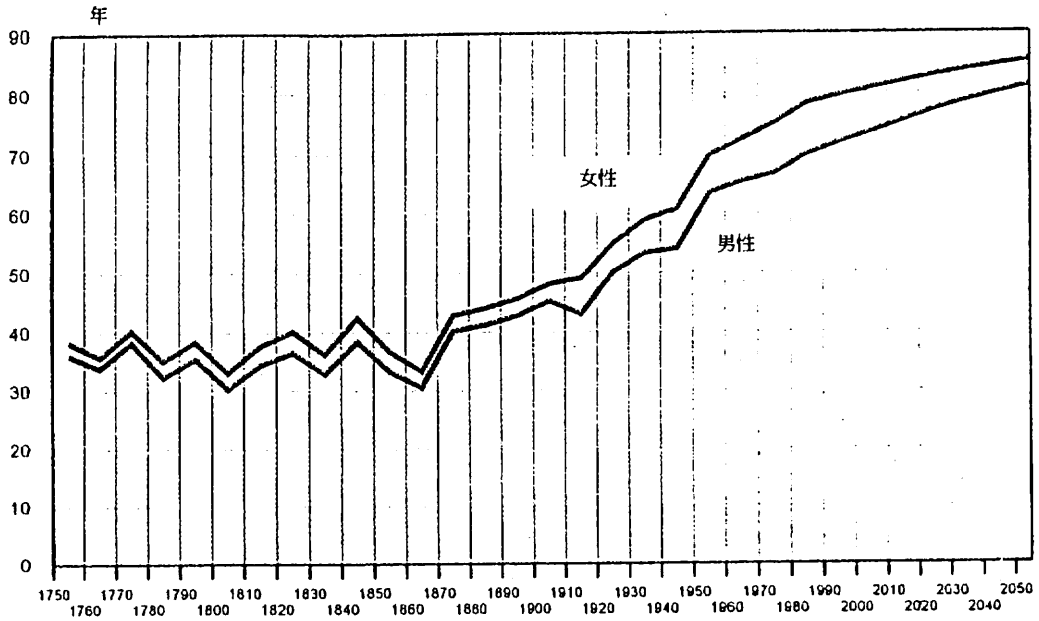


図4 0歳児の平均余命 1751-2050年

**1936年** 死亡証明書が導入される。

死亡に関するデータが、死亡証明書に基づいて登録される。死亡証明書は、地方の人口レジスターから検査のために県当局に、そしてさらに中央統計局へと渡される。

**母親の年齢別の出生**

1773年以降、出産数に関するデータは、母親の年齢に応じて収集されてきている。今では、データは、総ての生まれた子について母親の年齢別に生産されている。

**1938年** センサス法が発効する

センサス法(154/38)によれば、住居と不動産の数え上げとともにセンサスは、10年ごとに実施されるべきである。センサスは直接の調査に基づくべきであり、数え上げられた個人についてのデータは、特別な書式で配布されることとされている。フィンランドのセンサスは現在では完全にレジスターに基づいているが、センサス法は、なお有効である。

