

統計研究参考資料

No.105

統計の品質(7)
翻訳：フィンランド統計局
政府統計の品質ガイドライン

2010 年 1月

法政大学日本統計研究所
Japan Statistics Research Institute
Hosei University

はじめに

本資料は、この統計研究参考資料シリーズが、特にこの10年来とりあげてきている「統計の品質」の7冊目として発行する、フィンランド統計局の「政府統計の品質ガイドライン」(*Quality Guideline for Official Statistics*)の第2版(2007年)の翻訳である。

これまで本シリーズでとりあげた統計の品質関連の号は以下のとおりである。

61	統計の品質—「統計の品質」をめぐって—翻訳と論文	1999. 12. 20
79	統計の品質(2)—「統計」の品質をめぐって(2)	2002. 09. 17
89	統計の品質(3)—国際統計機関における統計の品質	2005. 09. 30
93	統計の品質(4)—IMF品質サイトとQ2004を中心に	2006. 07. 25
97	統計の品質論(5)—Q2006と2006サテライト会議から(翻訳と関係論文)	2007. 05. 31
102	統計の品質論(6): 論文と翻訳—ESSの統計品質論と実践	2009. 08. 31

ここでいう「統計の品質」とは、統計データの正確性等に関して、旧来の統計学等で主として注目されていた「統計の誤差=標本誤差+非標本誤差」論議から大きく転換したものである。それは、市民をふくむ広い統計利用者の必要・要請を徹底的に重視して、統計データが基礎にする概念・分類等が利用者の必要に見合っているか=「適合性(Relevance)」をはじめとして、利用者が統計を入手しやすいか=「アクセス可能性」や統計データが利用者にとってわかり易いか=「明瞭性」や、適時にそして全利用者に向けて同時に統計が公表されているか=「適時性」、などの品質基準にそって、統計データの継続的改善をめざす論議とその実践である。

この論議と実践の基礎には、政府統計の基本原則や専門家の倫理などの現代に発展させた諸規範がある。この点をふくめて統計の品質論は、政府統計活動の在り方を根本的に規定することはいうまでもなからう。新統計法を定め、第一次の統計基本計画を2009年の冒頭に定めた日本の統計改革の具体化過程で、しっかりと組み入れるべき論議と実践である。

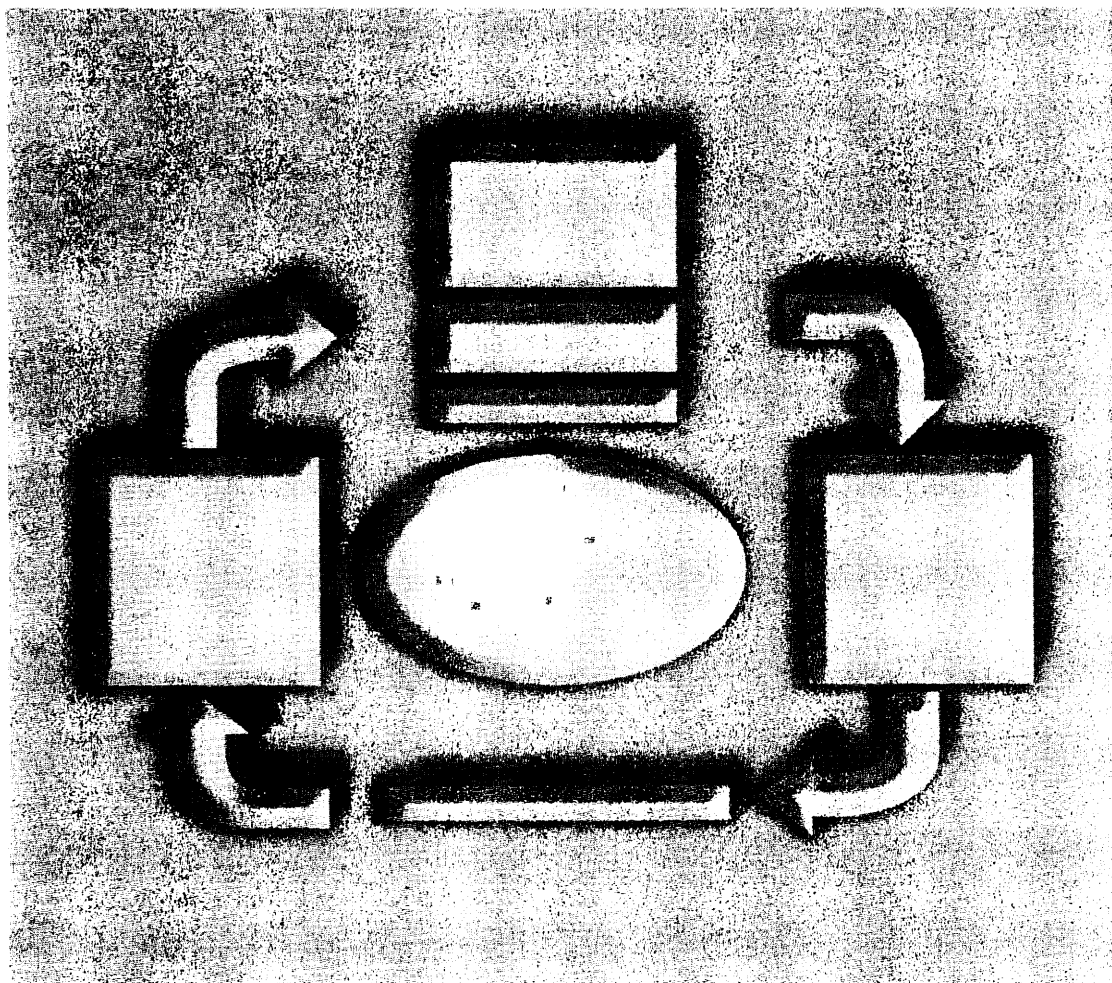
1990年代後半から国際的広がりをみせ、2000年代に加速しているこの統計の品質論議について、本研究所の上記資料は、この国際的活動の2つの推進体であるEurostatとIMFの論議、およびヨーロッパ統計品質会議、さらに、ひとまず体系化したようにみえるヨーロッパ統計システムについて紹介・論評してきた。国際的な全体的動向を中心にしてきた。しかし、日本での品質論の浸透や発展の必要を配慮すると、統計の品質に関する特定先進国に立ち入ることも必要と考えた。フィンランドのガイドラインをとりあげた理由である。

なお本資料全体の目次はつけていない。翻訳部分と138ページからの訳者論評からなる。

2010年 1月10日

法政大学日本統計研究所

Quality Guidelines for Official Statistics 2nd Revised Edition



政府統計のための品質ガイドライン

改訂第**2**版

Quality Guidelines for Official Statistics

2nd Revised Edition

フィンランド統計局
ヘルシンキ **2007**年

ハンドブック

問い合わせ

Johanna Laiho

Pentti Pietilä

Kari Djerf

(09) 17 341

Laatua.tilastoissa@tilastokeskus.fi

表紙図 Johanna Laiho

レイアウト Eva-Lisa Repo

© 2007 フィンランド統計局

出所がフィンランド統計局であることが示されていれば、引用は奨励される。

ISSN 0355-2063

= Handbooks

ISBN 978-952-467-743-1

Multiprint Helsinki 2007

序言

統計情報は多くの段階をなす複雑な過程の最終生産物である。統計の生産は、理論的および実際的手続きの集まりである統計システムの方法論に基づいている。統計専門家は、これらの手続きから適切な選択と適用をしなければならない。さらに、統計専門家は、利害関係者に対して使用した方法を明確に、そして包括的に説明できなければならない。政府統計の品質ガイドラインのこのハンドブックは、統計生産と利害関係者との相互関係の発展を支援することを狙っている。先行版と同じように、ハンドブックのこの版もまた、顧客と職員訓練で使われるだろう。

統計の生産過程は、われわれの活動の中核である。フィンランド統計局は、この生産過程、統計システムのデザインと実施に結びついた技能と知識に満ちている。フィンランド統計局の「政府統計のための品質ガイドライン」の中心目的は、既存の知識と現在の最善の方法を集めて局の共通の能力資産にまとめることである。機関全体の中での、この資産の共同の利用と生産過程の調整は、生産性を向上させて統計の品質を継続的に改善するための土台をなす。

このハンドブックは4つの主題についての章からなる。最初の章は、統計が生産される際の作業枠組み、政府統計を規制する規範、および普遍的な品質要請を検討する。第2章は、統計生産と調査過程、データの収集と処理から、統計、パラメーターと指標に至る機能的つながりにおける異なる段階に焦点をあてる。この目的は、過程を全体としてあばき、同時にこの段階が相互に関連しながら補い合う仕方を説明することである。統計調査を文書化する重要性を強調するために、第3章全体はこの主題にあてられる。文書化は、各統計専門家と各作業段階の活動のアプローチの不可欠の要素でなければならない。第4章は、統計情報の公表と配布、統計データを使って顧客に対して提供されるサービスを検討する。この章は、政府統計における顧客本位や顧客サービスの継続的な開発の重要性を強調する。このハンドブックは、巻末のアルファベット順の索引によって、さらに使いやすくなっている。

この書物は英語のハンドブックの改訂第2版であり、フィンランドの統計システムと作業枠組みに関心のあるすべてのわれわれの国際的パートナーや顧客に向けたものである。品質ハンドブックは絶えざる改訂を必要とし、この電子版は統計分野での継続的に変化しつつある作業環境と今日の増大する情報ニーズを考慮に入れて、更新され発展するだろう。

このハンドブックの殆どすべての材料は、名前をあげた責任者がひきいる小さなワーキング・グループにおいて編集された。ハンドブックのこの新しい版は、以前の第1版に基づいている。22名の著者と15名のコメンテーターが、ハンドブックのこの改訂版に貢献している。フィンランド統計局の無数の従業員がこのために貴重な原資料を提供し、著者たちに適切な専門家を紹介し、専門家は作成過程の異なる段階で貴重なコメントを提供した。これらの人々のすべてが、その熟知した投入と実際的コメントの点で心から感謝したい。

Carl-Erik Särndal教授 に対しては、本書の完成に向けての、作業の種々の段階での継続的な激励と有益なコメントに特別の感謝をしたい。特に、英語版の作成にあたっての彼の助言は、は

かり知れなく貴重であった。

この編集の主たる責任は、上級研究員 (Senior Researchers) の Johanna Laiho にある。さらに、編集作業とプロジェクト全体の進行を支えたのが編集助言委員会の 上級助言者 (Senior Advisers) Pentti Pietilä , Markku HuttunenとKari Djerf, およびRisto Lehtonen教授であり、彼らには、このプロジェクトの完成への知識豊かで継続的な作業、および統計の品質の多方面にわたる発展に向けて専門家の大きな集団を鼓舞してくれたことについて心からの感謝を表明したい。翻訳者のAila Hanley, Mia KilpiöとDavid Kivinenにも深く感謝したい。

ヘルシンキ, 2007年

フィンランド統計局長
Heli Jeskanen-Sundström,

著者

Faiz Alsu hail

Kari Djerf

Susanne Hellman-Ketola

Markku Huttunen

Riitta Kaisio

Janika Konnu

Vesa Kuusela

Johanna Laiho

Timo Laukkanen

Risto Lehronen

Jussi Melkas

Sirkku Mertanen

Heli Mikkela

Perttu Muurimaki

Pekka Myrskylä

Antti Pasanen

Oentti Pietila

Ritta Poukka

Anna-Leena-Reinikainen

Antti Suopera

Tuula Taivainen

Tuula Viitaharju

目次

	序言	3
	目次	5
1	全般的作業枠組み	6
1.1	政府統計の規範	6
1.2	統計の品質	16
2	統計調査と統計の生産	31
2.1	統計調査の定義	31
2.2	データ保護	37
2.3	分類、概念、定義と調整	45
2.4	枠組みとカバレッジ	49
2.5	行政記録とレジスター	53
2.6	標本抽出法	61
2.7	データ収集法	67
2.8	調査票の設計とテスト	73
2.9	調査への回答	77
2.10	統計的エディティングと補定	79
2.11	ウェイト作成と無回答の調整	85
2.12	統計的推定と分析	90
2.13	統計データの表示	94
2.14	時系列と季節調整	101
2.15	指標と指数	105
3	文書化	113
3.1	統計調査の文書化	113
3.2	文書保存	117
4	統計利用の促進、出版、配布と顧客サービス	120
4.1	統計利用の促進	120
4.2	出版	122
4.3	顧客サービスにおける品質	127
	索引	130
	略号のリスト	134

* 訳者注: ページ数は原書のもの。原本ページ数マイナス3ページが本翻訳資料のページにはほぼ対応する。

1 全般的作業枠組み

1.1 政府統計の規範

1.1.1 法律

フィンランド統計法 (280/2004) は、全ての政府統計に適用される国の一般法である。その目的は、国内と国際の両方のデータニーズを考慮した効率的で統一された政府統計の生産に向けての前提条件を創り出すことである。統計法は、データの提供者と利用者に適用される作業政策を作り出す明確な信任を規定している。それは統計編集の異なる段階に関する規制を含み、特にデータ収集の合理化、データ提供者との関係、統計倫理とデータ保護の実施に焦点をあてている。

統計法は、統計は、行政レジスターといった他の脈絡で収集されたデータを最初に利用することによって編集されなければならないと述べている。データの提供者は、他の手段で得ることのできないようなデータの提供を直接的に求められるだけであろう。法律で指定されているデータの提供の法的義務は、国家機関、公共サービス生産者、商業的取引あるいは職業の実施者、自治体あるいは国家の会社、合同自治体参事会(joint municipal board)および非営利機関に関してのものである。データ提供の義務を負う機関の代表者は、データ収集の新しいか改訂されたラウンドについて前もって、十分な協議を受けるべきであり、(国の政府を除いて) 統計の形でのフィードバックをもらうべきである。

統計法はまた、優れた統計実践、国際勧告、および一般的に適用された方法に従うべきことを要求している。それらの規制は、統計倫理の原則に従うことに言及している。

統計法は、統計の品質に関する節をふくんでおり、これに従えば、統計は可能な限り信頼できるものであり、社会の諸状況とその発展を正確に描写しなければならない。統計は完成後できるだけ速やかに公表されなければならない。結果とされる統計データの比較可能性を保証するためには、調整された概念、定義および分類が、可能な場合にはどこでもそれらの編集物で使われるべきである。

ヨーロッパ共同体の統計に関する理事会指令(EC) No. 322/97 が、ヨーロッパ共同体の統計プログラムにしたがって編集される統計に対して適用されなければならない。フィンランド統計法の中心的原則と共同体統計に関する協議会指令のそれとは、互いに一致している。

文献

ヨーロッパ共同体の法規 (European Community legislation)

Council Regulation (EC) No. 322/97 of 17 February 1997 on Community Statistics.

http://forum.europa.eu.int/irc/dsis/bmethods/info/data/new/legislation/stat_law.html

フィンランドの国家法規

Statistics Act (280/2004) -http://www.stat.fi/meta/lait/lait_statsiticsact04.pdt

1.1.2 専門的倫理

視野と目的

統計の分野での専門的倫理は、統計機関 (statistical authorities.) がその活動の基とする礎石をなしている。それらは、統計法規とその適用の基礎であり、また統計機関が採用する活動原則と活動方向に影響する。

統計分野での専門的倫理は、1985年に国際統計協会 (ISI: International Statistical Institute) が採択した専門的倫理宣言に集約されている。この宣言は、フィンランド統計局自身の専門的倫理に関するガイドラインの基礎になっている。欧州理事会はまた、統計分野の個別データの保護に関する理事会勧告 [Council Recommendation No. R (97) 18] を承認した。これはデータ保護法規だけでなく、専門的倫理についての上記の原則を基礎にしている。これに加えて、欧州委員会は、国および共同体の統計機関の独立性、高潔性 (integrity) および説明責任に関する勧告を採用した (COM217, 2005)。勧告の主な部分は行動規約である (1.13節参照)。

原則

統計的専門性の目的は、データに関わりにある人々のプライバシーを尊重しながら、社会についての偏りのない、多様で、信頼性のある描写を与えることである。継続的で速やかに変化している作業環境は、新しい種類の倫理問題を生み出し、古くからの問題を浮上させる。専門的倫理についてのガイドライン (Statistics Finland, 2006) の目的は、これらの倫理問題の解決を促進することである。

ガイドライン

専門的倫理についてのガイドラインによれば、フィンランド統計局とその従業者は、次のような倫理原則を守らなければならない。

- 不偏性 (impartiality) は、統計生産の客観的および独立した仕方を意味し、政治的その他の利害関係集団からのいかなる圧力から免れていること、特に、定められた目的の達成に最も適合した技法、定義および方法の選択に関して語られる。それは、すべての利用者に対して、遅れを最小限にして、統計が利用可能にすること意味する。
- 信頼性 (Reliability) は、統計が表現しようと企画された現実を可能な限り忠実に反映している統計の特徴を意味する。それは、出所、方法および手続きの選択のためには科学的基準が使用されることを意味する。カバレッジ、方法、手続きおよび出所についてのあらゆる情報がまたデータの信頼性を改善するだろう。
- 適合性 (Relevance) は、統計の生産は、目的によって決定される明確に定義された要請の関数であることを意味する。それらの要請が、分野、適時性および統計の大きさ (scale) を決定するが、それらは、あらゆる時に、新しい人口的、経済的、社会的および環境的開発と平行を保つべきである。データの収集は、望まれる結果を達成するために必要なことに限られるべきであ

る。目的にとって関心を失った統計の生産は断念すべきである。

- 費用効率性（Cost-effectiveness）は、すべての利用可能な資源の極大の利用と回答者負担の最小限化を意味する。統計生産が必要とする仕事量と費用は求められる結果/便益の重要性と釣り合っているべきである。
- 統計的秘匿性（Statistical confidentiality）は、直接的に統計目的で、あるいは行政的あるいはその他の出所から獲得された単一の統計単位に関するデータを、秘匿性の権利のあらゆる違反から保護することを意味する。それは、獲得されたデータの非統計的利用や不法な開示の防止を意味する。
- 透明性（Transparency）は、法的基礎、データが要求される目的と採用されている防御措置に関する情報を、回答者が持つ権利を意味する。統計の収集に責任を持つ機関は、そういった情報を提供するためのすべての手段をとるものとする。

EUの理事会指令No.R(97)18もまた、そのすべての加盟国も国家統計機関をしるる対応する原則をふくんでいる。

文献

ヨーロッパ共同体の法規，国際的ガイドライン

European Community legislation , international guidelines

Commission of the European Communities. 2005. Recommendation of the Commission on the independence , integrity , and accountability of the national and Community statistical authorities. COM(2005) 217.

http://eur-lex.europa.eu/LexUriServ/site/en/com/2005/com2005_0217en01.pdf

Council Recommendation No. R (97) 18 and Explanatory Memorandum of the Committee of Ministers to Member States Concerning the Protection of Personal Data Collected and Processed For Statistical Purposes.

Council Regulation (EC) No. 322/97 of 17 February 1997 on Community Statistics.

http://forum.europa.eu.int/irc/dsis/bmethods/info/data/new/legislation/stat_law.html

International Statistical Institute (ISI). 1985. Declaration on Professional Ethics.