**1939年** 出生統計の改革

生まれた個人についてのデータは、今では母親の年齢に応じて、1歳グループごとに収集されている。



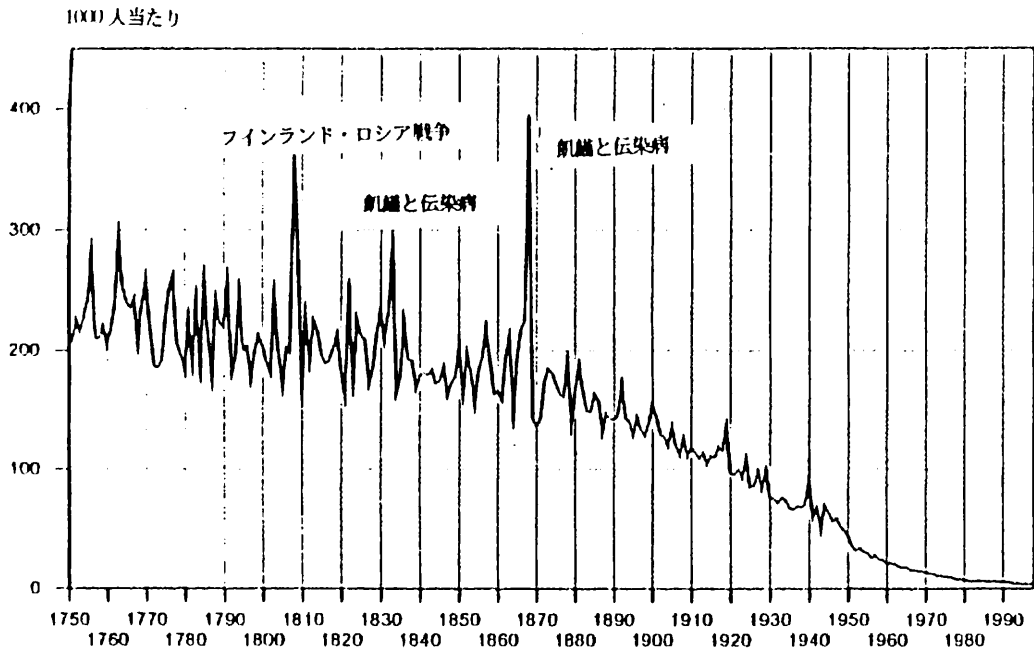


図5 幼児死亡率 1751-1998年

### 1940年 センサスが中止される。

戦争によって、センサス法に基づく最初のセンサスが中止される。人口の大きさと構造についての統計は、地方の人口レジスターから収集されたデータによって作成される最後のものとなる。

### 1949年 全国についての人口予測

全国についての人口予測が、種々の他の人口変化の選択肢についての計算とともに公表される。この公表はまた、人口予測の信頼性の問題を提起し、古い人口数の修正をもふくんでいる。この古い数字は、1890年以降国内に住んでいる人口数の計算によって獲得されてきたものである。

### 1950年 人口・住宅センサス

全国についてフィンランドでは最初の調査票ベースのセンサスが行われる。このセンサスは国連が承認した概念、定義、分類に基づいている。女性の出生力に関するサーベイがこのセンサスとの関連で行われる。データは自治体よりも小さな人口センターからも収集される。