<http://isi.cbs.nl/ethics.htm>

フィンランド国内法

Statistics Act (280/2004).

http://www.stat.fi/meta/lait/lait_statisticsact04.pdf

フィンランド統計局のガイドライン

Statistics Finland. 2006. Guidelines on Professional Ethics. Rev. ed. Handbooks 30b. Helsinki.

http://www.stat.fi/org/etiikka/eettinenopas_en.pdf

1.1.3 実践規約

視野と目的

統計の品質評価の際の客観性と比較可能性は、広い協力と相互に承認して適用された品質基準を必要とする。ヨーロッパ統計システム内の統計機関が生産し、配布する統計は、国のレベルと共同体レベルの両方で品質と信頼性に関する厳密な基準に見合わなければならない。この理由で、委員会はヨーロッパ統計実践規約を発行した。これはシステムが、情報化社会のニーズによりよく奉仕することができるように、システムの発展を導くものである。ヨーロッパ統計システムの狙いは、効率的に機能して、独立性、高潔性および説明責任に関する要請の集まりを満たすことである。

実践規約は、自己規制的な装置である。というのは、それはヨーロッパ統計の主な生産者が準備し、支持したものである。実践規約は共同体統計の生産において観察されるべき15の原則からなる。実践規約は2つの目的を持つ。すなわち、

- (i) 国家統計機関とEurostatの両方の独立性、高潔性および説明責任への信用と確信、そしてそれらが生産し配布する統計の信用性 (credibility) と品質を改善すること、
- (ii) ヨーロッパ統計の品質を高めるために、そのすべての生産者による最善の国際原則、方法と実践の適用を促進すること。

実践規約はESS品質宣言 (Statistical Programme Committee, 2001) に基づいて作られた。宣言は以下を示している。

ヨーロッパ統計システムの使命

「われわれは、ヨーロッパ連合と世界に対して、ヨーロッパ、各国および地域レベルでの経済と社会についての高品質の情報を提供し、その情報を意思決定の目的、研究および討論のためにすべての人に利用可能にする。」

ヨーロッパ統計システムのビジョン

「ESS は、統計情報活動で世界のリーダーそしてヨーロッパ連合とその加盟国に対する最も重要な情報の提供者になる。科学的原則と方法に基づいて、ESS は民主的過程や社会の前進にとって不可欠の基礎をなす調整されたヨーロッパ統計のプログラムを提供し、継続的に改善する。」

ESSの品質宣言は、ヨーロッパ統計システム内で、EFQMモデル (EFQM, 2003) に沿って総合的品質管理アプローチの発展に向けたガイダンスを提供している。EFQMモデルを採用することはまた、品質に関するリーダーシップグループ (LEG) の勧告 (LEG on Quality, 2000) で支持されてきた。LEGの勧告の実践に向けてのガイダンスはLEG実施ワーキング・グループ

(European Communities, 2002) で提供されている。

原則

実践規約は、統計の生産者が、ヨーロッパの品質基準に見合い、国と、ヨーロッパの機関、政府、研究機関、関連するビジネス界や一般大衆の両方のニーズに奉仕する点でガイドとなる。統計の品質は、その適合性、正確性と信頼性、適時性、整合性および地域と国の間での比較可能性によって決定される。これに加えて、統計の品質が評価される時には、容易なアクセスが不可欠である。高い品質の統計は、それらの基準にそって、あらゆる利用者のニーズに見合うものでなければならない。

実践規約の原則は、制度的環境、統計プロセス、および統計生産物にわたる3つの分野に区分されてきた (COM217, 2005)。

I 制度的環境

原則1 専門的独立性—統計機関が他の政策、規制、行政部門や機関から、また民間部門の取扱者からの専門的独立性は、ヨーロッパ統計の信用を保証する。

原則2 データ収集のための義務—統計機関は、ヨーロッパ統計目的のための情報の収集のための明確な指令(mandate)を持たなければならない。行政、企業、世帯および一般公衆は、統計機関の要請によって、ヨーロッパ統計目的のためのデータへのアクセスあるいは配布を許すことを法律によって強制されることがありうる。

原則3 資源の十分性—統計機関が利用可能な資源は、ヨーロッパ統計の要請に見合う十分なものでなければならない。

原則4 品質公約—すべてのヨーロッパ統計システムへの参加国は、ヨーロッパ統計システムの品質宣言に定められた原則に従って自ら活動し、協力する公約をしている。

原則5 統計的秘匿性—データ提供者(世帯、企業、行政その他の回答者)のプライバシー、彼らが提供する情報の秘匿性、統計目的のためだけの使用は、絶対的に保証されなければならない。

原則6 公平性(Impartiality)と客観性—統計機関は、ヨーロッパ統計を、科学的独立性、客観的、専門的、かつすべての利用者が同等に扱われる透明な方法で、生産され、配布されなければならない。

II 統計的過程

原則7 堅実な方法(Sound Methodology)—堅実な方法は統計の品質を支える。これは、十分な道具、手続き及び専門性を必要とする。

原則8 適切な統計手続き—データの収集から確認までに採用されている適切な統計手続きが、統計の品質を支えなければならない。

原則9 過重でない回答者負担—報告負担は、利用者のニーズに釣り合っており、回答者に過重な負担となってはならない。統計機関は回答者負担を監視し、時とともにそれを削減するため

の目標をたてている。

原則10 費用効率性—資源は有効に使用されるべきである。

III 統計生産物

原則11 適合性—ヨーロッパ統計は利用者のニーズに対応しなければならない。

原則12 正確性と信頼性—ヨーロッパ統計は正確かつ信頼できる形で、現実を描くべきである。

原則13 適時性と時間厳守性—ヨーロッパ統計は適時的かつ時間厳守的な形で配布されるべきである。

原則14 整合性と比較可能性—ヨーロッパ統計は、時間の経過においても内部的に一貫しており、地域と国の間で比較可能であるべきである。異なる出所からの関連するデータを結合し、つなぎ合わせた利用が可能であるべきである。

原則15 アクセス可能性と明瞭性—ヨーロッパ統計は、明瞭で理解可能な形で示され、適切で便利な形で配布され、支援となるメタデータとガイダンスを伴っており、公平な基準で入手可能でアクセス可能であるべきである。

ガイドライン

高い品質の統計生産物を保証するためには、統計の生産者は、その活動と目標を、実践規約にしたがう実現を定期的に監視しなければならない。統計の高い品質を維持するために、その生産は、利用者のますます骨の折れる、複雑な情報に見合うために継続的に発展しなければならない。

実践規約は、その原則の各々について指標を規定している。統計機関は、それらを固守し、それらに対応して作業環境の発展に貢献しなければならない。

統計品質の基準に見合うためには、統計生産者は、実践規約（COM217, 2005）に規定されている以下の指標を考慮しなければならない。

1. 適合性

- 利用者との協議し、既存の統計が利用者のニーズに対応する点での適合性と実際の効用を監視し、利用者の新たなニーズと優先度について助言する過程が整っている。
- 優先度のニーズが満たされ、活動プログラムに反映されている。
- 利用者満足度調査が定期的に行われている。

2. 正確性と信頼性

- 原データ、中間的結果および統計的生産物が評価され、確認されている。
- 標本誤差と非標本誤差が、ヨーロッパ統計システムの品質構成要素の枠組みに応じて測定され、体系的に文書化されている。
- 改訂の研究と分析が定期的に遂行され、統計過程を通知するために内部的に使用されている。

3. 適時性と定時性

- 適時性はヨーロッパおよび国際的な最高の配布基準にそっている。

- ユーロッパ統計の公表に向けて基準になる日の時間が定められている。
- ユーロッパ統計の周期は可能な限り利用者の要求を考慮している。
- 配布時間の予定からのいかなる乖離も前もって公表され、説明されて、新しい発表日が定められる。
- 受け入れ可能な総合的品質をもつ予備的結果が、有効と考えられるときには配布されてよい。

4. 整合性と比較可能性

- 統計は内的に整合的かつ一貫している（例えば、算術的および計算的同一性が認められる）。
- 統計は合理的な期間にわたって、整合的でありかつ調整されている。
- 統計は、異なる調査や出所の範囲、定義、単位および分類に関して、共通の基準に基づいて作成されている。
- 異なる調査や出所からの統計は比較され、調和されている。
- データの国家間の比較可能性が、ヨーロッパ統計システムと他の統計システムとの間の定期的交換を通じて保証されている。方法論的研究が加盟国とEurostatとの間の緊密な協力の下に遂行されている。

5. アクセス可能性と明瞭性

- 統計は、適切な解釈や意味のある比較を促進する形で提供されている。
- 配布サービスは、現代的な情報・通信技術を、適切な場合には伝統的なハードコピーを使う。
- 実行可能な場合には、顧客の企画した分析が提供され、公表される。
- 研究目的のためにマイクロデータへのアクセスが許されることがある。このアクセスは厳しいプロトコルに従っている。
- メタデータが標準化されたメタデータシステムにしたがって文書化されている。
- 利用者は統計的過程の方法とヨーロッパ統計システムの品質基準に関して統計的生産物の品質について常に知らされている。

文献

ヨーロッパ共同体の法規，国際的ガイドライン

Commission of the European Communities. 2005. Recommendation of the Commission on the independence, integrity, and accountability of the national and Community statistical authorities. COM(2005) 217. http://eur-lex.europa.eu/LexUriServ/site/en/com/2005/com2005_0217en01.pdf

Leadership Group (LEG) on Quality. 2000. LEG on Quality Recommendations.

Statistical Programme Committee. 2001. ESS Quality Declaration.

追加的情報

European Foundation for Quality Management (EFQM). 2003. The EFQM Excellence Model.Brussels.

European Communities. 2002. Quality in the European statistical system – The way forward.Eurostat

& European Commission. Luxembourg: Office for Official Publications of the European Communities.

Eurostat. 2006. The Eurostat Quality site.<http://ec.europa.eu/eurostat/quality> [referred to on the26.9.2006]

1.1.4 政府統計

視野と目的

政府統計は、定期的あるいは十分な頻度の間隔をもって生産され、全国を代表する、不可欠で品質の高い社会統計からなる統計体系である。入手可能な統計は、実際には利用者のニーズと優先度を反映し、統計の編集についての国際的基準と勧告に従わなければならない。多くの国において、政府統計は一般的には、異なる利用者集団、通常の市民、行政、経済界と研究界に奉仕すべきものとして定義される。

ほとんどの場合、政府統計は中心的トピックと社会の主題分野を通じて説明される。トピックスへの分割は、国連の統計系列の大分類—すなわち、人口・社会統計、経済統計、およびその他の統計、その下で、統計は副次系列あるいはアルファベット順、あるいはその他の定められた順序に従って区分される—に基づくことができる。系列の区分は、その中で生産される統計が大きな改訂を行ったとしても、相対的には安定するように設計される。

原則

政府統計を生産するためには信頼できる機関が必要である。一般的には、政府統計の編集は、特にこの目的のために設置された公的機関に委ねられるか、他の場合には、この仕事は異なる行政部門内の機関に分散されているかも知れない。殆どの場合、政府統計は、中央の調整機関としての国家統計局を持つ中央政府機関によって発行される。

政府統計の編集に適用される原則は、1994年に国連統計委員会で採択された「政府統計の基本原則」である。

- 1 官庁統計は、経済・人口・社会・環境の状態についてのデータを政府、経済界及び公衆に提供することによって、民主的な社会の情報システムにおける不可欠な要素を構成している。この目的のため、公的な情報利用に対する国民の権利を尊重するよう、政府統計機関は、実際に役に立つ官庁統計を公正にまとめ、利用に供しなければならない。
- 2 官庁統計への信頼を保持するために、統計機関は、科学の原理と専門家としての倫理を含む厳密に専門的な見地から、統計データの収集、処理、蓄積及び公表の方法及び手続を決定する必要がある。
- 3 データの正しい解釈を促進するため、統計機関は、統計の情報源、方法及び手続に関する情報を科学的基準に従って提示しなければならない。
- 4 統計機関は、統計の誤った解釈及び誤用に関して意見を述べる権利を有する。
- 5 統計を作成するためのデータは、統計調査又は行政記録などすべての種類のデータ源から入

手し得る。統計機関は、品質、適時性、費用及び報告負担の観点からデータ源を選定すべきである。

- 6 統計機関が統計作成のために収集した個別データは、自然人又は法人に関するものであるかによらず、嚴重に秘匿されなければならない、統計目的以外に用いてはならない。
- 7 統計システムを運用するための法律、規則及び諸手続は、公にされなければならない。
- 8 国内統計機関間の調整は、統計システムの一貫性及び効率性を達成するために不可欠である。
- 9 国際的な概念、分類及び方法を各国統計機関が用いることは、官庁のすべてのレベルの統計体系の整合性及び効率性を向上させる。
- 10 統計における二国間及び多国間協力は、すべての国の官庁統計システムの改善に寄与する。

ヨーロッパ連合の委員会の実践規約は、国家および連合の統計機関の独立性に関する自己規制的装置の基準をふくんでいる。規約は、独立性の原則に、さらに、統計の配布と入手可能性を特に言及した専門的独立性の概念を導入している。専門的独立性の原則は以下のように叙述されている。すなわち、

この原則の適用を監視する指標は以下をふくむ。

「専門的独立性—統計機関が他の政策、規制、行政部門や機関から、また民間部門の取扱者からの専門的独立性は、ヨーロッパ統計の信用を保証する。」

- (i) 独立性の要素の法律への明記
- (ii) 統計機関の長の地位と、統計方法、基準、手続きをふくむ機能的責任
- (iii) 統計公表の内容と時期の決定
- (iv) 統計活動プログラムの公表

政府統計は、社会の異なる部門の社会的・経済的状況と変化を叙述する。費用が大きい長期にわたる決定は、統計に基づいて行われることが多い。したがって、統計が信頼できることが不可欠である。この信頼は、統計の品質と客観性に依存する。

ガイドライン

品質への献身のサインは、品質改善手段、一般に承認された品質基準と道具の使用、品質分析、体系的な品質評価、政府統計の十分な文書化、および利用者に対する公開性と透明性、がある。利用者にとってこれらは、政府統計を生産する機関およびそのサービスへの利用者の確信が大き

く依存する専門的能力のサインでもある。

- 政府統計は、ある品質基準に合致しなければならず、その生産者は信用できて、顧客は情報に依存することができるものでなければならない。
- 信頼できる統計の生産の他にも、統計の編集者はまた、品質評価と報告とともに、その効率的な配布に責任がある。
- 品質報告は、一般的に承認されている品質指標にしたがって行われなければならない。ヨーロッパ統計システムの統計の品質構成要素には、i)適合性、ii)正確性と信頼性、iii)適時性と定時性、iv)整合性と比較可能性、v)アクセス可能性と明瞭性、がある。
- 利用者の信頼を獲得するためには、政府統計の編集者とその指導者は、その内部の作業文化において品質への堅実な献身を示さなければならない。

文献

Commission of the European Communities. 2005. On the independence , integrity , and accountability of the national and Community statistical authorities. COM(2005) 217.

United Nations Statistics Division. 1994. Fundamental Principles of Official Statistics. Supplement No. 9. Report of the Statistical Commission on its Special Session. New York 11.–14.4.1994. E/CN.3./1994/18 , E/1994/29.

追加的情報

Advisory Board of Official Statistics of Finland. 2006. Quality criteria for Official Statistics of Finland. (TK-41-1227-06).

United Nations. 1982. Directory of International Statistics. Vol. 1. Statistical Papers. Ser. M. No.56. Rev 1. New York.

United Nations Statistics Division. 2001. Handbook of Statistical Organization. The Operation and Organization of a Statistical Agency. 3rd ed. ST/ESA/STAT/SER.F/88 (Final Draft).

1.2 統計の品質

1.2.1 総合的品質管理 (Total quality management)

視野と目的

フィンランド統計局では、品質とは伝統的に、統計の品質、他の言葉でいえば、生産物の品質を意味してきた。フィンランド統計局の戦略的な総合的品質管理 (TQM) は、この生産物の品質の展望を捕うために定義されてきた。TQMの目的は組織のパフォーマンスの優秀性の達成であり、その中心的目標の1つは、品質表彰基準に見合う品質の達成である。同時に、品質表彰枠組みはまた、総合品質管理の近代的定義である。

TQMにおける焦点は、個々の統計生産物とその生産から、統計生産の全システムや他の中核的過程へ、そこから全体的管理システム、人的協力、および資源へ広がった。それらの諸側面の体系的管理によって、品質を強調し、そこからまた、最終生産物の品質を改善する作業システムを創り出すことが出来る。

良いパフォーマンスの能力は次のような特性を基礎にして建てられる (EFQM,1999)。すなわち、

結果重視 (Results orientation) :

- 優秀性は、すべての人々：従業員、顧客、データ提供者および社会一般のニーズとのバランスを持ち、満足させることに依存する。

顧客重点 (Customer focus) :

- 生産物、サービスおよび行動方向を計画するときに、現在および潜在的な顧客のニーズに明確に焦点をおく。

リーダーシップと目的の不変性 (Leadership and constancy of purpose)

- 組織の指導者の行動は、組織内の目的の明瞭性と統一、および組織とその人々が秀でることのできる環境を創り出す。

プロセスと事実による管理 (Management by processes and facts)

- 組織は、すべての相互に関係する活動が理解され、体系的に管理されるときに、より効果的に動き、現在の活動と計画された改善に関する決定が信頼できる情報を使って行われる。

職員の開発と参加 (Personnel development and involvement)

- 価値の共有と信頼とエンパワーメントの文化は、すべての者を激励し、組織の目標と職員の利害とを合流させる。

継続的な学習、革新と改善 (Continuous learning, innovation and improvement)

- 継続的な学習、革新および改善という文化の中での管理と知識の共有は、組織を最大限発展させ、他よりも優れたものになることを可能にする。

協力の発展 (Development of partnerships)

- 組織は、信頼と知識と統合の上に築かれた、パートナーとの相互に利益をもたらす関係を有するときにより効果的に動く。

公共的責任 (Public responsibility)

- 模範的な企業の市民権。

ISO 9000 の定義によれば、品質は、声明されているか暗黙かのニーズを満足させる能力を持つパフォーマンスをふくむ項目、生産物およびサービスの総ての特徴から構成されている。品質管理は、組織が品質を保証するためにとるすべての手段におよぶ。高品質の生産物とサービスの生産と供給は3つの要素からなる。

- **品質計画 (Quality planning) :** 顧客のニーズの定義、それらの最善の方法的解決への翻訳、および正確性の具体的レベルでの費用効率的なデータの収集・処理と配布を保証する適切な構造とシステムの創造によるそれらの導入。

- **品質保証/検査 (Quality assurance/verification)** : 顧客のニーズと計画に照らした最終生産物の品質の分析による生産物と生産の監視。保証は、生産過程内での個別的過程、活動および作業を、それらが適切に実施されていることを保証するために調べることをふくむ。
- **品質改善 (Quality improvement)** : 観察された品質問題の除去をめざした活動。それは、システムの全体的な更新あるいは再設計（矯正あるいは改善活動）をふくむ。-

品質評価は、大部分が設定された目標あるいは品質基準に関連づけられる。それらは、内部的に実施される自己評価でありうるし、あるいは評価は全部あるいは部分的に外部の団体に委ねることがありうる。評価は、質的あるいは数量的尺度あるいはそれらの結合に基づくことができる。

評価は通常、評価されるべき手続きの書面による記述に照らして行われている。この手続きは、活動が計画に合致しているか、効率的に実施され、記述された目的からみて適切であることを確認するために使用される。評価はまた、改善の尺度の状況を検討する。それらはまた、採用した手続きが最善の実践に照らしてどれだけ優れているかの評価もふくむ。

品質はまた、統計データの誤差の異なる発生源を調べることによって検査できる。リスクとその確率の大きさの確認もなおもう一つのアプローチである。得られた情報は、品質を改善するために統計の編集者が、そして、適切な解釈をするために、統計データの利用者が利用できる。.