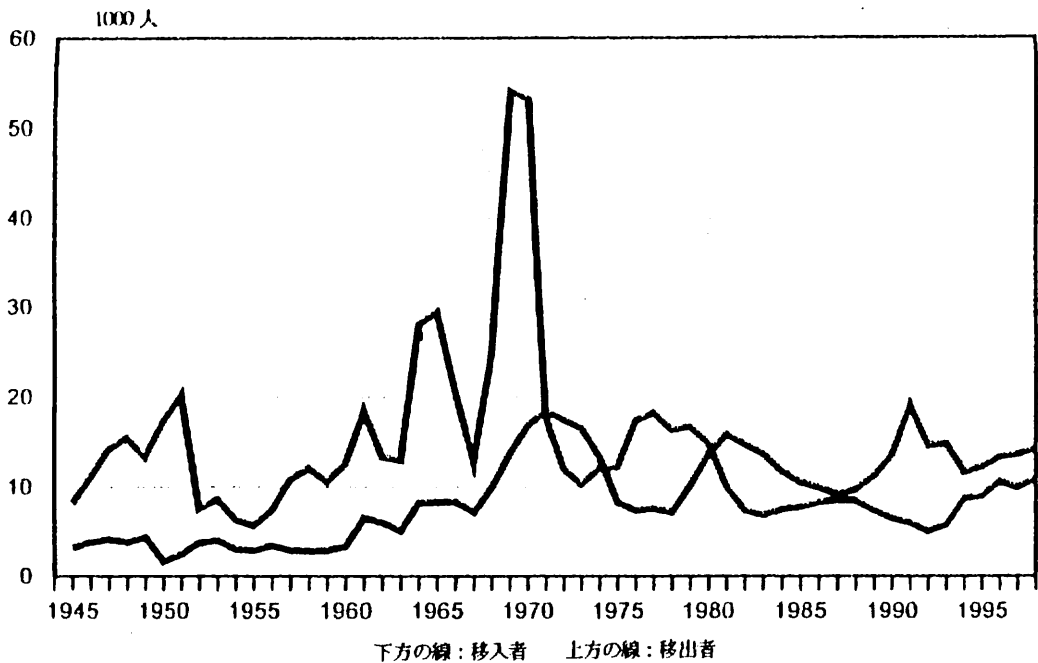


図6 移入および移出者数 1945-1998年

### 1951年 結婚と移住の統計の改革

それぞれの結婚契約についてのデータは別個に入手されている。  
自治体間の移住のデータは、サーベイを使って収集されている。

### 1954年 北欧共同労働市場

北欧共同労働市場協定が発効するとともに、北欧諸国間の旅行においてパスポートは不要になった。北欧諸国内での移住がもはや統計に登録されなくなったので、パスポート一覧表に基づく移住統計はかなり不正確になる。

### 1956年 統計委員会の報告

統計委員会が、人口統計の作成に関して種々の改革を提案し、幾つかの欠点について注意を払っている。欠点のうちのもっとも中心的なものは、移住統計がパスポート一覧表のみに基づいていることと、死因に関する統計の作成に医療専門家が欠けていることである。委員会はまた、都市居住地についての新しい定義を提案し、地域人口予測の作成を開始すべきことを示唆している。

**1960 年**

### 人口・住宅センサス

個人調査票に基づく2回目のセンサスが行われる。データはコンピュータによって生産される。

### ローカリテイについての地理情報

ローカリテイについての北歐共同定義がはじめて、正方形法を使って適用される。すなわち、ローカリテイは、基本地図での1km<sup>2</sup>の正方形の地域によって形成される多角形によって境界を画される

**1964 年**

### 識別番号の導入

国民年金機関の健康保険制度が、個人識別番号を導入する。これは、フィンランドに永住しているすべての人に指定される。個人識別番号は、個人の誕生日と個人番号（男性は奇数、女性は偶数）と照合番号からなる。

### 地域人口予測

中央統計局は、最初の地域人口予測を公表する。予測の際の地域分類は統計区域に基づいている。

**1966 年**

### 人口変化を報告する新しい書式

人口変動についてのデータは、単一個人にそくした1つの書式によって収集される。移民データもまた今や、これまでの標本法ではなく、すべての移住のケースを獲得する。この書式は地域人口レジスターにおいて記入される。

**1969 年**

### 人口レジスター・センターが創設される

中央人口レジスターのデータは、住居レジスターとの関連で収集され、教区レジスターから獲得されたデータによって完成する。中央人口レジスターはまた、個人の住居や不動産に関して識別可能なデータをふくんでいる。中央人口レジスターは、内務省の1部署である人口レジスター・センターが維持している。