原則

品質管理システム (Quality management system)

品質管理システムは、組織の活動のための共同で同意した座標となるような理論体系 (frame of reference) あるいは組織の活動である。それは、組織の外的および内的活動環境と組織の役割を輪郭づける。品質管理システムの狙いは、顧客の利益を守るために最善の実践を追求し、移し、維持し、発展させる体系的方法を創り出すことである。それはまた、望まれた作業モデルに透明性を与える努力をし、それらを組織全体についての知識に持ち込む。

品質管理システムは、すべての人に利益をもたらす、顧客と利害グループとの関係で組織の行動様式の改善を促進するように、人々の協力を導く行動様式を組織化することを試みる。それは、組織自体、顧客そして他の利害グループの見地から不可欠である専門家の間でのコミュニケーションを激励する。

品質管理システムは、組織が方法的アプローチ、現在の最善の方法、監視システム、継続的な改善、を構築し、総合するための前提を創り出す。それは、組織内のいわゆる格差シンδροームを、すなわち、そのヒエラルキーにおける機能と水平な顧客の方向での機能の間の分離を軽減し、組織内の論議のつながりと暗黙の知識の配分を改善する。

優れた品質管理システムは、行うべきことを確認することを助け、それらがどう行われるべきでありかについて、経験によって再定義されたアイデアを与え、パフォーマンス能力をさらに強めるために利用されうる価値あるフィードバックの知識を作り出す。

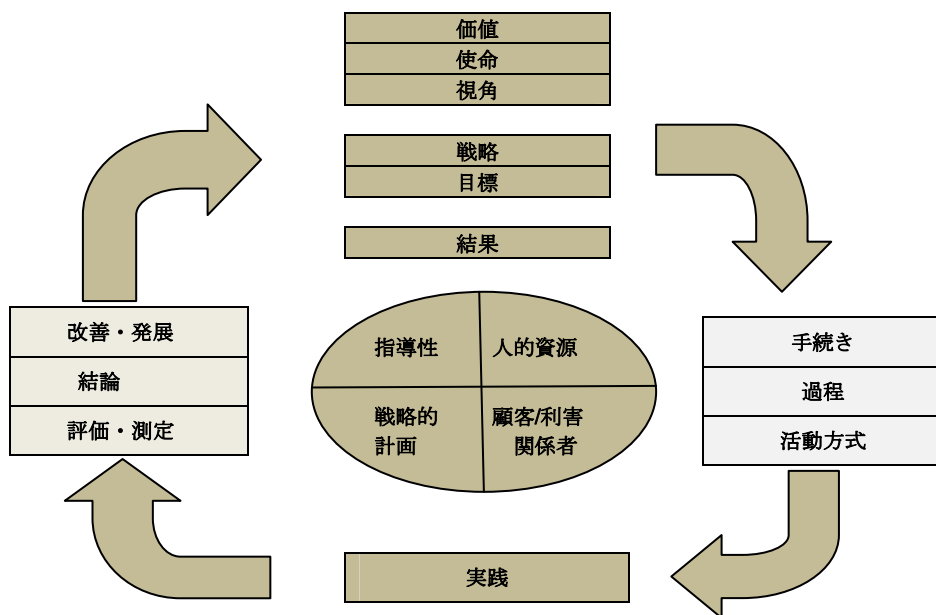
品質管理システムと一般の品質は、以下のように、**な組織**、**中核的過程**、あるいは**個別の過程**

の見地から検討することができる。

組織レベル

組織全体に影響を与える広い、戦略的な改善プロジェクトが、組織レベルで探求され、実施されるべきである。組織レベルの品質管理のための基礎は、強みと改善分野を探るための自己評価である。バランスある得点カード(BSC : balanced scorecard) は、コミュニケーションの道具であるとともに、戦略計画を創り出し、それを実施するための道具である。総合的品質管理システムは、それらの結合を描いている (図 1.1)。

図 1.1 総合的品質管理システム



中核的生産過程レベル

一貫した、費用効率的で、なだらかな生産過程と、関連するロジスティックは、中核的生産過程レベルの一つの中核的過程として発展・維持されるべきである。ここでの不可欠の諸要素は、過程の確認、定義、モデリング、文書化、評価、測定 (metering)、品質保証および改善である。操作的要素としての中核的過程は、統計方法、道具、データ、ソフトウェア・プログラム、メタデータ、マニュアルおよびハンドブックによって影響を受ける。

統計過程レベル

統計的調査は、中核的生産過程の副次的過程である。統計過程レベルでの検討は、個別の統計調査が最適に機能しているのかどうかを確かめるを試みる。しかし、トップの組織になることは、いくつかの個別的統計過程だけを改善することではなく、上述の全過程の発展を必要とす

る。

ガイドライン

- TQMの適用における礎石は、各人の別々の品質世界の進展を避けることである。狙いは、活動（品質）の総合的管理システムの創造であり、その中で、良い品質原則が統計生産システムに導入されることである。
- 組織全体に影響を与える重要な品質目標が戦略的目標に取り入れられ、それらが各活動の計画と監視の一部になることである。
- BSCは総合的品質管理への主な応用と測定の道具として使用される。
- 実践においては ISO 9000:2000 基準シリーズは、品質表彰とBSC枠組みと一致する品質システムの発展のための構造を提供する。
- 総合的品質思考においては、組織の活動は、その中核的過程とその副次的過程に助けられてモデル化される。
- 十分に動機づけられ、能力ある職員は、パフォーマンス能力を本質的に改善する組織の戦略的目標を理解し同意する。
- 変化を現実のものとするには、方向、確固とした基礎的知識と道具を必要とする。管理者は、方向を決定し、適切な（品質）道具の獲得を援助すべきである。

文献

European Foundation for Quality Management (EFQM). 2003. The EFQM Excellence Model. Brussels.

追加的情報

Anttila J. & Vakkun, J. 2001. Good Better Best. Tampere: Sonera Oyj/Tammer Paino Oy.

Anttila, J. & Vakkun, J. 2001. ISO 9000 for the Creative Leader. Tampere: Tammer Paino Oy.

1.2.2 統計の品質の評価

視野と目的

統計生産の品質の評価は、主として政府統計の品質と発展に関する国際的勧告や義務に基づく。国連、OECD、IMFやEurostatといった国際機関が、政府統計や国家統計サービスの品質についての目標を設定した。

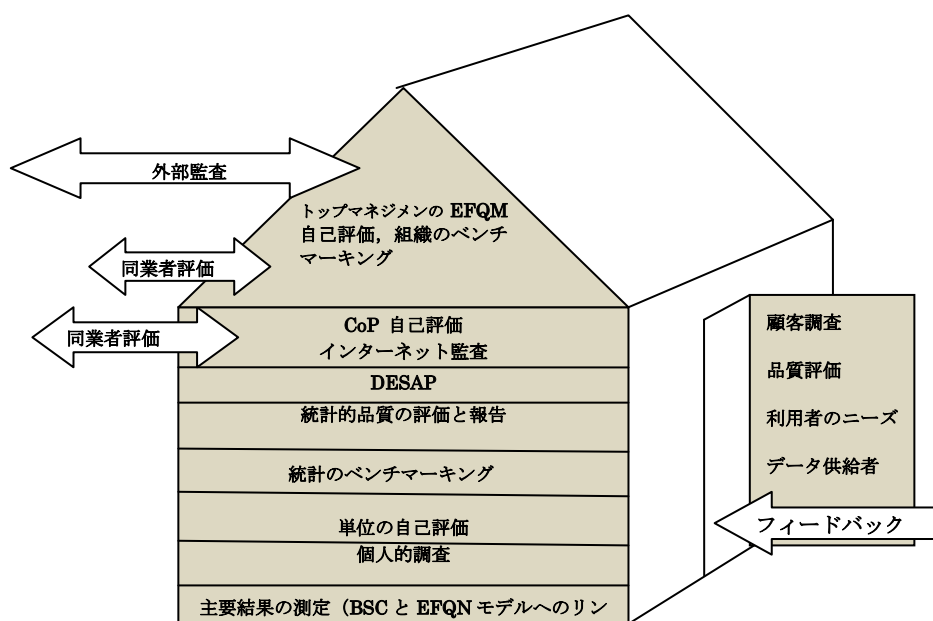
統計機関、その統計生産あるいはそれが創り出す統計を様々な視点から評価するために使うことのできる品質の測定のためには代替的道具がある。品質評価は、全組織のレベルで、個別統計の発展について、あるいは顧客の見地に焦点をあてることによって行うことができる。ほとんどの道具は管理者および組織の専門家が使うことができるが、あるものは、外部評価の採用を要求する。

品質評価は、統計生産の管理を支援するために直接的に利用され、開発に向けてのニーズの場を定め、成功を評価する。これらの出発点は、活動の発展、公開性および均一な手続きの利用である。幾多の評価の結合的利用は、組織と統計の生産の両方への交差する多様な見地をもたらす。異なる評価の結果の比較と分析は、見地の類似性と多様性ととも、得られた結果の信頼性と均一性（uniformity）を評価する助けとなる。品質評価の比較によって明らかにされた結果の均一性あるいは異質性は、機関の中の異なる関係者の見方によって、単一の評価で可能になる以上により完全な分析を可能にする。品質道具の結合した利用は、統計生産の品質の包括的発展にとって決定的な情報を作り出すことが多い。

原則

多くの機関は、高い品質とそのパフォーマンスの改善に向けて継続的に努力しなければならないという課題に直面している。成功のためには、組織はインフラストラクチャや資源のみならず、継続的な発展を促進する組織的文化を必要とする。これは、高い動機を持つ従業員、変革の管理者、透明性、効率的なコミュニケーションや、従業員が活動環境や実践を改善することの激励を必要とする。現在の新しい実績、開発のアイデア、対話と強い動機が、組織が求める発展

図 1.2 政府統計の品質の評価方法 (Laiho and Nimmergut,2004 から編集)



過程をさらに促進する (Laiho and Nimmergut,2004)。

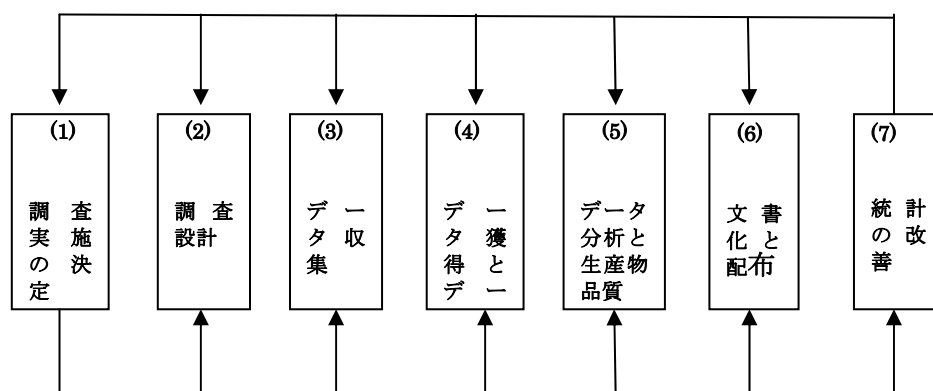
原則的には、すべての開発活動は、組織内の品質文化に基づく。新しい改善の示唆を示す責任や勇気は、信頼、尊敬、新しいアイデアに対する公開性に基づいて作られる。

DEAP¹あるいはEFQM²の自己評価、および内部の専門家による統計的監査のような自己評価の道具は、そういったアイデアを引き起こし、招き入れることを狙っている。それらはまた、組織内の対話を支える (Laiho and Nimmergut,2004)。

品質評価は大部分、標準、勧告あるいは目的との関連で行われる。それらは、組織自身によって、自らのあるいは外部の資源によって、あるいは二つを結合して行うことができる。評価は、品質および量的尺度および記述の両方に基づいて行うことができる。図1.2は、その統計生産を支えるために、ヨーロッパの国の統計機関とフィンランド統計局によって使われる品質評価法を描いている。

組織の管理の品質を評価のための広く使用され、標準化された手続きは、標準化された構造に従う、EFQMの自己評価として知られている (EFQM,2003)。その目的は、組織の強さと、改善が必要な分野を正確に指摘することである。それはまた、管理レベルでの組織間のインフォー

図 1.3 DESAP 自己評価の構造 (Körner and Nimmergut,2003)



マルな比較を可能にする。そういった比較の調整や焦点となる分野は、比較される機関の間での合意することができる。さらに、2005年にヨーロッパの国家統計機関は、CoP (実践規約) の自己評価を、2006年には、それに基づく同業者評価 (European Commission, 2005) を国家統計機関の間で、開始した。これらは、上述の機関の比較よりも、本質的により体系的であり、実践規約の内容から引き出された明確な構成にしたがっている(1.1.3参照)。

統計の品質に関する情報を顧客に伝える品質記述は、通常は、統計生産のための品質評価道具とみなされている。しかし、統計的監査 (Eiderbrandt,2004)や、統計生産に適切なDESAP自己評価といった、品質への実際の貢献者に焦点を置く品質評価法が、総合的品質の点からはより重要である (Korner and Nimmergut, 2003)。

監査の原則は、総合的品質管理に基づいている。それは個別統計の生産に集中する。それは、各部分の統計についての最善の専門家は、当該の統計部門に居ること、しかし、内部的な同業者

¹ DESAP(Development of a Self-sessment Programme)
² EFQM(European Foundation for Quality Management)

評価のためには、他の統計部門の専門家をあてることができることを認めている。この活動の目的は、外部的視角を利用して、発展を必要とする統計分野を確認し、それを優先し、統計における質的改善を達成するための関連手段を実施すること、にある。この評価は、評価手続きの書かれた方法論的記述と自己評価に基づく。そういった評価は、活動が計画どおりであるか、効率的に実施されているか、一連の狙いの点から適切であるかどうかを確認するために使用される。それらは、採用された手続きが、ベスト・プラクティス都の関連でそれだけ優れているかを検討する。監査はまた、組織全体へのグッド・プラクティスの広がりを探求する。

DESAPの自己評価は、統計生産の必要に合うように行われる。他の専門家が使われる統計的監査と違って、DESAPは、一定の統計の集まりに責任を持つ者が行う。このようにして、それは、外部資源抜きに、関心ある統計について、その品質の改善を支援するために、その生産日程に適切に独自に時期設定して実施することができる。DESAPは統計の生産に焦点を当て、特にヨーロッパ統計システム（ESS）の範囲内で、統計の体系的品質評価のために設計された。このように、組織内の総合的品質の一般的自己評価とは異なる。DESAPは統計の品質に寄与する生産の要因をとりあげる。それは、ESSの品質要求に対応し、統計データの品質に関連する主な側面からなっている。チェックリストとして、それは、トピックあるいは特定の研究方法に関係なく一マイクロデータが収集されるすべての個別統計に適用可能であるという意味で普遍的である。それは、研究に責任を持つ人々にとって、統計の品質の評価と改善の手段の計画の際の助けとなるための1つの道具として設計されている。対応する自己評価は、過去には、例えば、1980年代および1990年代の転換時のフィンランドにおけるように、国のイニシャチブに基づいて行われた。DESAPは、すべてのヨーロッパ統計システムの改善と調整のための調和のとれた道具として作成された。

統計の生産を改善するために行われる活動において、データの提供者と顧客から受け取ったフィードバックを正当に考慮することが是非とも必要である。データ提供者の回答負担を最小にすること、データ収集の調整、およびデータ収集装置の開発、は統計生産の継続的改善に必要な領域である。顧客との継続的対話の維持と育成は、統計の適合性、理解可能性、およびアクセス可能性にとって不可欠である。顧客の評価と分析、およびフィードバックの積極的収集は、品質を評価し高めるための主な方法と道具である。品質記述あるいは別個の方法論的記述（1.2.3節）が、統計の品質について利用者に伝えるために使われることが多い。統計の品質はまた、評価の焦点を、比較される統計の整合性と比較可能性に向けることを可能にする比較によって評価できる。

組織の品質を評価する他の方法は、部門の自己評価（Holma,2003）と職員調査であり、その結果は、上述したすべての生産者と一緒に利用されるべきものである。自己評価は、他の形の評価に支えられることが必要である。すなわち、同時に使用されるべき他の評価方法には、異なる利用者集団や利害関係者からのフィードバック、同業者評価、ベンチマーキングおよび独立した外部の関係者あるいは顧客が行う評価（Holma,2003）がある。

ガイドライン

- 組織全体にわたる統計生産の品質の評価においては、定められた手続きと道具を利用すべきである。品質評価の効率的で統合的な活用はまた、調整と操作への継続的インプットを必要とする。
- 生産のサイクル内の統計生産の品質評価の適切な時期は、十分な便益がそれらから獲得され、そこに投下される時間は、統計生産に負担を与えないことを保証することが不可欠である。
- 獲得された結果は、統計生産の管理と発展に利用されるべきである。主要な結果と決定された手段についての情報は、組織内で伝達されるべきである。
- 公開性は、すべての品質評価における不可欠の前提である。秘匿性と公開性原則の遵守の程度は、自己評価において前もって指定され、明確に伝達されるべきである。
- 品質評価における公開性原則の遵守は、組織の文化の継続的な発展を必要とする。
- 活動の品質の評価は、データの提供者、顧客および他の利害関係者からの情報の体系的な収集とともに、内部的評価を関与させる。それぞれにおいて、このフィードバックは、組織内の評価結果との関連で分析されるべきである。

文献

Eiderbrandt, G. 2004. Statistical Auditing at Statistics Sweden. Q2004 conference proceedings. Mainz, Germany.

European Foundation for Quality Management (EFQM). 2003. The EFQM Excellence Model. Brussels.

Holma, T. 2003. ITE2 – opas uudistuneen itsearviointi- ja laadunhallintamenetelmän käyttöön.

(ITE2 – Guide to the use of the renewed self-assessment and quality management method; in Finnish only).

Helsinki: Association of Finnish Local and Regional Authorities.

Körner, T. & Nimmergut, A. 2003. DESAP. Development of a Self Assessment Programme. Final report. Eurostat & Federal Statistical Office Germany – Institute for Research and Development in Federal Statistics.

Laiho, J. & Nimmergut, A. 2004. Using Self Assessments for Data Quality Management – DESAP Experiences.

Q2004 conference proceedings. Mainz, Germany.