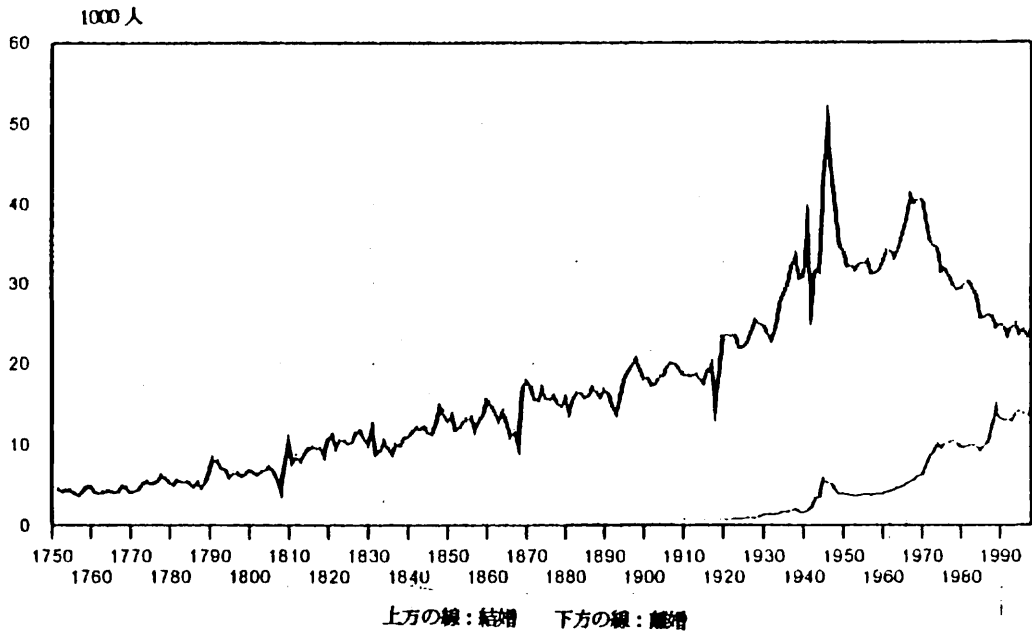


図7 結婚と離婚 1751-1998年

### 自治体についての人口予測

中央統計局は、自治体についての最初の人口予測を完成する。自治体についての以前の人口予測は、1963年からのものであり、それは、国家計画機関によって公表された。

## 1970年

### 人口・住宅センサス

センサス調査票は、中央人口レジスターに登録されている総ての人に対して郵便で送られている。個人識別番号がはじめてセンサス目的に使用され、宗教や出生地データは中央人口レジスターから直接的に獲得される。所得データもまた、税務当局から獲得された。最終学歴と学位についてのレジスターが創設される。

### 地域ベースから点ベースの統計体系へ

1970年センサスでは、建物が地域統計の基本単位として使われている。建物についての地図座標の採用は、統計的地域の決定をより柔軟にすることを可能にする。

## 非行政的統計領域の発展

地域コードとと並ぶ地図座標上の統計単位によって、ローカリテイー、地方自治体の下部地域、郵便番号地域、グリッドそして他の非行政的地域別の統計を生産することが可能になる。

**1971年**

### 人口動態統計の基本データに識別番号が付け加えられる

1971年以降、個人識別番号が、人口変動に関する総てのデータに含められる。これによって、人口変動についてのデータと個人識別番号を使っているあらゆる資料とを結びつけることが可能になる。

### 人口センサスからの背景情報

移出民と移入民についてのデータが、1970年センサスからの背景データと結びつけられている。

**1975年**

### 人口・住宅センサス

このセンサスは、1970年センサスとほとんど全く同じ形で行われている。出生日、性、婚姻上の地位、母言語、国籍に関するデータは、中央人口レジスターから、他のデータは調査票によって収集されている。教育に関するデータは、フィンランド統計局の最終学歴・学位レジスターから、所得データは、税務署のレジスターから獲得されている。

**1977年**

### レジスターに基づく家族統計

家族統計は、中央人口レジスターから得たデータに基づいて作成されている。家族統計は、このときから2年毎に、1994年からは毎年発行されている。

**1979年**

### 住宅状況と死亡率

1970年センサスで獲得された社会・経済的データが、死因に関する1971-1975年統計と結合されている。サーベイ資料が、死亡率の社会・経済的要因に関する広いサーベイの基礎を提供している。この広いサーベイは、最近の諸センサスからのデータと死因統計のデータを結びつけている。

## **1980年**      **人口・住宅センサス**

このセンサスは、様々なレジスターから直接的に獲得されるデータにますます依存している。人口レジスターと税当局のレジスターとが、国民年金庁 (the National Pension Institute) のレジスターと結びつけられている。

人口レジスター・センターが建物・住居についてのレジスターを作成する。

## **1981年**      **地域データベース**

直接的にアクセス可能な地域データベースが、統計データの利用を促進するために創設される。

## **1985年**      **人口・住宅センサス**

書式が郵送で配布され、返送される。様々な行政レジスターがますます多く利用される。仕事と職業に関するデータだけが書式によって収集される。

## **1987年**      **地域就業統計**

統計家は様々なレジスターから獲得されるデータを使って年次地域就業統計の作成を開始する。

## **1989年**      **生来住所の最後の登録**

最後の生来住所登録が1989年1月1日に行われている。これは、450年にわたる古い伝統の終わりである。中央人口レジスターのデータは、もはや毎年チェックはされない。