追加的情報

Eurostat. 2006. The Eurostat Quality site. <http://ec.europa.eu/eurostat/quality>

1.2.3 フィンランドの政府統計と品質基準

視角と目的

フィンランド政府統計 (OSF: Official Statistics of Finland) は、フィンランド社会の全面的な描写をひきだすことを狙いとし、社会がその活動で必要とする統計情報を生産して社会の奉仕することである。OSF統計は、利用者のニーズと社会における一般的情報ニーズの両方を反映する。OFSは、その内容にしたがって26のトピックに分類される。

フィンランド政府統計の諮問委員会 (Advisory Board) は、政府統計の体系を調整し、OSFに

受け入れられている統計と出版物のリストを保全する。OSFの生産者には、フィンランド統計局と19の他の政府機関がある。

原則

フィンランド政府統計（OSF）は、国連の政府統計の基本原則とヨーロッパ共同体の委員会の実践規約に準じている（1.1.3節）フィンランド政府統計諮問委員会は、フィンランド政府統計に受け入れられた統計に対する基準を特定してきた。

フィンランド政府統計で発行された統計系列は、次の基準に合致していなければならない（Advisory Board of OSF,2006）。

OSFの基本的基準

OSFの基本的基準1:

この統計は、フィンランド統計局が管理するOSFのリストに含まれる機関等（agency or institution）が生産する。

OSFの基本的基準2:

発行機関は、統計の内容とデータの正しさ（correctness）に責任を持ち、したがって、当該統計の生産を継続的に保証する。OSF統計はいかなる個人の名前で公表してはいけない。

OSFの基本的基準3:

OSF統計はフィンランド政府統計の諮問委員会によって承認された更新された品質記述をふくむ。

OSFの品質基準

1. **適合性**： OSF統計は、それが表わすトピックについて社会のニーズに対して不可欠で全国的に包括的なデータをふくむ。
2. **正確性と信頼性（Accuracy and reliability）**： OSF統計は、それが関わる現象の状況と変化を正確かつ信頼性をもって記述する。統計の編集において適用される方法は、明確に報告されるべきであり、統計の信頼性に関係がありうるすべての事実は、分析され報告されるべきである。OSF統計でおそらく探り当てられた誤差を修正する手段は、フィンランド政府統計の諮問委員会が承認した勧告を遵守すべきである。
3. **適時性と迅速性（Timeliness and promptness）**：OSF統計のデータは、可能なだけ現在のなものでなければならず、その発表時点は、前もって周知され、フィンランドの政府統計の諮問委員会が発行した支持によって定められた手続きにそって、その発表者が保持している発表日程に発表されなければならない。
4. **整合性、一貫性および比較可能性（Coherence, consistency and comparability）**： OSF統計は他のOSF統計と論理的に一貫しており、時間と地域単位の間で可能な限り比較可能でなければならない。OSF統計は、基本的に一般的で、定められた国際的概念や統計分類を使用しな

らばならない。統計の一貫性や比較可能性に貢献する諸要因は、文書化され、統計に明細に示されるべきである。異なる定義を持つデータ集合は、異なる概念によって明確化されるべきである。

5. アクセス可能性と明瞭性：OSF統計は、より明瞭で、透明で、理解可能な形で示され、適切で便宜的な仕方では発表されるべきであり、その利用者への指示とともに、それらを支持するメタデータが、普通にアクセス可能でなければならない。

品質評価

OSF統計の生産者は、生産する統計の品質を、上に明示した品質基準に照らして定期的に評価しなければならない。これに加えて、彼らは、その顧客のニーズと満足度を、その統計の生産過程の便宜性を評価し、監視しなければならない。

統計の品質報告書と別個の方法論的報告書

OSFの統計の各集合は、その品質、信頼性と異なる目的に向けての適用可能性の簡潔な評価を提供する品質報告を伴うべきである。品質報告の主要な目的は、当該統計が、OSFの品質基準にどれだけ合致しているかを示すことである。

さらに、OSF統計の各集合について、別個の詳細な方法論的報告書とそういった報告書の必要な頻度に対する内的小および外的ニーズを考える必要がある。編集の方法において大きな修正が行われた統計に対しては方法論的報告を付けることが取り分け重要である。

表1.1は、品質報告と方法論的報告の構成、およびその違いを概説している。方法論的報告書の内容と範囲は、問題とされている統計集合の明確な特徴とその最終利用者のニーズに大きく依存する。

表1.1 品質報告書と方法論報告書の構成の提案

	品質報告書		方法論的報告書	
1. 適合性	1.1	情報内容の要約とその統計の利用目的	1.1	情報内容の正確なまとめと統計の利用目的。統計が叙述する現象の説明とその歴史。
	1.2	統計を理解するために不可欠な主要概念と使用された分類への案内と、調査の目的、データ収集者と提供者	1.2	統計を理解するために不可欠な主要概念と使用された分類への案内と、調査の目的、データ収集者と提供者
	1.3	その統計の編集が基礎にしている何らかの法規、布告および勧告	1.3	その統計の編集が基礎にしている何らかの法規、布告および勧告
	1.4	利用者をヒアリングと、当該統計の適合性と有用性の監視の方法の叙述	1.4	利用者をヒアリングと、当該統計の適合性と有用性の監視の方法の詳細な叙述
			1.5	統計の適合性が集団の要請とどのように対応するか、および現象におけるあらゆる変化が、それについての

				統計の編集においてどれだけ許されるかの評価
2. 統計調査の方法論的記述	2.1	適用された方法の明確な報告，すなわち，母集団，基本的データ，調査のデザイン（センサス調査あるいは標本調査），（標本設計），データ収集法，推定方法および標本調査での加重係数の使用。	2.1	最終結果を生産する際に使用される（標本抽出），データ収集，エディティング，補定，ウエイティングおよび推定方法についての品質報告以上に，より徹底的な叙述。
	2.2	標本誤差と非標本誤差の影響に関する報告。	2.2	標本誤差と非標本誤差の測定とその結果への評価報告。
			2.3	出所データ，暫定的結果，および生産された統計の評価と確認。
			2.4	使用された方法の正当化とそれらの改訂（時系列へのその影響の評価をふくめて）。
			2.5	使用されたデータ出所の，追加的情報をふくめた叙述。
			2.6	統計に向けての全生産過程の記述。
3. データ正確性の信頼性	3.1	統計が測定すると想定されている現象を測定していることの提示。	もし必要なら，品質報告書の項目3.1，3.2項に関して，より詳細な記述。	
	3.2	主要な不確性要因，すなわち，誤差のありうる発生源（例えば，無回答あるいは過小カバレッジ）とともに，統計の信頼性に意味を持つかもしれない総ての事実の記述。		
	3.3	必要ならば，推定のために使われた分類別の統計的パラメータの表示（例えば，標本調査の標本設計が許す標準誤差，平均分散誤差（MSE:mean square errors），標本設計の有効性を推定するパラメータ（deff））。	3.3	推定のために使われた分類別の統計的パラメータの表示（例えば，標本調査の標本設計が許す標準誤差，平均分散誤差（MSE:mean square errors），標本設計の有効性を推定するパラメータの提示と説明（deff））。
			3.4	主要な不確性要因，すなわち，誤差のありうる発生源とその公表推定値への影響の評価：すなわち， <ul style="list-style-type: none"> － 標本誤差 － 非標本誤差 － 枠組み － 測定 － 処理 － 無回答
	3.4	標本調査における抽出標本の無回答あるいは回答率についての報告。	3.5	追求されている母集団と対象母集団との間の枠組みの同等性の評価と利用された枠組みの評価。
			3.6	無回答，エディティングおよび補定からの最終統計への影響についての量的指標の提示。
4. 公表	4.1	公表統計が叙述する時点あるいは時	4.1	特定時点あるいは期間に関するデータが公表された

データの適時性と迅速性		期の確認(identification)。		ときの指示。
	4.2a	データが暫定的か最終的かの確認。	4.2	暫定データと最終データのありうる差の分析。
	4.2b	データが将来、例えば、季節調整で変わるかも知れないなら、これはデータの利用者に明示されるべきである。	4.3	生産スケジュールがOFS統計の勧告に従っているかどうかの検討。もし統計がOFS統計に向けて勧告されている目標スケジュール内に発行されていないか、その発行が遅れているなら、生産と公表における遅れの理由の説明。
	4.2c	統計が暫定値としてはじめて公表されたなら、最終データが利用可能になるときの指示。	4.4	必要なら、時系列データが時とともに変わったかどうかの検討（例えば、季節調整が理由になって）。
5. 統計の整合性、一貫性、比較可能性	5.1	時間にわたるデータと他のデータとの比較可能性の叙述。	5.1	時間にわたるデータと他のデータとの比較可能性に関する品質報告書よりもさらに詳細な叙述。
	5.2	統計の同じトピックの他の統計との一貫性と比較可能性の評価。	5.2	統計の同じトピックの他の統計との一貫性と比較可能性に関する品質報告書よりもさらに詳細な評価。
			5.3	統計の一貫性と比較可能性に影響を与えた変化と、例えば、使用された生産過程、概念および分類へのそれらの意味の検討。
			5.4	比較可能な時系列が生産された出所統計からの。利用可能な時系列の長さについての情報。
6. データのアクセス可能性と明瞭性	6.1	出版物の経路（チャンネル）の紹介と情報サービスのための連絡先の詳細の提供。	6.1	出版物の経路（例えば、電子的あるいは印刷出版物、データベースおよび/あるいはオンラインサービス）の紹介。その他の情報サービスの経路の追加的記述。
	6.2	最終利用者向けの統計生産で使用された方法と、ヨーロッパ統計システムおよびOSFの品質基準に対応する生産された統計データの品質、とに関する情報。	6.2	最終利用者向けの統計生産で使用された方法と、ヨーロッパ統計システムおよびOSFの品質基準に対応する生産された統計データの品質、とに関する情報。
	6.3	統計調査のメタデータの入手可能性とデータ出所の叙述。	6.3	統計調査のメタデータの叙述とより詳細なメタデータを伴うデータベースの参照。
7. 文書化			7.1	利用できる保管されたデータファイルの叙述。

ガイドライン

フィンランドの政府統計（OSF）の出版物は、OFSの諮問委員会が承認したフィンランド政府統計の出版ガイドラインを遵守し、OFS基準に見合ったインターネットで公的に入手可能な品質記

述への参照をふくむべきである。これに加えて、それらはヨーロッパ共同体委員会が発行したヨーロッパ統計実践規約の1.1.3節に準じるべきである。

別個の、詳細な方法論的報告書への内的および外的ニーズ、その内容と公表頻度が別個に評価されなければならない。

統計集合の品質は、基本的に、国際的に勧告された品質指標あるいはEurostatが明示した品質指標に照らして評価されなければならない。

統計的品質は、以下のOSFの品質基準に対応する品質報告に記述された。すなわち、

1. 適合性
2. 正確性と信頼性
3. 適時性と迅速性
4. 整合性、一貫性および比較可能性
5. アクセス可能性と明瞭性

文献

ヨーロッパ共同体の法規と国際的ガイドライン

Commission of the European Communities. 2005. Recommendation of the Commission on the independence, integrity and accountability of the national and Community statistical authorities.(COM(2005) 217.

http://eur-lex.europa.eu/LexUriServ/site/en/com/2005/com2005_0217en01.pdf

United Nations Statistics Division. 1994. Fundamental Principles of Official Statistics. Supplement No. 9.

Report of the Statistical Commission on its Special Session. New York 11.-14.4.1994.

E/CN.3./1994/18,

E/1994/29. <http://unstats.un.org/unsd/statcom/doc94/e1994.htm>

Eurostat. 2005. Standard Quality Indicators. Doc/ESTAT/02/Quality/9/EN.

国のガイドライン

Advisory Board of Official Statistics of Finland. 2006. Basic and quality criteria for Official Statistics of Finland. (TK-41-1227-06).

Advisory Board of Official Statistics of Finland. 2004. Suomen virallisen tilaston julkaisuohje (Publishing guidelines for Official Statistics of Finland; in Finnish only)

フィンランド統計局のガイドライン

Statistics Finland. 2004. Tilastokeskuksen julkistamisohje. (Statistics Finland's publishing principles; in Finnish only).

Statistics Finland. Internal memorandum. Helsinki.

追加的情報

Eurostat. 2006. The Eurostat Quality site <http://ec.europa.eu/eurostat/quality>

2 統計調査と統計の生産

2.1 統計調査の定義

視角と目的

本章は、統計生産の一般的性格と統計の品質に基本的に影響を与える諸要因を検討する。その目的は、読者に統計生産と統計の品質の間の結びつきの全体的描写を、広く使われ確定した定義によって与えることである。したがって、本章の節は、統計の品質に貢献する諸要素を反映する。

統計調査という術語は、ここでは社会的あるいは経済的意思決定、科学研究あるいは国際比較のニーズに向けて、社会についての記述的あるいは分析的情報を作成する統計調査 (stataistical studies) に限っている。それらの目的は、検討中の主題事項に関して母集団全体の状況を記述することにある。母集団は、個人、世帯、企業と事業所、あるいは他の経済的および/あるいは経済的単位からなる。

統計調査は伝統的には統計的トピックスによって区分されている(1.1.4節参照)。目的が、全統計生産および/あるいは個別的統計過程を計画し、維持し、強め、質的に改善することであるときには、異なる統計調査過程の類似性と特徴の概略を描くことが有益である。

これらの統計調査は、調査(統計的調査)という術語、すべての実証的、社会的統計的研究と生産からなる広い術語によって、文献において言及されている。より正確には、調査は、目標母集団およびそれについて収集したデータに適用される(Marriott, 1990)。定義によって、統計調査という術語は、次のようなタイプのデータをふくむ。すなわち、

- 総合的データ (センサス調査), ここでデータ収集は母集団の全単位をふくむ, ,
- 標本データ (標本調査), ここでデータは目標母集団から(通常は任意で)選択される,
- 行政レジスターから作成されたデータ (行政記録) ,
- 導出された統計データ (導出された統計活動), ここでデータは推定されたか, モデル化されたか, さもなければ, 既存のデータのプールから引き出されたか, である。(Statistics Canada, 1998).

統計的調査とはこのように、レジスター・データに基づく調査をふくめて、すべての調査を取り上げる高度に一般的な概念である。母集団の各単位は統計調査(サーベイ)のために全体データに基づいて標本抽出される。標本ベースの統計調査においては、目標母集団の単位のあるものが、目標母集団全体のレベルまで一般化できる代表的結果を目指して、選出される。統計調査は、標本あるいは全母集団およびそれらのアプローチの組み合わせからの直接的および間接的データ収集をふくむ。

統計調査は、繰り返される調査と一回限りの調査に分けることもできる。繰り返しの統計調査はある間隔で行われ、これを基礎にして時系列を作成することも可能になる。統計調査はまた繰り返しの調査系列、一回の、および二次的系列に分割することができる。第一のタイプでは調査

が同じ質問群によって、幾度も繰り返されている。一回の統計調査の場合には、調査の全体過程は、最初から包括的に構成されなければならない。二次的調査は他の調査あるいは行政レジスターからのデータを統合する。調査票のテストにおいては、フィンランド統計局の調査実験室を使用し、そして/あるいは試験調査をすることが勧告されている(2.8節を参照)。

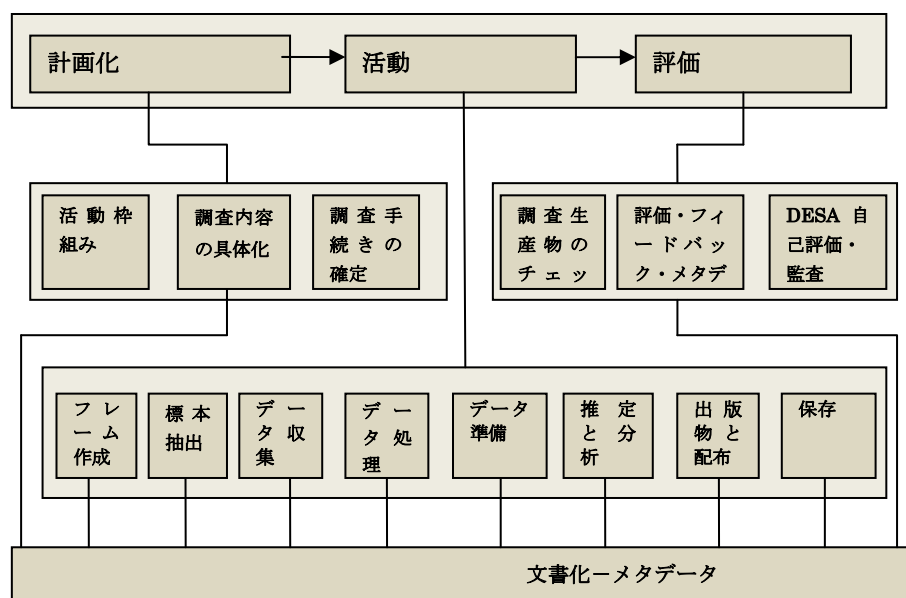
原則

計画と目標の設定のために、統計調査は3つの主な過程、すなわち、計画化過程、作業および評価からなるすべての過程と見なされるべきである(図 2.1参照)。

計画化過程 (Planning process)

統計調査の計画化は社会的あるいは経済的意思決定あるいは科学研究のニーズに基づくことが多い。顧客、研究者、あるいは利害集団のデータ・ニーズは、計画化段階で確認され、明示されるべきである。フィンランド統計法(280/2004)が定めた作業の様式、概念、分類および基準は、統計調査の作業枠組みを決定する(1.1節参照)。図2.2はフィンランド統計局の計画過程を描いている。統計調査の目的は、現実的であるべきであり、目標母集団および調査の内容は**研究**計画の中に明示されるべきである。計画化段階の重要な領域は、内容の定義、データ収集方法についての戦略的決定、およびデータ収集の計画である。予算、職員やその他の装備といった資源もまた、計画化段階の早くに、定められるべきである。同時に、過程、データ保護の方法、および日程も統計調査について定められるべきである。すべての努力は、統計に対する品質基準の集団を満たすことができることを保証するためにすべての努力が払われるべきである(1.2節参照)。

図 2.1 統計調査の全体過程 (Sundgren が準備した United Nations 1999 から引用¹⁾)



1) 原出所 (Sundgren が準備した United Nations 1999) に4つの追加をした。統計調査過程に「文書化」と「メタデータの蓄積」が付け加えられ、「作業枠組み」が計画化段階に、「監査」が評価段階に加えられた。

データ収集方法の選択は、統計調査を計画する際に戦略的に重要な決定である。それは、データ収集の性格、質問の設定、調査票の計画の他の側面、統計調査の存続期間、およびその費用に対して大きな意義を持つ(2.7と2.6節参照)。データ処理方法もまた計画段階で、最終的データ集団の生産に必要なエディティングや補定の方法とともに、詳細化される。(2.9節参照)。計画段階の他の重要な要素は、調査員の訓練と現場活動の組織である(Statistics Finland, 1999)。

これに加えて、標本に基礎をおく調査においては、計画段階はまた、標本設計と推定方法を決定する標本抽出の計画化を必要とする。

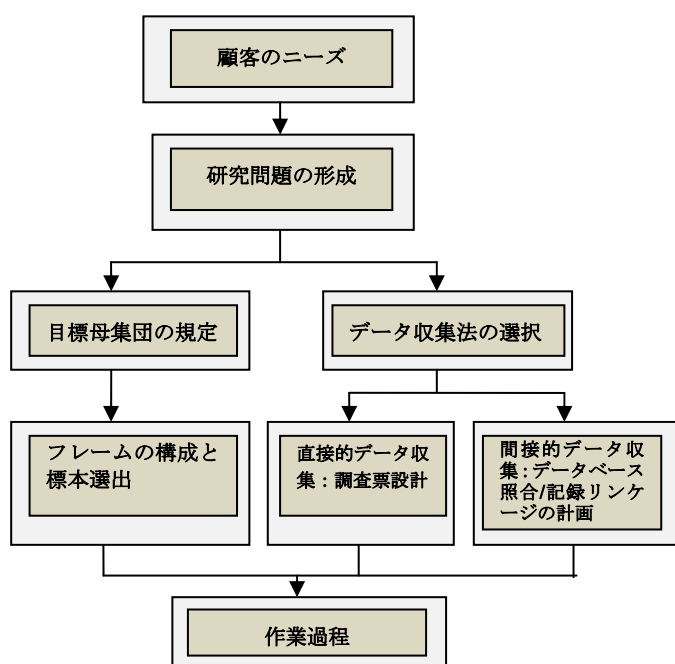
計画化の主要な目標は、可能な限り正確で正しい調査結果の獲得を助けることである。起こりうる問題は、計画過程で予想されるべきである。例えば、面接調査においては、無回答に対してと、それによって必要となる是正方法に対して準備するべきである。統計調査の計画化の文書化は、技術的報告、すなわち、生産物記述、品質報告、および利用者マニュアルの用意を速やかなものにする。

作業過程 (Operational process)

統計調査の作業段階は次の過程からなる(図2.3参照)。これらの過程は最終的結果の品質に本質的関係を持つ。

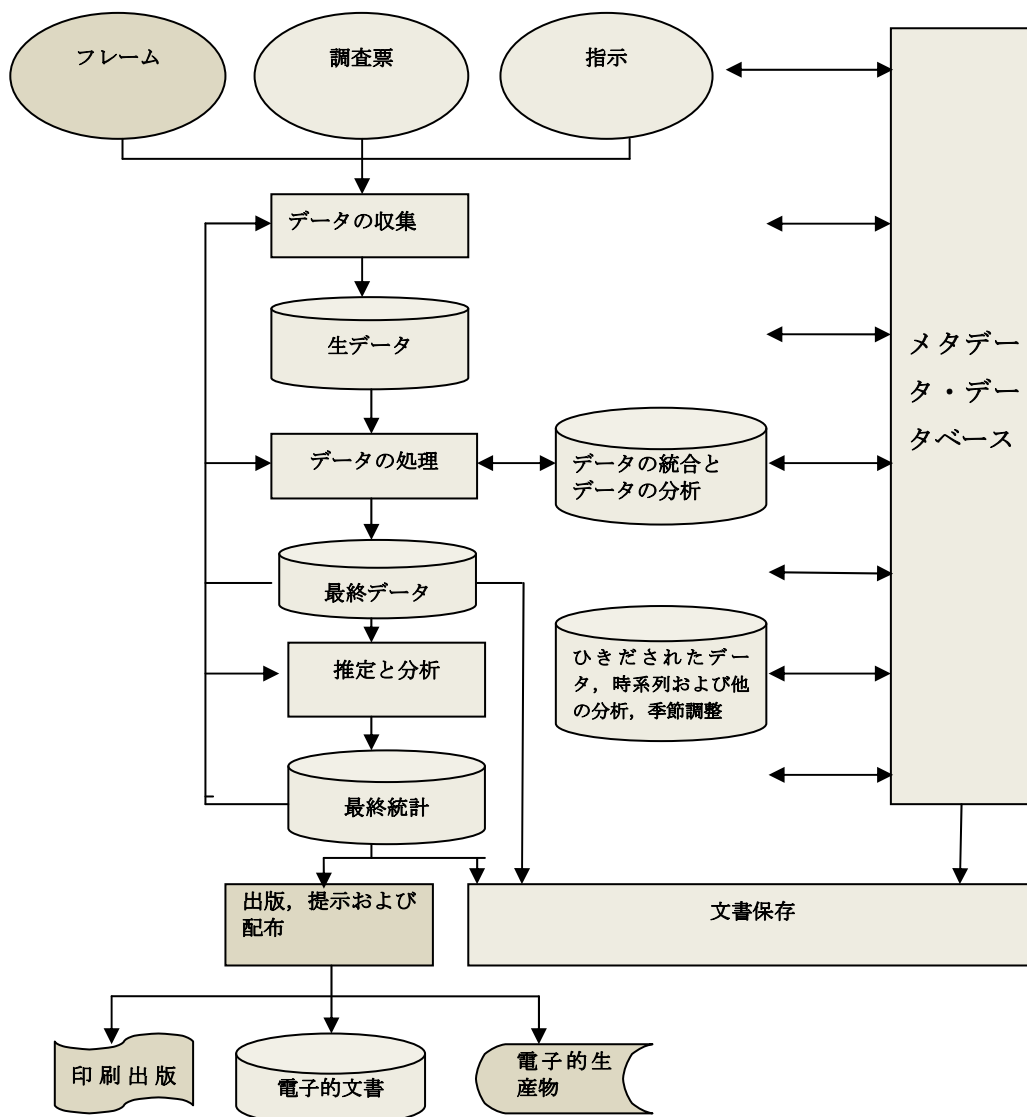
- 標本フレームの構築と必要な補助的情報の併合 (2.4節参照) ,
- 標本の抽出 (2.6節参照) ,
- 調査票の作成/計画と作業指示の準備 (2.8節参照) ,
- データ収集とその調整 (2.7節参照) ,

図 2.2 統計調査の計画過程



- データの処理 (データの収集, コーディング, (論理的)検査とデータのエディティング), (2.10節参照)
- 最終データの生産 (2.10, 2.11,2.12節参照)
- 推定と分析(2.12節参照)
- 最終結果の発行と統計情報の配布 (4.2節参照) ,
- 保存 (3.2節参照) .

図 2.3 統計調査の主要作業過程 (Sundgren が用意した United Nation,1999²から採用)



2) メタデータの結合に関する節は、過程を通じて、メタデータが収集され、結合するように、この図に変えられた (Sundgrenが用意したUnited Nations,1999)。これに加えて、最終データの「文書保存」、最終統計とメタデータが原図に付け加えられた。

データ管理とシステムワークは、中心において統計調査の各段階とつながっている。図2.3は、統計調査の文書化がどうはじまり、メタデータが計画化段階ですでに生じており、その後、作業過程を通じて蓄積され続けることを描いている。ここでメタデータとは、定義、分類、基準、サンプリングフレーム、補助的情報、調査票、作業指示、回答者と面接者のコメント、エディティング、および変数の転換 (variable transformations.) をさす。

評価過程 (Evaluation process)

統計調査の評価段階は、計画段階で定義された最終生産物がどのように生産され、発表されたか、そして品質基準がどのように達成されたかを調べる。さらに、評価過程は、メタデータの文書化の検査と評価、データの貯蔵、分類と概念の貯蔵、そして顧客満足度調査の分析をふくむ (United Nations, prep. by Sundgren, 1999)。

統計調査全体の実施と品質基準を満足させること~~逆行~~も、統計に責任を負う者のDESAP自己評価、組織内での統計生産の内部監査、統計の専門職員あるいは独立した専門家が一緒になって統計調査の全過程を評価して、改善すべき領域を確認することを狙う別個の監視、とによって評価することができる。

ガイドライン

統計調査における高い品質は以下によって達成できる。

- 統計調査の各段階に正しく資源を配分すること、
- 計画段階に投資をすること、
- 顧客ニーズの注意深い分析、
- 方法的システム活動によって特徴づけられる正しく資源配分された作業段階、
- 個別作業と全体作業の交互作用へ投資すること、
- 個別の活動段階を文書化し、活動段階が他の統計専門家によって繰り返すことが出来るようにすること (3.1節参照) 、
- それが文書貯蔵されたときにはじめて完成するものとして統計調査を承認すること (3.2節) 、
- 統計調査のアウトプットを評価すること、
- 他の統計調査と協力し、優良な実践と知識を共有すること。 .

文献

ヨーロッパ共同体の法規、国際的ガイドライン

United Nations Statistical Commission & Economic Commission for Europe. 1999. *Information systems architecture for national and international statistical offices. Guidelines and recommendations.*

Conference of European Statisticians Statistical Standards and Studies. No. 51. Geneva: United Nations. http://www.unece.org/stats/documents/information_systems_architecture/1.e.pdf

フィンランドの国内法, 内部的ガイドライン

Statistics Act (280/2004).http://stat.fi/meta/lait/lait_statisticsact04.pdf

Guidelines of Statistics Finland

Statistics Finland.1999. *Tilastohaastattelijan opas*. (Guide to statistical interviewers;in Finnish only). Internal memorandum. Helsinki.

他のガイドライン

Statistics Canada. 2003. *Statistics Canada Quality Guidelines*. Ottawa.

<http://www.statcan.ca/english/freepub/12-539-XIE/free.htm>

Sundgren, B. 2001. *Documentation and Quality in Official Statistics*. Conference CD. Stockholm

14.-15.5.2001. Statistics Sweden & Eurostat.

追加的情報

Dodge, Y., Cox, D., Commenges, D., Davison, A., Solomon, P., Wilson, S. 2003.

The Oxford Dictionary of Statistical Terms. 6 th ed. Oxford, New York: Oxford University Press.

2.2 データ保護

視角と目的

データ保護とは、その総てがデータ提供者のプライバシー、事業あるいは商業上の秘密を保護することを狙う法規、規制、方法およびその応用を意味する。データ収集段階とその計画時に、統計生産に必要でないデータは収集されないことを保証することによって、すでにデータ保護は考慮されている。データは、データ保護原則にしたがって処理されているが、これは、秘匿データは、データ処理のあらゆる段階で、外部者には開示されないことを意味する。

データ保護は、政府統計のひとつの中心的原理である。データ保護原則の狙いは、データ提供者との秘匿の関係を維持することである。これは、統計編集のための網羅的で信頼できるデータの入手可能性を保証する。データ保護原則は、どのデータを集めることができるか、データがどう処理されるか、どういった形でそれらは発行されるか、を指定する。かくして、データの保護は、統計の編集で使われる手続きと公表されるデータの内容に影響を与える。

政府統計に関するデータ保護の制定法は、フィンランド統計法（280/2004）やヨーロッパ共同体統計についてのEC規制（EC）322/97に基本的に含まれているが、フィンランド個人情報法（523/1999）と政府活動の公開性に関する政令（621/1999と1030/2004）もまた、データ保護に関する条項をふくんでいる。法規や規制に加えて、データ保護に関するガイドラインも、政府情報安全管理委員会(VAHTI)によって発行されている。政府統計に関する主なVAHTIのガイドラインは、行政におけるデータ資料の処理（財務省、2000）およびセンシティブな国際的データ資料（財

務省,2002) に関するデータ保護のガイドラインがある。

フィンランド統計局は当然、統計法と個人データ法の適用のためのガイドライン (Statistics Finland,2005aと2000a) や、企業と個人の表形式のデータの保護に関するガイドラインといったデータ保護に関する幾つもの内部のガイドラインを持っている。データ保護に関する一般的な運用原則は、統計局のデータ保護原則(Statistics Finland, 2005b) と、EU,国連および国際統計協会 (ISI) が承認した統計的倫理の諸原則に基づく専門的倫理についてのガイドライン(Statistics Finland, 2006) に与えられている。このうち、国際的ガイドライン、特にISIの専門的倫理に関する宣言(ISI, 1985), 国連委員会が承認した 「政府統計の基本原則」 (UN,1994), およびヨーロッパ委員会の実践規約(COM 217, 2005)は、フィンランド統計局の専門的倫理についてのガイドラインを作るときの主な出発点であった。

原則

データ保護規定の観察は、データの収集から統計情報の公表と配布にいたる統計調査の全過程についての注意深い事前の計画を必要とする。関連するトピックを叙述するために必要なデータだけがデータ供給者から収集されてよい。収集される必要のあるデータ量を削減するためには、可能な場合にはいつも標本抽出や他の方法が使われなければならない。データの処理は、誰のプライバシーも、企業あるいは商業上の秘密も危険にさらされていないように実施されなければならない。統計を編集する政府職員は、データは計報な処理や露出がないように十分に保護されることを保証しなければならない。フィンランド統計局においては、統計部門からの自分自身以外に関するマイクロデータの利用もまた許可を必要とする。一般に、統計的目的で獲得されたすべてのデータは秘匿される。中央と地方の政府機関の活動および政府機関のサービスの生産を描くデータは、公的なものである。ビジネス・レジスターにふくまれているデータもまた公的である。秘匿データは、科学目的あるいは統計的研究に向けて、個人は直接的であれ間接的であれ、個人が識別されない形でだけ開示されることができる。1つの例外として、年齢、ジェンダー、教育と職業に関するデータは、科学的研究あるいは統計的研究に向けての識別データを伴って開示されることができる。データはまた、自らの分野の活動についての統計を編集する目的で、他の幾つかの統計機関に対して開示できることが統計法に規定されている。行政的意思決定あるいは他の類似の活動での利用のための秘匿データの開示は明確に禁止されている。データを開示するために応用される方法と原則は、フィンランド統計局のデータ・ファイルへの利用者の権利の許可について、局長が支持したガイドラインに従っている (フィンランド統計局, 2005c)。

ヨーロッパ共同体への加盟は、共同体決議(EC.No.281/1997)ーここで、特定の統計に関するEUの規制がそれを要求しているのだがーによって、共同体統計の生産者であるEurostatに対して統計データが公表されるべきことを意味する。共同体統計に関する理事会規制 (EC, No. 322/1997) は、ヨーロッパ統計システムの一般的原則と、データ保護に関する条項をふくんでいる。これに加えて、加盟国からEurostatへの統計データの提示は、ヨーロッパ共同体統計局への統計的秘匿性に従うデータの移動に関する理事会規制によって規定されている (EC, No. 1588/1990)。個人デ

ータの保護に関する条項は、ヨーロッパ議会と理事会の指令（EC, No. 46/1995）と個人データに関する理事会勧告（EC, No. R(97) 18/30 1997年9月）の両方に公表されている。さらに、実践規約（COM 217, 2005）は、統計的秘匿性に関する絶対的勧告を与えている。科学的目的のための秘匿データへのアクセスに関して、共同体統計に関する理事会規制を（EC）No. 322/97 を実施する、委員会規制（EC, No. 81/2002）は、状況と手段を指定しており、これが満たされた時には、Eurostatが所有しているデータを研究者に示すことを許している。データは、供給者あるいは供給者に関する秘匿情報を開示しない形式でだけ公表される。秘匿データが開示される例外的なのは、関連する供給者の同意がある場合だけである。この可能性は、主として大会社のデータに適用される。統計が研究目的に開示あるいは公表されるときには、秘匿データを守る必要は常に確かめられるべきである。ごくわずかの努力でマイクロデータの開示の危険があるならば、データは、適切な統計的開示コントロール法によって保護されなければならない（Hänninen, 1997）。データ保護における考慮は、データの性格と敏感性、およびデータの使用に適せるな方法の選択に対して払われなければならない。というのは、開示コントロール法は、統計の内容に影響を与えるからである（Konnu, 2006）。

統計的開示コントロールの方法

秘匿データを保護するための開示コントロール法は様々ある。適切な方法を選択する際には、保護されているデータへのことなる方法の影響は大きく異なりうるので、データの将来の利用を意図した考慮が払われるべきである。すべての開示コントロール法の狙いは、記録あるいはその属性の開示を不可能にする形式におくことである。表データとマイクロデータについては異なる開示コントロール法があり、それらの方法は別々の論議されることが必要である。以下は、フィンランド統計局でテストされ/あるいは使用された開示コントロール法の例である。

I. 表データの開示コントロール

表データに関しては、変数の敏感性が、適用されるべき開示コントロール法の強さを決定する。表データの保護ニーズの評価には、表のセルにおける場合の数とともに、変数の秘匿性、母集団の大きさ、変数の数、統計が関わる地域の大きさ、およびデータが区分される詳細度を考慮しなければならない。表が公表される前に、そこにおけるデータをもう一度チェックするのが良い考えである。実際表が保護されていることに加えて、関連する他のすべての表もまたチェックされるべきである。統計の発表者は、秘匿データが、表を比較・検討することによって開示されないことを、保証しなければならない。

表データの保護を考える際には、頻度表と大きさの表が区別されるべきである。頻度表は、表の各セルに場合の数を与え、大きさの表は、特定通貨での所得に平均値あるいは合計といった数値的値を示す。事前に設定された域値—これよりも小さなセルの頻度は圧縮されなければならない—を適用することによって保護されることが最も多い頻度表よりも、大きさの表を保護することに、より多くの時間や他の資源が通常は必要である。ときとして、分類区分を結合し、そこで

の新しい、結合された分類のセルの頻度が、域値を越えるようにすることによって、圧縮を避けることができる。

大きさの表を保護する際には、表の値のみならず、記録の値についての情報も必要である。表の敏感性を規定する二つの基準がある。すなわち、一セルにおける場合の数と一セルにおける場合の支配である。表データの保護を考える際には、頻度表と大きさの表が区別されるべきである。頻度表は、表の各セルに場合の数を与え、大きさの表は、特定通貨での所得に平均値あるいは合計といった数値的値を示す。事前に設定された域値—これよりも小さなセルの頻度は圧縮されなければならない—を適用することによって保護されることが最も多い頻度表よりも、大きさの表を保護することに、より多くの時間や他の資源が通常は必要である。ときとして、分類区分を結合し、そこでの新しい、結合された分類のセルの頻度が、域値を越えるようにすることによって、圧縮を避けることができる。

大きさの表を保護する際には、表の値のみならず、記録の値についての情報も必要である。表の敏感性を規定する二つの基準がある。すなわち、一セルにおける場合の数と一セルにおける場合の支配である。もし、ひとつあるいは少数の記録だけがあるセルの値に寄与しているなら、その値は圧縮されなければならない。これに加えて、セルの値は表の限界的合計を使って計算されてはならないこともまた確かめられなければならない。もしそういった計算が可能ならば、限界的合計に貢献している最初のセル以外の他のセルは圧縮されるべきである、言い換えれば、補足的圧縮が使用されるべきである (Duncan et al. 2001)。支配ルール、あるいは(n,k)ルールを適用するときには、セルに貢献しているすべての記録とセルの値に貢献している数字についての情報が必要である。例えば、支配ルール (2,80) の場合には、セルの値に貢献している2つの最大の記録が、全セルの値の80%以上を数えるなら、セルは保護されなければならない。

大きさの表に対するより正確な開示コントロール法の目的は、公表された数字の殆どに貢献している記録がお互いの値を開示する可能性を防ぐことである。例えば、その分野で2番目に大きな会社は、それがそう望むなら、会社によるその分野の他の会社の売上高の推定値が正しいのなら、その最大の競争相手の売上高を極めて正確に評価することができる。そういった露出は、セルの性アックな数値を発表せず、それがある区間を示すことで防ぐことができる。「p%ルール」(Merola, 2005)は、保護を必要とするセルを探知するために使うことができる。そのデータが検討されたセルの値に貢献する単位の一つが、セルの値に貢献する他の単位の1変数の値を、p%の正確性を推論できるなら、p%ルールは、保護されるべき問題のセルを必要とする。上の例の保護を80%ルールと対応して検討しよう。この場合は、売上高を含むセルは、その分野の第二に最大の会社は、その最大の競争相手の売上高を80%の正確性で推定できる。Argusといったように、SuperCossソフトウェアに含まれているデータ保護に利用可能な多様なソフトウェア・プログラムや、小地域統計のための開示コントロール方法がある。

II. ミクロデータの開示コントロール

ミクロデータは非摂動的的方法 (non-perturbative methods) あるいは摂動的的方法 (perturbative

methods) よって保護することができる(Crises, 2004a and 2004b)。その名前が示すように、非摂動法は、データから秘匿情報を取り除くことに基づいている。マイクロデータを開示から保護する最も効果的な方法は、全体データの代わりに情報を識別しないで標本データを公表することである。全体データは、ある資格を持っているか、ある領域に住んでいるすべての人を指すことができる。フィンランド統計局において研究目的で公表されたすべてのデータは、全体データの標本である。

フィンランド統計局がマイクロデータのために使用している他の方法には、全体的記録化(global recoding)、すなわち、データをより広い分類で区分する、トップコーディング(top coding)、すなわち、最高の値を持つ区分を結合する、ボトム・コーディング(bottom coding)、最低の値を持つ区分を結合する、方法がある。例えば、80歳以上のすべての人々は、データの残りが年齢5歳のコーホートで区分されている場合でも、個人データの1区分にまとめることができる。

データの変数の値がユニークな像を持つ場合は、稀な値の部分的圧縮を使用する。例えば、データがある職業の1つの代表を含むだけなら、その人の職業に関するデータは、欠測値として表わすことができ、これによって通常は識別の危険を除去する。この方法の弱点は、乱入者(an intruder)、すなわち、開示を企てる人は、容易に開示される記録を捜すときに、欠測値の記号を活用出来ることである。したがって、圧縮された値は、他の欠測値に対して使われた記号とは異なる記号ではなく、理由はどうあれ、総ての欠測値について使われたものと同じ記号で置き換えるべきである。

データが非摂動法で保護されるときには、保護された値以外のいかなる見地においても分析結果を変える危険はない。しかしながら、あまりに広い分類に区分されたデータ、あるいは余にも多い欠測値をふくむデータは、利用するのが困難でありうる。摂動法は、利用者が特に厳密なデータを使う必要がある場合に使用されるべきである。これらの方法は、さらに体系的摂動法とランダム摂動法に区分できる。体系的方法は、それが使われるすべてのときに、データに同じ変化を生みだし、ランダム法は、各回に異なるマスクされたデータを作り出す。

体系的開示コントロール法のひとつは、マイクロ総計(Crises,2004c)、であり、ここではデータ記録は、類似の記録のグループに配列されて、記録の変数の値は、グループの平均値で置き換される。グルーピングするには種々の選択肢があり、グループの大きさは、変えることも一定とすることも許される。ランダム摂動法の典型的例は、変数の値にノイズを加えるものである(Brand, 2002)。これは、最初の変数が受け取った値にランダム変数を加えることを意味する。多様な修正の手段によってノイズを加えることは、保護効果を強める。摂動法のもうひとつの簡単な例は、特定の変数の値のあるものが二つの単位の間で交換し合うときのデータスワッピングである(Fienberg & McIntyre, 2004)。データスワッピングは、ランダムにか、あるいは変数の値が交換の前に上向順に配列されるときの、配列スワッピングをを使って適用できる。さらに進んだランダム摂動法には、保護が、事前に選択された確率マトリックスに従った変数値の変化に基づくPRAM法がある。