## 家族形成と出産力サーベイ

家族形成と出産力に関する中央のデータが標本抽出によって収集される。このサーベイはまた、態度の形成についての質問もふくんでいる。女性がはじめて参加している。1992年の同じサーベイは男性について行われている。このサーベイは、ECEが組織した家族と出産力調査(FFS)の一環である。

## ロンジチュージナル・センサス・データ

センサスから得られる個人にそくしたデータは、いわゆるロンジチュージナル・データ、すなわち、1970年以降の種々のセンサスから収集される個人にそくしたのデータ、を形成する。現在では、1970～1995年のロンジチュージナル・センサス・データは、個人の人口学的データだけでなく、所得や経済活動、また家族や世帯のデータもふくんでいる。1950年以降収集されている標本データは、実際のロンジチュージナル・センサス・データとして、同じような個人にそくして歴史的な追跡情報をふくんでいる。

## 1990年

### レジスターだけに基づく人口センサス

フィンランドは、センサスを様々な行政レジスターだけに基づいて行うことをはじめた世界で2番目の国である。センサス目的で使われるレジスターの中で最も重要なレジスターは、人口レジスター、税務署レジスター、中央年金庁の就業レジスター、大蔵省の就業レジスターと自治体年金機関の就業レジスター、社会保障機関の年金生活者レジスター、種々の学生レジスター、求職者レジスター、フィンランド統計局の企業レジスター、および最終学歴・学位レジスターである。

## 1992年

### GIS—統計における地理

地理情報技術の実施が、定められた地理情報プロジェクトによって開始された。

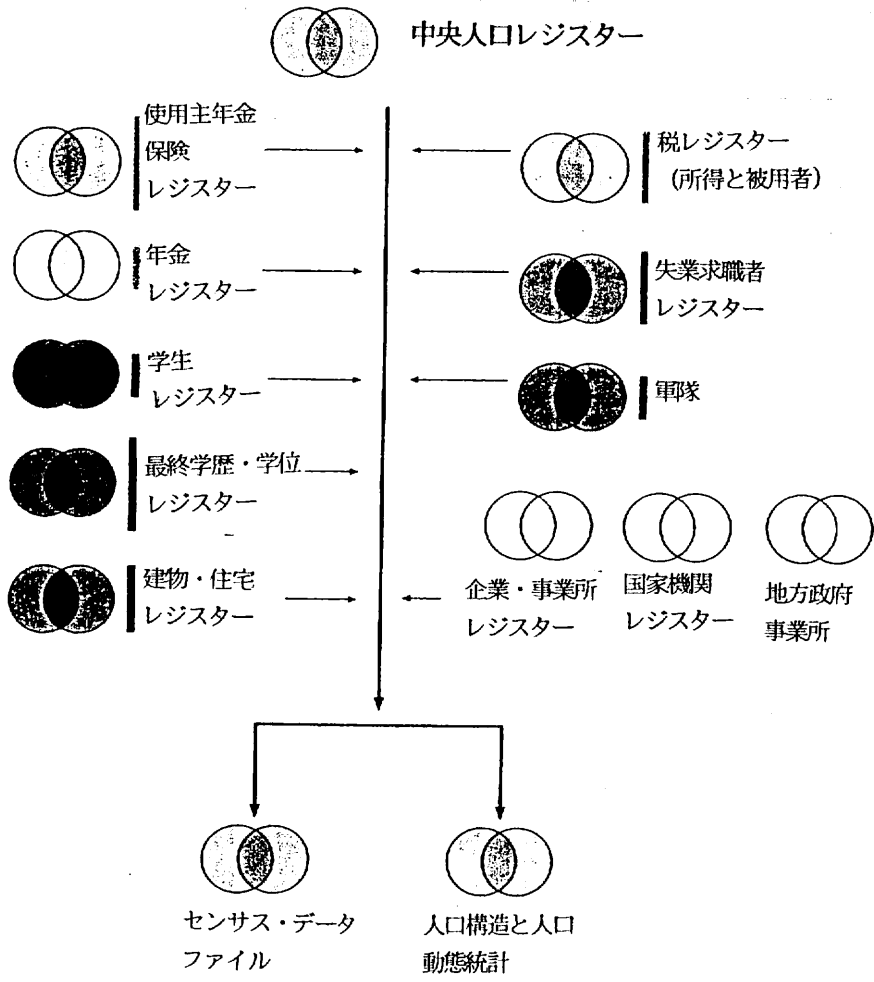


図8 レジスター・ベースの人口統計システム

1999年 人口統計の基礎としてのレジスター・データ

最初の人口統計が公表されてから250年後に、フィンランドは、今やセンサスや人口構造や人口変動についての統計を様々なレジスターを使って作成するシステムを使っている。



## 接触先情報

フィンランド統計局	住所 : Tyopajakatu 13, Helsinki, Finland
郵送先住所	FIN-00022 STATISTICS FINLAND
電話	+358 9 1734 1
テレファックス	+358 9 1734 2750
インターネット・ホームページ	<a href="http://www.stat.fi">http://www.stat.fi</a>
E-mail	firstname.lastname@stat.fi

## 訳者あとがき

1. 本資料は、フィンランドのレジスターベースの人口・住宅センサス（および GIS）に関する最近の以下の4つの論文と1つの小冊子を訳出したものである。

(1) Riitta Harala (1997), "Statistical Properties and Quality of register-based Census Statistics in Finland"

(2) Riitta Harala (1994), "Evaluation of the results of the register-based population and housing census 1990 in Finland" *Statistical Journal of the United Nations ECE 11*

(3) Riitta Harala and Marja Tammilehto-Luode (1999), "GIS and Register-based Population Census", from Juha Alho ed. *Statistics, Registries and Science-Experiences from Finland*

(4) Risto Lehtonen and Ari Veijanen (1999), "Use of Register Data to Improve the Estimation in a Sample Survey" from Juha Alho ed. *Statistics, Registries and Science-Experiences from Finland*

(5) Statistics Finland (1999), *250 Years of Population Statistics - History of Population Statistics*

(1) はレジスター・ベース統計生産の概略のイメージを与え、(2)はこのデータの品質とその評価法を語っている。この2つによって、読者はフィンランドでのレジスター・ベースでの人口センサスの全体像をおおよそ把握できるかと思う。(3)は、さらにレジスターベースデータの GIS との結合、さらに(4)は、標本調査における推計での活用を論じている。この2つによって、読者はレジスターベース統計の発展の可能性・方向の一端を見ることができると思う。(5)は、以上をふくむ、フィンランドでの、人口統計とセンサス関連事項の歴史的経過を簡条書的に要約したもので、便利であると思う。

2. 調査票を調査対象である世帯や事業所に配布し・回収する調査票ベースの統計生産の形態は、20世紀中において支配的であった。しかし、調査協力の低下、多大な費用、統計の正確性への疑義などの困難、そして情報技術の発展を基礎にして、1980年代あたりから、レジスターベースの統計生産への注目が高まってきた。この形態は、行政過程等で政府関係諸機関が獲得している登録等記録（レジスター）からの情報を収集して統計にまとめるというものである。

この形態が北欧諸国における採用、このシステムがふくむ問題、日本との違い等は、工藤弘安氏が1966年のはやくに若干の言及をされたのち、1980年代末から

一連の論考にまとめられまた学会等で発表されてきている【①(1966)「住民登録による統計作成の新しい可能性—スウェーデンの例—」『統計情報』15(12)、②「レジスター・ベースの統計制度」『研究所報』16、日本統計研究所、1987年、③「統計調査における情報提供(II)—事例研究 デンマークその一」成城大学『経済研究』108、1988年、④「レジスター・ベースの人口住宅センサス」成城大学『経済研究』127、1995年、⑤「統計調査における情報提供(III)—事例研究 デンマークその二」成城大学『経済研究』128、1995年、⑥「レジスター・ベースの統計」『グローバル統計の編成可能性の検討い』科学研究費・重点領域研究成果報告書】。その他さらに、西村善博氏がフランスでの行政データ利用を紹介している【①科学研究費重点研究「統計情報活用のフロンティアの拡大」「マイクロデータ利用の社会制度上の諸問題」資料No.7、日本統計研究所、1996年6月、②翻訳・ジャン=ルイ・フォル「労働者と企業に関する統合ソースを作成するための行政データ利用」『大分大学経済論集』50(6)、1999年6月】。