マイクロデータに関する限りは、すべての状況に適用できる網羅的なデータ保護ガイドラインを与えることは不可能である。というのは、求められる手段は常に、データの量、変数の数、およ

び分類の正確性の程度といった幾多の寄与要因に依存するからである。マイクロデータの保護はまた、Argusプログラムといったソフトウェアのプログラムで実施できる。

ガイドライン

一般原則

フィンランド統計法は、統計目的で獲得されたデータすべてが秘匿されるという基本原則に基づいて機能している。以下の基本的原則は、データの収集、処理、公表において観察されるべきである。

- 統計生産に必要なデータだけが収集されるものとする。
- レジスターのために既に収集されたデータが可能な限り活用されるものとする。
- 識別情報は、絶対的に必要なときにだけ収集し、利用するものとする。
- 間接的な識別を促進する変数もまた、データが発表される前に、除去するか、集計されるものとする。
- データ開示コントロール法は常に文書化されるものとする。

一般的なデータ処理の原則に加えて、データを公表するか提供するときの以下のような詳細なガイドラインが考慮されるべきである。機関の外部にどのデータを提供するかを評価するときには、表形式をマイクロデータと区分する必要がある。

表データに関するガイドライン

- 個人についての表データの保護に関するフィンランド統計局のガイドライン(Statistics Finland, 2002) あるいは企業についての表データの保護に関するフィンランド統計局のガイドライン(Statistics Finland, 2000b) に、表データの提供のときに従うものとする。
- 個人あるいは企業についてのデータの統計表は、それらのセルが不必要に低い頻度をふくまないように編集されるものとする。個人、家族、住居単位あるいは企業を識別する危険は、データが公表されるか提供される前に常に評価されなければならない。
- 表データは、保護の後でも可能な限り高い有用性を保持する適切な方法で保護されるものとする。
- 例えば、多次元の表が高度に詳細な分類レベルを使って構成されているときには、表データはマイクロデータに近づく。そういった例では、その提供において利用者の権利を認める必要のあるマイクロデータと解釈するかどうかについては、ケース・バイ・ケースの検討が必要であることとする。

マイクロデータに関するガイドライン

- マイクロデータの提供の際には、フィンランド統計局のデータファイルへの利用者の権利を認めるガイドラインに(Statistics Finland, 2005c) したがうものとする。
- マイクロデータは科学的目的あるいは統計的研究に対してだけ提供されるものとする。

- マイクロデータの提供は、常に利用者の権利を認める手続きに従うものとする。利用者の権利は、定められた期間について特に認定された研究目的に対して許可される。
- 全データ集合はフィンランド統計局の外部には提供されないものとする。外部者に対しては部分集合のみを提供できるが、研究者は、フィンランド統計局の建物内での監督のもとに全データ集合へアクセスできる。
- 提供されるマイクロデータは、どんな記録も識別できない形式のものとする。
- 開示コントロール法は、間接的識別を防ぐために使用されるものとする。企業に関するデータ集合については、一般的には、間接的識別の可能性を除去することはできないので、フィンランド統計局の建物内での監督のもとに作業する研究者にだけ提供できる。
- マイクロデータの提供は常に、データ保護原則に沿った確認に従うものとする。

文献

ヨーロッパ共同体の法規と国際的ガイドライン

国際的ガイドライン

Commission Decision (EC) No. 281/97 of 21 April 1997 on the role of Eurostat as regards the production of Community statistics.

Commission of the European Communities. 2005. On the independence, integrity, and accountability of the national and Community statistical authorities. COM(2005) 217.

Council Recommendation No. R (97) 18 and Explanatory Memorandum of the Committee of Ministers to Member States Concerning the Protection of Personal Data Collected and Processed for Statistical Purposes.

Council Regulation (EC) No. 322/97 of 17 February 1997 on Community Statistics.

Commission Regulation (EC) No 831/2002 of 17 May 2002 implementing Council Regulation (EC) No 322/97 on Community Statistics, concerning access to confidential data for scientific purposes.

Council Regulation (Euratom, EEC) No. 1588/90 of 11 June 1990 on the transmission of data subject to statistical confidentiality to the Statistical Office of the European Communities.

Directive of the European Parliament and of the Council (EC) 46/95 of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data.

International Statistical Institute (ISI). 1985. Declaration on Professional Ethics.

United Nations Statistics Division. 1994. Fundamental Principles of Official Statistics. Supplement No. 9. Report of the Statistical Commission on its Special Session. New York 11 – 14 April 1994. Reg. No. E/CN.3/1994/18, E/1994/29.

フィンランドの国内法

Act on the Openness of Government Activities (621/1999).

- Decree on the Openness of Government Activities and Good Data Management Practice (1030/1999).
- Personal Data Act (523/1999).
- Statistics Act (280/2004).
- フィンランド統計局のガイドラインと他の国家的ガイドライン
- Ministry of Finance. 2000. Valtionhallinnon tietoaineistojen käsittelyn tietoturvallisuusohje. (Data protection guidelines for processing data material in public administration; in Finnish only). Guidelines VAHTI 2/2000.
- Ministry of Finance. 2002. Arkaluonteiset kansainväliset tietoaineistot. (Sensitive international data material; in Finnish only). Guidelines VAHTI 4/2002.
- Statistics Finland. 2000a. Henkilötietolain soveltamisohje. (Guidelines on applying the Personal Data Act; in Finnish only). Reg. No. TK-00-578-00. Helsinki.
- Statistics Finland. 2000b. Ohje taulukkomuotoisen yritystiedon suojaamiselle. (Guidelines on the protection of tabulated enterprise data; in Finnish only). Reg. No. TK-00-888-00. Helsinki.
- Statistics Finland. 2002. Ohje taulukkomuotoisten henkilötietojen suojaamiselle. (Guidelines on the protection of tabulated personal data; in Finnish only). Reg. No. TK-00-314-02. Helsinki.
- Statistics Finland. 2005a. Tilastolain (280/2004) soveltamisohje Tilastokeskuksessa. (Guidelines on applying the Statistics Act (280/2004) at Statistics Finland; in Finnish only). Reg. No. TK-00-198-05. Helsinki.
- Statistics Finland. 2005b. Tietojen suojaaminen Tilastokeskuksessa. Tietosuojaohje. (Data protection at Statistics Finland. Data protection guidelines; in Finnish only). Reg. No. TK-00-334-05. Helsinki.
- Statistics Finland. 2005c. Ohje käyttölupien myöntämisestä Tilastokeskuksen perusaineistoon. (Guidelines on granting permissions to use Statistics Finland data files; in Finnish only). Reg. No. TK-00-128-05. Helsinki.
- Statistics Finland. 2006. Guidelines on Professional Ethics. Rev. ed. Handbooks 30b. Helsinki.
- 統計的開示コントロールの方法
- Brand, R. 2002. Microdata Protection through Noise Addition. Teoksessa: *Inference Control in Statistical Databases*, toim. Domingo-Ferrer, J. Berlin, Heidelberg: Springer. 97-116.
- Crises, G. 2004a. Non-Perturbative Methods for Microdata Privacy in Statistical Databases. Research Report CRIREP-04-007.
- Crises, G. 2004b. Perturbative Masking for Microdata Privacy Protection in Statistical Databases. Research Report CRIREP-04-008.
- Crises, G. 2004c. Microaggregation for Privacy Protection in Statistical Databases. Research Report CRIPREP-04-005.
- Duncan, G.T., Fienberg, S.E., Krishnan, R., Padman, R. & Roehrig, S.F. 2001. Disclosure Limitation Methods and Information Loss for Tabular Data. Teoksessa: *Confidentiality, Disclosure, and*

- Data Access: Theory and Practical Applications for Statistical Agencies*, toim. Doyle, P., Lane, J.I., Theeuwes, J.J.M. & Zayatz, L.V. Amsterdam: North-Holland.
- Fienberg, S.E. & McIntyre, J. 2004. Data Swapping: Variations on a Theme by Dalenius and Reiss. Teoksessa: *Privacy in Statistical Databases*, toim. Domingo-Ferrer, J. & Torra, V. Berlin, Heidelberg: Springer.
- Gouweleeuw, J.M., Kooiman, P., Willenborg, L.C.R.J. & de Wolf, P.-P. 1998. Post Randomisation for Statistical Disclosure Control: Theory and Implementation. *Journal of Official Statistics*. Vol.14. No. 4. 463–478.
- Hänninen, M. 1997. Tilastolliset tietosuojamenetelmät ja niiden käyttö. (Statistical disclosure control methods and their use; in Finnish only). Statistics Finland. Reviews 1997/3. Helsinki.
- Konnu, J. 2006. Mikroaineistojen tilastolliset tietosuojamenetelmät henkilötalustoissa. (Statistical Disclosure Control Methods; in Finnish only). Master's theses in Statistics. University of Jyväskylä: Department of Mathematics and Statistics.
- Merola, G.M. 2005. Safety Rules in Statistical Disclosure Control for Tabular Data. Teoksessa: *Monographs of official statistics, Work Session on Statistical Data Confidentiality (Geneva, 2005)*. Luxembourg: Office for Official Publications of the European Communities.

2.3 分類, 概念, 定義と調整

視角と目的

統計は通常、個別的観察の大きな集まりから作成される。結論をだすためには、経験的観察は、一貫した形で、たとえば、それらの類似性あるいは差異によって分類されるなど、配列される必要がある。分類は、このように統計の不可欠の一部であり、標準分類は政府統計の重要な道具である。

概念、定義、分類はともに、統計の生産の不可欠の作業の枠組みを形成する。使用された概念は通常、理論的あるいは統計的な座標となるような理論体系に基づいており、主題、描かれるべき統計単位および/あるいは調べている母集団を定義するために使われる。

統計に使われる概念は、一般的には標準的統計概念であるが、定義されるべき標準語の概念でもありうる。科学的概念の使用は、困難の原因になることが多い。ある概念は幾多の異なる現象にあてはまるかも知れず、ひとつの概念あるいは言葉は幾多の意味を持つか、あるいはすべての概念が経験的に等しいものを持たないのである。

最も簡単には、分類は、名前を付された集団やその識別項(コード)からなる。一般的には、分類の記述はこれよりも広く、(i)集団の定義(名前、コード、またしばしば集団の説明、言い換えれば、その文の形の定義)、および複雑で広い分類の場合には、(ii)基本的概念、分類の原則と

基準の提示，からなる。

分類は概略的に以下に区分できる。

- 人口，人，そして彼らの生活状況を説明する分類（人口動態的分類），
- 経済活動，および国民経済を説明する分類，
- 地域分類と環境統計向け分類，
- 疾病，関税専門語，犯罪専門語といった特別な分類

原則

分類原則

データの観察は，編集された統計が，現実—社会の構造，その現在の状況とそこでの変化—に対応する現実社会の構成された理解可能な描写を生み出すように，グループ化されるか分類されるべきである。統計に使われる分類は主な利用者の目的に適合的であるべきである。というのは，統計は社会研究と意思決定において大きな役割を果たすからである。

分類は体系的であるべきである。すなわち，それは観察値を同意された基準を使って一貫した形で分類するべきである。この分類は通常，大分類レベル（ないしは1桁レベル）—さらにこれはより低い分類レベルに区分される—の位階順になっている。

分類グループは，あいまいさが無く，総てを尽くし，相互に排他的でなければならない。あいまいさが無いことは，観測値が定められた分類原則と基準に基づいてあるグループに明確に区分できることを意味する。悉皆性は，観測データのすべての場合が，分類できることを意味する。悉皆性は，グループの定義が1つのケースを2つあるいはそれ以上のグループに分類されることを排除するときに達成される。

「他に区分できない」といった指定できないグループあるいは残余グループは，分類レベルの他のグループに比べて相対的に少ないケースをふくむべきである。残余グループの大きさが大きくなる場合には，分類システムは改訂されるべきである。

分類に関連する概念や変数は定義されなければならない。この定義は，本質的に一般的であるか，観察値の処理をあるデータの見地からの説明する操作的定義からなりたつことがありうる。

分類基準

標準統計分類は，ほとんど例外なく国際勧告あるいは同意に基づく。ヨーロッパ共同体の加盟国が使っている分類のあるものは，規制として確認されており，これは，ヨーロッパ共同体の加盟国が，政府統計において，そしてヨーロッパ共同体の統計局，Eurostatに情報を提供するときに，関連する分類を使うことを義務付けている。規制として確認されているそういった分類には，例えば，産業および生産物分類，改訂国民勘定体系(ESA95)の基本分類がある。対照的に，人口とその生活条件を描く分類は，人口センサスの勧告に基づいている。さらに，所得分布と消費を叙述する分類は，国民勘定の枠組みおよび所得統計に関する勧告に基づいている。

フィンランド統計局は，分類データベースを保持しており，これはフィンランド政府統計で使

われている分類基準をふくんでいる。このデータベースは、フィンランド統計局の内部的利用だけを意図している。フィンランド統計局のインターネットサービスは統計の分類に関して外部の利用者向けの情報をふくむ。http://tilastokeskus.fi/luokitukset/index_en.html。

中心的統計基準は、フィンランド行政的勧告として確認されている。この手続きは、社会のすべての部門そして特に行政における標準分類の使用の拡大をめざしている。一般的統計標準は、行政情報システムにおいて広く、そして民間企業と組織においてもますます使用されている。これは政府統計にとっては非常に重要である。というのは、行政、企業および会社は基本的統計データの重要な提供者だからである。

ガイドライン

- 標準分類はフィンランド政府統計において可能なときはいつでも使われなければならない。統計を発表するときには、分類と関連する中心的概念が常に提示されなければならない。
- 新しい定義と概念を伴って新しい分類が導入するときに、グループの見出しに、国内あるいは国際的標準分類ですでに使われている分類の見出しを与えていないことを確かめなければならない。
- 何らかの理由で標準的概念と分類に十分に従うことが不可能なときには、利用者には、統計出版物の説明において、あるいは他の何らかの適切な仕方で、通知されなければならない。EU規制で確認されている分類からの乖離は、ヨーロッパ共同体の統計局、Eurostatによって正式に承認されなければならない。
- 「その他」と言う用語は、残余の集団についての見出しとしては避けるべきである。というのは、それは自由な文章検索においては弱く機能するからである。
- 観察データの大きさによって、詳細分類レベルに適用することができないなら、分類は詳細ではないレベルで行われなければならない。これは同じレベルでグループを合計するか結合するかによって行うことができる。そういった指定されないグループについての普通の表記は0で終わる識別項である。これに対応して、9で終わる識別項が、残余グループに与えられることが多い。

比較可能性、調整 (Comparability, harmonisation)

- 国際勧告あるいは合意に基づく分類を使うことによって、異なる国で生産される統計情報が、社会とその構成の違いにもかかわらず、比較可能であることが保証される。
- 国際分類の各国版は通常、国の特徴点や国の状況に関わる事項の具体化をとり入れることを可能にするために作成される。
- 各国版は通常、最下位の分類レベルとして国別の詳細レベルを加えることによって準備される。この方法においては、構造と高い分類レベルは、なお国際分類と一致している。

個別的トピックスについてのガイドライン

国別コード:

国籍、出生や居住、移民、住居の国、外国貿易、あるいは企業の本社の所在国についての統計を生産するときには、国別コードに国際基準ISO3166を使うことが可能である。コードシステムは、3桁数字、2字および3字のコードをふくむ。国連統計部が発行している3桁の数字コードの使用が、統計システムにおける使用のために何よりも推薦される。2字コードは一般的使用向けを意味する。3字コードは、地域の識別を促進するため、あるいは明瞭性を増すために使うことができる。文字コードは、主として国の英語の名前に基づいている。

このシステムは独立国と非独立地域(国)の両方をふくむ。諸国は6つの大陸と19の主要地域にグループわけすることができる。この区分は主に人口統計に適している。諸国はまた他の方法で、例えば、経済政策の基礎あるいは国の連盟にしたがって、分類することができる。

地域分類とClandの特別な状況の考慮:

ヨーロッパ共同体の公式 NUTS 地域分類システム(NUTS,Nomenclature des Unités Territoriales Statistiques,これは、ヨーロッパ議会と理事会の規制(EC)No.1059/2003に規定されている)は統計において主として使用されている。フィンランドの NUTS 地域区分においては、大陸フィンランドとCland がNUTS1 地域レベルを、大きな地域がNUTS2 レベル、地域がNUTS3レベル、下位地域単位がNUTS4 レベル、そして市が NUTS5 レベルを形成する。

地域と、市の統計的区分が、統計に対して勧告された第一の地域分類である。市のデータ(NUTS5)についての表を生産するときには、市は地域 (area) 別に、通常は地区(region)別にアルファベット順に示される。Cland は常に最後の地区として示されている。地域的統計データを示す印刷出版物には、NUTS1区分 (大陸フィンランド,Cland) が下位の見出しとして、対応する小計を伴っても伴わなくても、使用することが勧告されている。

市の統計的区分:

市の都市化の程度は、市の統計的グループ化—これは、市を、都市部、準都市部、および地方都市に区分している—をもって叙述される。以前に使われていた「町/地方都市」へのタイプ別の市の分類は、1997年に停止された。それは、かなりの期間、町/地方都市という概念と同義としてうまく機能しなかったし、それぞれの区分が法規から取り去られたからである(法は、都市を町と呼ぶか呼ばないかは、市だけが独立して決定できることを認めている)。

世帯:

世帯概念は、フィンランドで面接調査と郵送調査においてだけ使用できる。住居単位概念は、レジスターを基礎にする人口センサス、年次統計および行政記録に基づくその他統計で使われている。

産業と生産物分類：

ヨーロッパ共同体の加盟国においては、産業別統計はNACE (NACE, Nomenclature Générale des Activités Economiques dans les Communautés Européennes) あるいはそこからひきだされた国内分類を使って作成されなければならない。2003統計年以降、この分類の更新版、NACE 2002 (NACE Rev1.1) が、それに基づいた国内産業分類、TOLSIIC 2002.とともに使われている。

生産物分類、生産物の活動別分類 (CPA: Classification of Products by Activity)は、例えば、Eurostatに生産、消費、外国貿易と分配に関する統計データを提出するときに、報告の分類として使われている。

産業生産についての統計は、標準産業分類の大分類C,D,E (鉱業と採石業、製造業、電気・ガス・水供給) の生産物項目をとりあげるPRODCOM リストを利用している。国内的必要によって、PRODCOM リストは、PRODCOMから落ちているか、8桁のPRODCOMの項目の下位項目である追加的項目で補われている。これらの追加的な国内項目は10桁である。

文書化

分類、概念と定義、および処理規則は注意深くすなわち、分類は分類データベースに、概念は概念データベースに、文書化されなければならない。処置規則と統計調査と結びついた他の作業指示は、それらが統一されたデータファイルで叙述されていない限りで、統計アルヒーブに貯蔵される。これによって、統計データと関連する分類、概念等々が、後に研究のために利用可能となる。

文献

追加的情報

フィンランド統計局の分類とメタデータサービス単位は、分類データベースに分類を維持し、ハンドブックとして分類を発行し、イントラネットとインターネットで分類に分類の情報を提供している。主要な分類はまた、例えば、顧客の訓練イベントや調整の会合に提出できる分類へのガイドリーフレットでも提供される。

分類データベースは、すべての分類基準とそのフィンランド語、スウェーデン語、英語版、そして異なる長さで、分類メタデータ、大きな分類のためのインデックス (探索用語) や異なる時期の分類の間の転換キーを含んでいる。分類データベースは、ハンドブックといった分類生産物のための生産環境とサービスの拠りどころである。分類データベースは、フィンランド統計局の総ての従業者が使用できる。

分類サービス部門は、自らのインターネット・サイトを持ち、これは、フィンランドの政府統計で使われる主な標準分類と関連メタデータを示す。インターネットのアドレスは、

http://stat.fi/luokitukset/index_en.html

Eurostatが使う分類基準と幾多の国際的に広く使われている分類基準へのリンクが、RAMON分類デー

データベースに取り入れられた。インターネットのアドレスは、<http://europa.eu.int/comm/eurostat/ramon/> 同様に、国連が発行している分類基準は、国連統計部（UNSD）のインターネットページから <http://www.un.org/Depts/unsd/class/> で入手可能である。

2.4 枠組みとカバレッジ

視角と目的

統計調査の目的は、目標集団と関心ある単位を決定する。それらの単位は母集団を形成する。ほとんどすべての統計調査の目的は、最終結果を全目標母集団に対して一般化することである。というのは、統計はほとんど共通して、全目標母集団について計算された分布、合計、平均、分位と比率の形で表現されるからである(Levy and Lemeshow, 1991)。

統計調査の結果の一般化可能性、適合性および正確性は、データ収集における正しいサンプリング枠組みの利用を必要とする。全数調査と標本調査の両方において、枠組みのカバレッジ、完全性、適時性、情報内容および正確性は、その適性(suitability)に関して決定的な要因である。また、フレームの単位と調査で測定された単位の間につながり、そしてその定義の違いと類似性を評価することは、中心的重要性を持つ。実際の問題は、住居単位に基づくレジスターにおける世帯と家族構造の識別にあらわれるかも知れない(Kish, 1965)。

この章は、フィンランドの政府統計のための統計調査で使われているフレームの主要な特徴を規定し、調査結果の一般化可能性と品質との関連でフレームの役割をかえりみる。フィンランド統計局で使われている異なるフレームの性格と特別な特徴は、2.4節でより詳細に記述されている。

原則

統計調査の目標母集団は、各母集団単位の算入を明確に決定できるように定義するべきである。調査の計画における次の段階は、代替的フレームとその適性を評価することである。統計調査の性格はまたこのフレームの適性に対する効果を持つ。

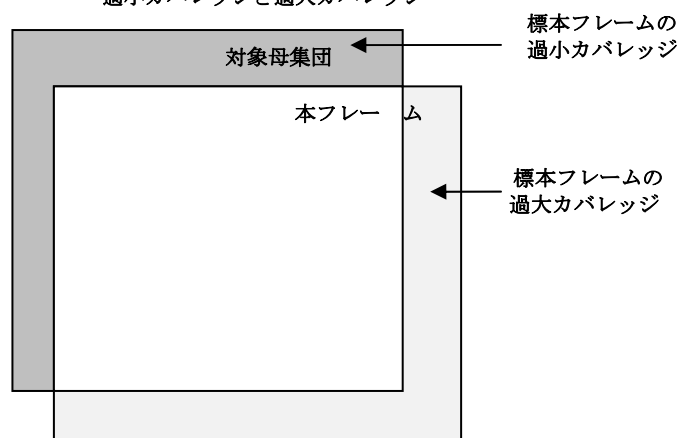
- 全数調査では、フレームとそれにリンクされる補助的情報は、統計調査の内容要求に対応しなければならない。
- 標本調査では、フレームが選択された標本抽出法に必要な情報、例えば層化標本をひきだすために必要な情報をふくむことが不可欠である。

目標母集団の全体は、調査によってカバーされるべきだが、実際においては、完全なカバレッジを達成することは稀である。したがって、目標母集団の概念とフレーム母集団の区別をしなければならない。目標母集団は、関心あるすべての単位をふくむのに対して、フレーム母集団は、把握できる単位をふくむ。最も一般的に使われている標本フレームは行政レジスターである。

フレームにふくまれている情報は更新されるべきである。レジスターを使うときには、階層と分類は、最新の可能な情報に対応して更新されるべきである。標本フレームが、レジスターを結合して作られるときには、レジスターデータを結び付けるために使われる選択された人の個人識別コードは検査されなければならない。例えば、これは個人および世帯調査に適用される。レジスターに含まれているデータは、調査にとって特に重要な人口の下位グループに対する標本抽出を行うために使われるべきである。この目的のためには、層化標本抽出、2段階標本抽出、あるいは変数確率を伴う標本抽出を使うことが可能である。

行政レジスターが標本フレームとして使われるときには、目標母集団に属する単位とともに他の単位をふくむ可能性がある。そういった単位は過大カバレッジに属し、通常は、標本抽出の段階でレジスターから認定することができる。行政レジスターはまた過小カバレッジをもつ。すなわち、目標母集団に属する単位がレジスターによってはカバーされないということである。標本フレームの過大カバレッジは測定できるが、過小カバレッジは推定するのが難しい(Djerf, 2000). 過小カバレッジについては、可能なら、フレームのカバレッジを改善する方法を決定することが必要である。

図 2.4 統計調査における目標母集団、調査母集団とフレームの過小カバレッジと過大カバレッジ



母集団の変化は通常、個人と世帯に関するよりもビジネス調査においてよりも大きい。したがって、特別な注目が標本フレームの適時性に払われるべきである。ビジネス調査では通常、企業の大きさが標本を選ぶときに考慮される。これは、階層化あるいは企業規模に確率的に比例させて企業を選択することで行うことが出来る。最大の企業は、そこから総ての要素が標本に取り入れられる階層に置かれることが多い。これは総てを取り出す階層(take-all stratum)と呼ばれることが多い。フレームは、推定の段階のために、レジスターから引き出された補助的情報によって更新されるべきであるし、この情報はまた過小カバレッジを調整するために使用されるべきである。

調査の目標母集団が定義されたときには、標本フレームとそのありうるカバレッジ誤差が調査母集団の正確な定義のために検査されるべきである。いくつかの調査ではいくつかの単位は母集

団から除外される。

- **地理的根拠:** 例えば, Cland とフィンランドの外部の群島の市はいくつかの世帯面接調査からは除外されることがある。
- **単位に特化した人口動態的根拠:** 統計調査は, 成人人口といったある人口動態的特性によって識別される個人に向けられることがある。同じように, ビジネス調査は, 例えば, ある産業の企業を中心にするところがありうる。
- **グループに特定の根拠 (Group-specific grounds):** 標本単位がフレームのそれとは異なるタイプであるなら, それは, あいまいさなく定義されなければならない。次に, 調査の範囲内に明確にはいない単位は, フレームからは除外されるべきである。例えば, 幾つかの国では, 施設人口に属する人は世帯調査の標本フレームから除外される (Rossi, Wright and Anderson, 1983)。

ガイドライン

フィンランド政府統計で標本フレームを選ぶためのガイドライン:

- 標本フレームは, どのレジスターが標本フレームをつくりだすのに利用可能か, それを使用する費用, および, レジスターは, 最善の可能な調査フレームを形成するためどう合併されるか, を見出すために評価されるべきである。
- 個人レジスターは個人の標本を選出するときに適切であり, 世帯の標本はまた以下の方法でそれらのレジスターから引き出されるべきである。個人が選ばれる住居単位は調査にふくまれる。そのあと, 世帯の定義と選ばれた人あるいは世帯の他のメンバーから得られた情報に基づいて, 調査員が世帯構造を作るためのものになる。それにもかかわらず, 住居単位の大きさといった住居単位情報は, データ収集段階の前に世帯に関する最善の推定値として使うことが出来る (2.4節参照)。
- ビジネス調査に使われる標本フレームは企業・事業所レジスターあるいはそこからひきだされる他のいくつかのデータ集合である。行政レジスターにおける変化は, 企業・事業所レジスターに記録されるとき, 統計単位に転換される行政事象とみなすことができる。法的単位は必要なら統計単位に転換される。
- 農場標本を選出するために使われる標本フレームは農場レジスターである。
- 標本選出の後に, 実現した標本がフレーム分布と矛盾しないことを検査するために, 補助的情報の全体 (あるいは他のパラメーター) が, 標本と標本フレームの両方から考慮されるべきである。大きな標本誤差の場合には, 標本選出と検査は更新されるかも知れない。後の使用のために可能なだけ多くの補助的情報を標本に付与することが賢明である。
- レジスターと標本フレームの品質は収集データによって評価されるべきである。あらゆる 過大カバレッジ, 分類変数における誤差, 過小カバレッジは, 他の情報源を使って評価されるべきである。適時性とアドレス情報の完全性, および分類変数の正しさについての調査員あるいはその他どこからでも得られたフィードバックが続くべきである。

- フレーム母集団，標本フレームおよびそのカバレッジが調査結果との関連で報告されるべきである。

文献

- Djerf, K. 2000. *Properties of Some Estimators under Unit Nonresponse*. Statistics Finland. Research Reports 231. Helsinki.
- Kish, L. 1965. *Survey Sampling*. New York: John Wiley & Sons.
- Levy, P. S. & Lemeshow, S. 1991. *Sampling of populations: methods and applications*. New York: John Wiley & Sons.
- Rossi, P. H., Wright, J. D. & Anderson, A. B. 1983. *Handbook of Survey Research*. London: Academic Press.

2.5 行政記録とレジスター

視角と目的

フィンランドの政府統計のシステム内で，行政記録とは，何よりも政府統計生産以外の他の目的のために収集されたすべてのデータをさす。この定義の範囲内の下では，公共機関以外の機関は行政記録を集め，保持することができる。この定義は，政府統計に向けての行政記録の最も重要な特徴は何かに注意をひいている。というのは，データは他の目的のために収集されているので，それらを使用することは，回答負担を増やさないからである。統計の生産にとっての費用はまた直接的収集に比較して非常に低い。これらの長所は非常に大きく，行政記録の使用とむすびついた短所を通常は越えている。統計法（280/2004）はまた，統計生産において必要なデータが行政記録から入手可能な場合にはいつでも，回答者負担を最小にすべきことを強調している。

行政的レジスターは，行政的記録の部分集団である。行政記録が単位レベルのデータから成り立っているなら，それはレジスターと呼ぶことができる。この定義内では，データが関係データベース（*relation database*）に貯蔵されているとしても，われわれはレジスターについて語るることができる。関係データベース内のデータは，情報技術の意味においてレジスターを構成しないかもしれないが，われわれは，このデータベースは一つあるいはより多くの論理的なレジスターを構成すると言うことができる。

統計レジスターは，何よりも統計生産のニーズのために維持されているレジスターである。大部分の場合に，統計生産以外のいかなる他の目的にも統計レジスターを使用することはフィンランド統計法で実際に禁止されている。大きな統計レジスターは，行政レジスターを基礎にしている。その場合，統計生産に関係する変数は，1つあるいはそれ以上の行政レジスターから統計レジスターに集められる。

レジスターの利用における重要な概念は、コードシステム—このコードによって、各データ単位をあいまいさなしに識別できる—である。これは、ほとんどのレジスターがそれを使用するならば、一般的コードシステムと呼ぶことができる。個人識別コードあるいはビジネス識別コードのような一般的コードシステムは、収集されたデータとレジスターデータを直接的に結合することを可能にする。

レジスター

情報レジスターは通常、電子的形式の単位レベルの(合計的)データ記録であり、全体のうちのある集団ないし集合を構成し、定期的に更新される。

レジスターの過半数は行政的レジスター、すなわち、省庁のニーズに奉仕している。例としては、人口情報レジスター (Population Information Register)、税レジスター、雇用者年金、求職者、年金生活者および徴兵のレジスター、がある。レジスターのごく小さな部分だけが本来的に統計目的のために作成された。それらの例には、企業・事業所レジスター、中央政府機関と作業単位をカバーする関連するレジスター、地方政府の作業単位のレジスター、企業集団のレジスターがある。これに加えて、修了した教育と学位のレジスターは、人口の基礎レベル後の資格をふくみ、学生レジスターは、基礎レベル以後の教育下の学生をふくむ。

フィンランドにおいて極端に重要なレジスターは中央人口レジスター、現在の人口情報システムである。それは、1960年代後半に構築され、フィンランドに永住の場所を持つか、持っていた総ての者からなる。人口情報システムはまたいわゆる歴史情報、すなわち、移民と死亡をふくむ。人口情報システムは、人口変化の情報、すなわち、出生、死亡、入国移民、出国移民、国内移民、結婚、離婚、養子、氏名変更について毎日更新される。人口情報システムは、国の地方管轄区 (State Local Districts) の地方レジスターの結合であり、また州の地域事務所で更新される。1980年に人口情報システムは、建物、住宅、ビジネスの建物、および自由時間の住居をカバーするものへ拡大した。すべての住宅の建物と仕事場を持つ建物が含まれた。人々は住宅と、企業はビジネスの場と、識別可能なデータによって結びつけられている。

他の電子的データ集合もまた、実際には、クロス・セクション・データ集合であり、変更情報とともに更新はされないものであるが、レジスターと呼ばれることが多い、それらの例は、所与の年の職業訓練学校あるいは大学や上級二次学校への志願者と入学者をふくむ共同選択レジスターである。次の年に、このレジスターは、人についてのほとんど完全に異なる集合—その根拠は事実上、ただ電子的データ集合であるからである—からなっているだろう。同じことは、フィンランド統計局の年次的データ集合—それらはレジスターと呼ばれることが多いものであるが—の多くについて当てはまる。

税レジスター (現在の税データベース) は、1960年の終わりに構築された。それは、毎年の税を支払う人から構成されており、彼らの年間の納税の対象になった所得と財産の情報をふくんでいるのだが、レジスターと電子的データ集合の間の中間の形のものとして考えることができる。このデータ集合は、人口情報システムの死亡、除去、家族情報といった変化によって更新されているが、税データは、所得稼得者自身、彼らの使用主と年金機関から毎年収集されている。その

ため、前年のデータは更新されていない。しかし、税データのあるもの、例えば、資産税や売り上げからの利益を判定する資産はレジスターに保存されている。これらの項目は、毎年変化が起らないなら、毎年再度報告される必要はない。

企業統計のための最も重要なレジスターは税務所によって維持されている。事業税を意図した最初のレジスターは1970年代に開発されて、例えば、売り上げ税に責任を持つ事業をふくんでいる。1980年代に税務所は、事業税の情報システムを改訂し、電子的形式に転換し、取引登録コード(現在の事業・会社コード)を導入した。これらの改訂の後に、フィンランド統計局はほとんど悉皆的な企業・事業所レジスターを構築することができた。ここでは、事業の立ち上げと閉鎖に関する情報がほとんどリアルタイムで保持できていた。

1990年代までは、政府統計の事業税レジスターの利用は、企業・事業所レジスターを中心としていた。1990年代には、ビジネス統計における行政記録の利用において多くの前進があった。第一に、1995年から1996年にかけて構造統計の改訂があり、このとき事業税レジスターが重要な情報源になった。企業の売り上げと賃金表に関する月次統計の生産は1998年に出発した。それらは、税務所の支払い管理レジスター—これは、付加価値税と登録された使用主が支払った賃金・俸給についての情報をふくんでいる—に基づいている。

原則

単位を識別するコード

コード情報は行政記録とレジスター・データの単位を識別するために使われている。これによって、異なるレジスターの単位を互いにリンクさせることができる。例えば、人を住居と、住居を建物と、建物を不動産と、そして雇用者を企業と、である。識別の道具は、人のレジスターにおいては、個人識別コードであり、建物・住居レジスターでは不動産コード、建物コードあるいは住居コード、企業・事業所レジスターでは事業所コード、ビジネス識別コード、そして車輛レジスターでは、各自動車のレジスター番号である。それらのうち、最も知られているのは、正しく書かれていることを保証する管理文字をふくむ個人識別コードである。それは事実上変化せず、利用者には周知で、異なるレジスターで広く採用されていて、全く曖昧性を持たない。コードが無い場合は、氏名および/あるいはアドレス情報で、部分的に補うことができる（例えば、いわゆる年金計画番号を、ビジネス識別コードと連結するとき）。

他の情報を、個人識別コードによって、個人と結びつけることが可能になる。すなわち、他の情報として、産業の職場部門、就業および/あるいは失業の期間、病気期間、年金と学習、獲得資格、所得、資産、負債、所得支援の受領、車輛の所有、建物、住居、自由時間の住居の所有権、および住居単位情報がある。

統計生産に使われる主要な行政記録とレジスター

人口と個人についての情報

- 人口、建物、住居、事業の建物、および自由時間の住居についての情報を含んでいる人口情報システム（人口レジスターセンターによって維持されている）、
- 高齢、傷害および失業年金の人口、および住宅手当と疾病に関する情報（フィンランド社会保障機関によって維持されている）。
- 所得支授受領者のレジスター（国立社会福祉保健研究開発センター,Stakes, によって維持されている),
- ケア報告レジスター（Stakesによって維持されている)
- 徴兵レジスター（フィンランド軍隊防衛スタッフ局が維持),
- 雇用年金システムにおける雇用データ（中央年金保障機関，国家財務部，地方政府の年金機関およびいくつかの小さな情報生産者，によって維持されている)
- 求職者と労働市場訓練のレジスター（労働省によって維持されている)
- 学生レジスター(フィンランド統計局他が維持),
- 修了教育と学位のレジスター（フィンランド統計局が維持),

企業に関する情報

- （顧客データベースからの）納税顧客レジスター，（付加価値税と使用者支払についての月次データをふくむ）付加価値税のデータベース，ビジネス税データ，納税レジスター，所得と財産データ，使用主への年次支払管理データ。
- 税務所とフィンランド国家特許・登録委員会（National Board of Patents and Registration of Finland）が共同で維持しているフィンランドビジネス情報システム。
- フィンランド国家特許・登録委員会によって維持されている取引レジスター。
- 企業・事業所レジスター（国と自治体の単位をふくむ）(フィンランド統計局が維持).
- 農場・園芸企業レジスター（フィンランド農林省も情報センターによって維持されている)

他のレジスター

- 車輜レジスター（車輜登録センターが維持)

統計生産における行政記録とレジスターの使用

通常、レジスターは全数統計と変化統計の両方の統計生産に使われる。例えば、人口統計と人口構造統計は、レジスターの全員を数えたデータから生産される。人口動態統計（出生、死亡、国内移民、移出、移入、結婚と離婚）はレジスターの更新データに基づいている。同じように、建物のストックと生産に関して、建物と住居レジスターからひきだされた統計は、レジスターの維持システムに基づく。全員を数えるデータの定期的維持はまた、車輜レジスター、修了教育と学位のレジスター、企業・事業所レジスター、求職者レジスター—そこから総計と変化統計が生産される—において遂行される。

統計の編集における行政記録とレジスターの利用は、統計生産の費用とデータ供給者の負担の両方を減らす。これは、フィンランド統計法（280/2004）が、布告している主な原則のひとつである。例えば、レジスターに基礎をおく人口センサスの費用は、調査票を伴った直接的データ収集の費用のほんの一部分にすぎない。センサスデータの幾つか—センサス週の主な活動、家族、世帯—住居単位、家族の参照人、社会-経済的グループ、仕事場と住んでいる場の間の距離といった—は、様々なレジスターデータの助けを以て(たとえ、それらがそのレジスターにも直接的には現れないとしても)推定される。

幾多の統計は、例えば、人口、住居、建物そして企業のストックについての総てを数える統計と、それらの変化についての情報を作り出す。変化についてのそういった情報は、出生、死亡、国内移民、結婚、離婚、入国移民、出国移民、建物の完成と住居、登録車両、事業の開始と閉鎖についての年次データからなる。

全数データがひとたびレジスターに入ると、統計を生産するためには、変化が起きた単位だけを扱う必要があるだけである。例えば、新しい建物だけが建物レジスターに取り込む必要があり、他の建物はそのまま残される。税レジスター、企業・事業所レジスターや雇用者年金レジスターといったいくつかのレジスターについては、ほぼすべての単位について年次的に新しい情報が収集される。そういったレジスターは、年次データと大きくは変わらない。幾つかのレジスターは、すべての単位とそれらの属性についてのデータは毎年収集されているので、実際には年次ファイルである。そういったレジスターデータの例としては、住宅手当と学生財政援助および、年間の制度的ケアの受給者からなるケア報告レジスターをふくむ。同じように、学生レジスターは各年の教育施設に入学した者をふくんでいる。学生レジスターからひきだされる変化を叙述する統計は、志願者、合格者および新入生と卒業生についてのデータをふくむ。