1980年代末から1990年代の国際統計界においては、レジスター・ベースの統計生産への関心は一層広がり、研究や諸会議での論議がひんぱんに行われている(Pekka Myrskyala(1991), "Census by Questionnaire- Census by Registers and Administrative Records: The Experience of Finland", *JOS*, 7(4)のち、*Statistical Journal of the United Nations ECE* 1995の特集、*JOS* や *International Statistical Review* 誌等での取り上げ他、カナダやEU他でのワークショップやカンファレンス等々)。ごく身近なところでは、1999年10月30日に、本日本統計研究所が開催した「2000-2001年世界人口センサスに関する国際ワークショップ」において、ドイツからのDieter Bierau氏が、ドイツでの今後のセンサス実施では、レジスター・ベースが強く意識されていることを報告されている(日本統計研究所『研究所報』No.26、2000年1月参照)。こういった状況の中で、フィンランドは、デンマークについて、レジスター・ベースの統計生産への移行に努め、人口センサスでは、1990年センサスで全面的に依拠することになった。

3. 折から1999年8月10~18日の第52回国際統計協会(ISI)総会は、ヘルシンキ市で開催され、フィンランド統計局の統計動が大きく紹介された。ひとつには、参加者に対してフィンランド統計局の統計活動・統計研究に関わる15の論文を集めたJuha Alho ed. *Statistics, Registries and Science - Experiences from Finland* が配布された。もうひとつには、フィンランド統計局へのツアーが組まれ、概要の説明とともに質疑と局内の見学会を通じてである。統計局長は、統計局の活動が認められて、統計局の予算が増加していることを誇っていた。その際示された予算総額の変化が、人口センサス実施年にも非センサス年と変わりなく右上が

りの直線状態であることが非常に印象的であった。これはフィンランド統計局の統計生産がレジスター・ベースに移行していることからきている。本資料の解説の執筆者は、改めて局を再訪問して、4人のスタッフを持つジェンダー統計係と意見交換をする機会を持った。その折に、レジスター・ベースの統計生産にも関心があるので、上述論文集の Riitta Harala と Marja Tammilehto-Luode の論文に掲げられている文献の入手を依頼した。その後、送られてきたのが、(1)と(2)である。(2)は国連ヨーロッパ経済委員会・統計雑誌のいわば行政データ利用の4論文を特集している号に掲載されている。この内容は1993年のフイレンツイエでのISI総会の Invited Paper において示されており、工藤氏がすでに取り上げられている。(1)は比較的最近のもので、端的に概要を伝えるものであるが、その中には、レジスター・ベースの場合には、いわゆるロンジチュージナル・データが、調査票ベースからマイクロ統計データ・ファイルを作り出すよりも簡単に作成されることが示されている。(3),(4)は最近の状況を伝えている。(3)は、地理情報システム(GIS)がレジスター・ベースによって、一層豊富になることが論じられている。フィンランドでは、地理情報システムは、早くから発展しており、その成果である詳細な各種のカラフルな地図は非常に興味深いものがある。(4)は、上述の配布論文集の第3部「レジスターベースによる研究」に、「死亡における社会的不平等－死亡率研究におけるリンクされたレジスター」、「フィンランドの癌レジスター－統計生産と研究」、「フィンランドにおける健康研究におけるルーテインとして収集されるレジスター」の3つの論文とともに収録されているものである。レジスターからの補助的情報の活用によって標本からの推計の効率の改善に関して、国際的に現在進行中のテクニカルな議論にあてられている。フィンランド統計局が、統計手法の研究・開発にも力を注いでいる一端を示している。(5)はISIの会場で配布された小冊子である。上述論文を訳す過程で、この小冊子も、本統計参考資料に収録しておけば、特に歴史的経過の中にこのレジスタ・ベースの統計生産を位置づける上で、あるいはフィンランド統計史、フィンランド統計局の理解のために何かと参考になると考えた。以上によって本資料は、フィンランド統計局の活動の大きな特徴の1つであるレジスター・ベースの統計生産の内容を伝えようとした。

4. 政府の行政姿勢への国民の信頼が厚く、統一的識別番号に大きな抵抗を感じずに、多くのレジスターを作り上げていると見える北欧の体制と、例えば日本の分散型を骨格とする統計制度・統計状況とは、かなりの差異がある。筆者は、レジスター・ベースの統計生産への移行は、21世紀の遅くない時期に、多くの国で進む事態ではないか、と予想している。その際の大きな問題になるのは、統一的識別番号の採用、関連して個人のプライバシー保護やデータの秘匿性保持、そして昨今、サイバー・テロという形で警鐘が鳴らされた電子的生活の危うさ問題をふくむデー

タの安全性の保障、である。これらについては、論文(2)で簡単な言及があるだけである。社会的背景の大きな違いを見ながら、レジスター・ベースへの移行過程で、また現在体制のもとで、諸問題がどう措置されたか、されているか、についてさらに情報を得ること、日本をふくめて将来的にどのような手だてが必要なのかを、今後検討していく必要がある。

5 最後にこの翻訳に際して、文献を送付していただき、また翻訳を許可いただいた Riitta Harala 氏、Risto Lehtonen 氏、間にたつていただいた Eeva-Sisko Veikkola 氏、そしてフィンランド統計局の関係者の方、に深く感謝したい。また、同僚でフィンランドに詳しい法政大学経済学部の田淵洋氏には、地理情報システムの文献的成果を見せていただく他、幾つかのサポートをしていただいた。お礼を申し上げます。

6 本資料でとりあげた文献の翻訳は、2. を千葉敦志（法政大学大学院経済学研究科、博士課程）、4. を小野寺剛（法政大学大学院経済学研究科、博士課程）、1, 3, 5 を伊藤陽一が担当した。

2000年1月

伊藤陽一

## 統計研究参考資料（最近発行分）

- |    |  |          |
|----|--|----------|
| 45 | 国連（1984年）『女性の状況に関する統計と指標のための概念と方法の改善』  | 95/12/31 |
| 46 | フィリピンの農業統計                             | 95/12/31 |
| 47 | ロシアにおける統計制度・政策の改革（II）                  | 96/03/31 |
| 48 | 統一価格中国日本産業連関表 1985, 1987, 1990年        | 96/03/31 |
| 49 | 国連(1995年)『世界規模のジェンダー統計に関するワークショップ』（翻訳） | 96/04/30 |
| 50 | 英国統計制度関係資料                             | 97/03/31 |
| 51 | インストローとジェンダー統計                         | 97/06/30 |
| 52 | アメリカ合衆国労働省労働統計局『製造業生産労働者の時間当たり報酬』      | 97/07/31 |
| 53 | わが国における外国人の死亡特性                        | 97/12/15 |
| 54 | アメリカにおけるマイクロデータの提供                     | 98/01/15 |
| 55 | 民間統計ガイド                                | 98/03/15 |
| 56 | 韓日産業別購買力平価の推計                          | 98/04/15 |
| 57 | カールステン・シュターマー他『1990年物的産業連関表』（翻訳）       | 98/05/15 |
| 58 | 各国統計関係法規集                              | 98/11/20 |
| 59 | 産業・職業クロス表による全国及び都道府県の死亡分析              | 99/02/15 |
| 60 | 合衆国センサス局「所得と政策参加」1993年パネル調査票           | 99/11/20 |
| 61 | 「統計の品質」をめぐる一翻訳と論文一                     | 99/12/20 |
| 62 | 合衆国BLS：国際比較諸統計一翻訳一                     | 99/12/20 |

統計研究参考資料 No.63

フィンランドにおけるレジスター・  
ベースの統計生産一翻訳一

2000年 1月25日

発行所 法政大学日本統計研究所

194-0298東京都町田市相原町4342

Tel. 042-783-2325, Fax. 042-783-2332

E-mail: jsri@nt.tama.hosei.ac.jp

発行人 伊藤 陽一