一般に、機関が行う告示は、正確な時点を与えるのに対して個人の告示に基づく変化についての情報は、ある程度主観的である。すなわち、例えば、一緒に住んでいる人々は、同棲の結合に異なった見方を持っているかも知れず、働いている学生は彼/彼女自身を学生として報告するかも知れないが、彼/彼女は、雇用者（週に少なくとも1時間以上稼得労働に従事しており、したがって労働力概念にしたがうと、彼/彼女は、彼・彼女を夕方の掃除人<evening cleaner>として報告すべきである)の基準に合致しているのかもしれないのである。レジスター・データが使われるときには、回答者の主観的考えの影響は減る。雇用関係を持つすべての学生は、被用者として記録され、最高の所得を受けている者が家族の参照人として記録される。

データの正しさ、欠如および論理性(logicality)はまた、行政記録とレジスターにおいて、精査されなければならない。これは、統計の生産者にとっての課題であり、特に、異なる組織が情報システムを維持するときにそうである。これは、利用可能なデータ出所と体系的に開発された位階的推計規則がデータの検査において、効果的に利用されなければならない理由である。例えば、企業税データは、幾つかの企業を完全に落としているか、誤りである。データの正しさは、企業がその勘定を閉じたかどうかを検査することで確認されるべきである。それで、普通の誤差のタイプは、論拠薄弱な勘定の中か

ら探ることができ、この勘定はプログラム可能な論理性と推計規則によって正すことができる（2.10節参照）。企業税に関するデータのあるものは、欠けているか、財務声明においてバランスできないなら、それは売上高データから補定されなければならない。というのは、売上高は、企業と事業所レジスターからすべての企業についてひき出すことができるからである。この補定は、例えば、産業や企業規模クラスを考慮に入れる高度に精巧な方法によって改善できる。しかしながら、職務、職業および所得についてのデータが異なるレジスターから結合されるとき、職業や産業部門が同じ雇用関係を叙述ことは常には確かではない。これは一貫性問題を生み出しうる。

最も早くても、行政記録とレジスターからのデータは、調査対象時期の後、数週間か数カ月で利用可能になることができ（車輛レジスター、求職者レジスター、人口情報システムおよび税支払管理データ）、最も遅い完了のときには、1年にかかる（例えば、税データベース）。適時性は、事象が登録される前の時間の長さ、データがレジスターで更新される頻度によって影響を受ける。人口情報は日々更新されるのに対して、所得のデータは年に1度更新される。しかしながら、もし移動した人が、それについて機関に告知することを怠るなら、このレジスター・データは更新されない。建物も完成、すなわち、それが建物レジスターに入る時点、は問題のある概念である。というのは、建物は、最終検査で完成したと宣言される前に、長い期間居住していることがありうるからである。

適時性はまたデータの悉皆性に影響を与える。例えば、税支払管理データを統計機関がはじめて受けとったとき、データは、企業の最終売上げの約80%を、登録された使用主が支払った最終賃金額の75%をカバーしている。個別企業についての回答が、遅れのため、あるいは企業が活動を停止したために、失われているのかどうかを知ることは難しい。したがって、与えられた月についてのデータは、それらが全部になるまでに、6カ月にわたって政府統計のために繰り返し請求されている。しかしながら、これは適合性の課題を削除はしない。月次の指標の場合には、パネル法は適時性というこの課題への最初の答であった。

年次的に生産される構造的企業統計は、企業税レジスターや企業・事業所レジスターのデータへの直接的問い合わせを通じて得られたデータの結合から編集される。データは、すべての企業の数パーセントから直接的に収集されたに過ぎない。月次売上高や賃金合計指標は、付加価値税（VAT）と使用主の支払いレジスターおよび企業・事業所レジスターのデータへの直接的問い合わせから得たデータを結合して編集されている。この後に、データは、すべての企業の1パーセント以下から直接的に収集される。構造統計のためと月次指標のためにデータを収集する構成の類似性にもかかわらず、それら2つの統計の集合の方法論的課題は異なる。構造統計においては、補定は行政データを統計利用に適するものにするために、その欠点を正すために使う方法に置いて主要な役割を果たす。対照的に、パネル法は月次指標における適時性と過小カバレッジに関連する問題を調整するための主な道具である。これらの統計の品質は、方法論的作業を維持することで改善されてきた。

ガイドライン

行政記録とレジスターの活用の前提条件

行政記録とレジスターは統計生産において使用できる前に、ある前提を満たさなければならない。個々の前提条件のすべてが絶対的に必要なのではなく、それらは常に、他の前提条件および代替的収集方法という見地から考えられるべきである。一般に、総ての行政記録が統計の編集者にとっての品質の課題を示す。経験によれば、それらの課題は、データシステムの管理者との協力とともに、持続的な方法論的活動によって応えることができることを示してきた。

内容と概念の前提条件

- **悉皆性 (Exhaustiveness)** : 使用されているレジスターは、意図している目標母集団をカバーすべきである (1.2.3節参照)。
- **単位の具体性 (Unit specificity)** : 基本的なレジスターデータは単位が具体的であるべきである。企業の合併や分離は他の課題をつくり出す。というのは、行政でのそれらの扱いは、必ずしも政府統計の要件と一致しないからである。
- **利用者ニーズとの一致 (Correspondence with user needs)** : レジスターデータは統計利用に対して適合的でなければならない。レジスター・データで使われる概念は、統計生産で取り上げられているものと十分に一致していなければならない (1.2.3節参照)。行政記録のデータは代替変数とみなされるべきことが多い。ということは、変数の代替値が他に獲得することが困難であるということである。したがって、そのデータは、統計データの必要に見合うように、ある仮定 (理論) に基づいて変換されなければならない。

信頼性 (Reliability)

- **カバレッジ** : 使用されるレジスターは意図した目標母集団を可能な限りカバーしなければならない (2.4節参照)。カバレッジの絶対的なパーセンテージを規定することはできない。というのは、完全に悉皆性をもったシステムはなく、またほとんどのシステムは過大カバレッジ、すなわち、目標母集団には属さない単位を持っているからである。人口センサスのカバレッジは国によって異なり、95%から100%である。もし、レジスターに基づくセンサスを作り上げるために使われるレジスターのカバレッジは、これよりも弱いなら、それらの使用は問題をもつ。
- **データの信頼性 (Reliability of data)** : 個別のデータは可能な限り信頼できるものであるべきである (1.2.3節参照)。これは、コードデータと属性データにかかわる。
- **データの修正 (Correction of data)** : レジスターを更新するシステムもまた悉皆的であるべきである。例えば、人口統計システムは、出生、死亡、結婚、離婚についてデータを非常に信頼性あるように獲得するが、アドレスの変化は拒絶される。同じように職業データは、住所の更新と関連してだけ更新され、住居に対しては、建築許可なしに大きな変更が行われる。結婚と同等と

考ええられる同棲のカップルは、統計的に推定されなければならない。

レジスターデータの適時性

- フィンランド政府統計データは更新されなければならない(1.2.3節参照)。適時性は、その事象が登録されるのにかかる時間とそのデータがレジスター内でどれだけひんぱんに更新されるかによって影響される。
- 異なる統計機関の間の協力と適時性の短所を管理する方法は、統計生産において改善され、発展させられるべきである。行政記録は定義によって、統計目的のために収集されてはいないので、統計機関は適時性に関連した問題に直接的影響を持つことはできない。

技術的前提条件

- 文書化:** 利用者がレジスターの活動、構造および情報内容の描写を得ることができるように、レジスターについての十分に優れた詳細な説明が、行われるべきである (3.1と3.2節参照)。
- データ処理:** レジスターは、統計生産での利用が可能ないように電子的形式であるべきである。企業の雇用者レジスターや協会の加盟者リストをふくむカード・ファイル形式の手作業で処理されたレジスターは、統計には稀にしか使われない。
- コード:** レジスターの単位が広く使われるコードを持っていることは不可欠である。
- 分類:** 属性データはそれらが分類できるようにコードづけられるか数字書式にされるべきである。データが職業といった特性形式におかれているなら、それらは例えば、自動コード化によってコードづけできる。ある特性の数値見出しは、手作業の処理を必要とする(2.3節参照)。

行政的前提条件

行政記録およびあるいはレジスターの管理者はフィンランド統計局にデータを渡す義務を負っている。フィンランド統計法(780/2004)は、フィンランド統計局に、統計を作成するために、政府や幾多の他の機関から必要なデータを獲得する権利を認めている(1.1.1節参照)。

レジスターデータの費用

統計法(780/2004)によれば、「統計目的でデータが収集されるときに、第一に活用される出所は、一般政府の課題を実施する際に蓄積されているデータ、使用主、自営業者、会社および基金の正規の活動の結果として生産されたデータであるものとする」。レジスターの活用のための費用は、他のいかなる代替的収集法のそれを上回るべきではない。フィンランド統計局は一般的に、そのデータの利用から起きるいかなる追加的費用だけを、データの管理者に支払う。

文献

フィンランド国内法

Statistics Act (280/2004)

他の情報

- Elvers, E. 1999. *Frame errors. In Model Quality Report in Business Statistics. Vol. 1. Theory and Methods for Quality Evaluation. Assessment of the Quality in Statistics.* Doc. Eurostat/Supcom 97/lot6/99/Final Report/Vol.1. Eurostat.Luxembourg.
<http://amrads.jrc.it/WPs%20pages/Quality/Documents/eqpbvol.1.pdf>
- Harala, R. & Tammilehto-Luode, M. 1999. GIS and Register-based Population Census. In Alho, J. (ed.). *Statistics, Registries and Science. Experiences from Finland.* Statistics Finland. Keuruu. 55–73.
- Ilmakunnas, P., Laaksonen, S. & Maliranta, M. 1999. Enterprise Demography and Job Flows. In Alho, J. (ed.). *Statistics, Registries and Science. Experiences from Finland.* Statistics Finland. Keuruu. 73–89.
- Myrskylä, P. 1991. Census by Questionnaire – Census by Registers and Administrative Records: The Experience of Finland. *Journal of Official Statistics.* Vol. 7. No. 4. 457–474.
- Myrskylä, P. 1996. Virallisten tilastojen tiedonkeruu. (Data collection of official statistics). In Niemi, H. & Tourunen, K. (eds.). *Tilastoista tiedoks korkea-asteelle.* (From statistical data to information – a textbook for higher education; in Finnish only). Statistics Finland. Helsinki.
- Myrskylä, P. 1999. New statistics made available by the use of registers. *Statistica Journal of the United Nations ECE.* Vol.16. No. 2–3. 165–180.
- Myrskylä, P. (eds.) 2004. *Use of Registers and Administrative Data Sources for Statistical Purposes - Best Practices of Statistical Finland.* Handbook 45. Helsinki: Statistics Finland.
- Ragnarsøn, R. 2000. *Utilisation of Administrative Data in the Manufacturing Statistics. The Norwegian Experience.* Meeting of INSEE's and Nordic NSI's Business Statisticians. Helsinki 11 – 12 May 2000. Statistics Finland. 2001. *Population Census 2000.* Handbooks 35. Helsinki: Statistics Finland.

2.6 標本抽出法

視角と目的

様々な理由で、標本調査は統計庁においてセンサスを代替することが多い。費用を少なくすることができ、必要とする現象についてのより深い情報を獲得することができ、総ての母集団要素を調べなくても良いので、結果をより速やかに推定できる。それらはまた、レジスターあるいは他の行政データから既存のデータへの追加的情報を獲得するために使うことができる。適切に設計された標本調査は、利用者に対して研究中の変数についての正確なパラメータの推定値を提供し、同時に全体的な回答負担を減らすことになる(Kish, 1965; Särndal et al., 1992)。

現代の標本抽出法はランダム化原則に基づいており、このため、確率抽出法と呼ばれる。その優位性は、標本の結果を全母集団に推定適用し(extrapolating)、それらの推定値の信頼性に関する統計の計算を可能にすることである。標本抽出法の選択は、費用とともに、関心を置く現象、調

査設計、母集団特性値、利用可能な標本フレーム、データ収集の方法論に依存する。

原則

標本抽出法は多くの仕方で区分できる。最初の区分は、データを収集する方法によるもので、要素（究極の抽出単位）からの直接的な抽出か、集落と呼ばれる要素の結合体からの抽出か、である。もうひとつの分類は、標本設計のための補助的情報の利用に基づくもので、標本フレームからか、その他の何らかの行政的出所か、である。さらに、選択は費用と統計的精確性の間の費用・便益分析に依存するものである。

要素レベルの標本抽出法は、統計的に有効で、特に標本をレジスター・フレームから抽出するときに容易に適用することができる。集落抽出法は、一般に要素標本抽出法ほど効率的ではないが、より費用効率的であることが多い。

以下の標本抽出法は、要素抽出と集落抽出に等しく適用できる。この後に、われわれは階層化と集落化といった枠組みに関連するいくつかのトピックスを論じる。

単純任意標本抽出

基本的な任意抽出法は単純任意抽出(SRS : simple random sampling), であり、これは自動的にウェイトづけが行われる、すなわち、母集団のすべての要素が等しく標本にふくまれる確率を持つことを意味する標本設計である。

単純任意標本をひきだす便利な方法は、母集団の各要素に(仮の)乱数を割り当て、データ集合を乱数に従って区分し、最後に、第一の要素からはじめて、通常は、必要な標本の大きさに到るまで、母集団のどの連続的部分からでも必要な標本の大きさを選出することである。選出は、母集団の各要素に付与された乱数に基づいており、したがって、抽出手続きにおいて補助的情報は使用されない (Lohr,2003)。単純任意抽出は普通、個人あるいは世帯の調査において使われる。それはまた、他の標本設計の統計的効率性を評価するときに、参照する抽出法として役立つ。

系統抽出

系統抽出 (systematic sampling) は、フレーム全体にわたって、ランダムな出発点から一定の区間をとって母集団の要素を選択する過程である。この方法は標本を手作業の引き出しファイルから便宜的に標本を取り出すために開発されたが、今日のコンピュータ・データベースにもまた有用である。

系統抽出の適用は、母集団の分類順からの補助的情報を使う。もし要素がランダムな順序にあるなら、系統抽出は実際的に単純任意抽出に合致する。しかし、フレームはまた、1つ以上の変数にしたがって配列されていることがありうる。それらの場合、この方法は、インプリシットな階層化と呼ばれ、ランダムな順序よりもより正確な結果をもたらす。例えば、もし関心の焦点がフィンランド人口における空間的に相関している現象にあるなら、地理的に区分された人口レジスター (例えば、人口データベースの住居コード) からの系統標本抽出が正しい標本を作るだろ

う。

しかし、母集団が何らかの（隠れた）順序あるいはつながりをふくむなら、系統抽出は、真の母集団の多様性を反映しない非常に類似した要素から成り立つ標本をもたらす、これによってその標本は誤った、あるいは偏った結果をもたらすことがある。

規模比例確率抽出 (PPS : Probabilities Proportional to sizesampling)

もし標本を、単位の大きさが非常に異なる歪んだ分布の母集団から選出するときには、各母集団要素に、大きさの尺度に関連する包含確率(inclusion probability)を付与することによって正確性を改善する方法がある。これは、われわれを規模比例確率 (PPS) 抽出に導く (Särndal et al., 1992; Lehtonen and Pahkinen, 2004)。

PPS 標本抽出は、すべての母集団要素が包含確率を計算するために使われる大きさの尺度を持つことを必要とする。ときには、最大の母集団要素が支配的なので、その選択確率が1になることがある。そういった要素は、確実に選択され、残りの母集団要素の選択確率が再計算されるべきである。

PPS標本は統計的に非常に効率的である。しかし、そこには、大きさの尺度として使われる変数あるいは大きさの尺度と高い正の相関を持つ変数を分析するためにだけ適切なことが多いという欠点がある。

標本フレームの部分母集団への分割

標本フレームは、適切な補助的情報によって相互に排他的な部分集団に分割できる。これを行うふたつの理由がある。この分割が、各部分母集団が、さらにそれ自身で標本抽出フレームとされるように行われるなら、われわれは階層化を語っている。他方で、部分母集団は、我々の関心のある単位の「より高いレベル」の結合として役立つだろうし、われわれは、それらの標本を選択するが、究極的にはその各々の内部の要素を調べる。この場合、われわれは集落標本抽出について語っている。

集落抽出 (Cluster sampling)

標本選出で測定されるべき要素の自然な結びつきを利用することが良いアイデアであることがときどきある。そういった自然的結合とは、例えば、測定されるべき要素をふくむ世帯、事業所、学校あるいは他の単位でありうる。もし選ばれた各集落のすべての要素が調査されるなら、われわれは一段集落抽出を話している。そうではなく、選択された集落が、それに続く副次標本抽出のための一層の集団とされることがある。後者の場合に、われわれは二段あるいは多段の集落抽出を語っている。1例としてわれわれは労働条件の調査を考えることができる。すなわち、第一段標本は企業からなり、そこに属する雇用者が測定される要素である。各企業のすべての雇用者がふくまれるなら、われわれの設計は一段であるが、もし職業別等の雇用者の幾人かを副次抽出するなら、われわれは二段階集落設計をしている。

集落抽出は、選択された各单位から多くの要素についてのデータを得るので、データ収集費用を減らす。しかし、単純任意標本抽出と比較するとき、標本分散を増加させるという欠点を持つ。効率性の喪失は、階級内相関として知られている集落内の要素の類似性によって引き起こされる。このようにして、集落標本は、SRS設計で選択された標本ほどは母集団の多様性を必ずしも含まない。

階層別抽出

階層別抽出は、母集団の要素が補助的変数の助けを借りて階層と呼ばれる同質の副母集団に分割されることを意味する。歪んだ母集団の階層化（あるいはPPS標本抽出）によって、最大の単位からの正確な情報を得ることが常に望ましい。

階層化は基礎的に、すべての母集団要素が階層を構成することを許す情報を持っていることが必要である。究極的には、各要素は、ひとつの階層に属することができるだけである。しかし、標本設計あるいは標本抽出率は、異なる階層間で違いがあってもよい。ときとして、ある階層におけるすべての要素がセンサスにおけるように調査されるべきであり、一方で、他の階層には標本抽出だけ適応されることがある。世帯調査においては、例えば、地理的領域、年齢あるいはジェンダーといったかなり基本的な人口の階層化基準が使われる。

種々の階層の標本の大きさは、階層化標本選出での追加的手段として計算されるべきであり、配分(allocation)として知られる。もしすべての階層が同じ標本の大きさを持つなら、われわれは等配分を話している。同等割り当ては、階層間の比較について特に有効である。もし、標本の大きさがすべての階層の母集団割合に比例するなら、われわれは比例配分を語っている。それは、基礎的SRSよりもわずかに効率的であり、最小の階層ですら標本の中に要素の比例割合を持つことを保証する。もし母集団が大きく歪んでいるなら、最適配分、あるいはパワー配分のいずれかを使用できる。両法とも、標本を最大の、あるいはそうでなければ最も重要な単位に導き、統計的には非常に効率的である(Lehtonen and Pahkinen, 2004)。極端な場合には、最大の単位は、確率1で選出されて、自己表出階層(self-representing strata)として知られている。推定はそういった単位からの情報なしでは信頼できないものになる。

複合的標本設計

複合的標本設計は、基礎的設計の組み合わせからなる。典型的には、ひとつの複雑的抽出設計は、階層別多段集落設計である。ここでは第一段階の標本抽出単位は、地理的順序、そうでなければ分類された集落から、PPS抽出で選択される。究極の単位(測定されるべき要素)の選出は、ほとんどいつもSRSあるいは体系的抽出設計によって行われる。注意深い階層化、集落化、および種々の標本選出段階からの情報の最大の利用によって、非常に費用効率的な標本抽出設計を提供できる。

もし母集団がレジスターあるいは他の行政記録の中にあるなら、2あるいは多段標本設計を使用することも可能であろう(Särndal et al., 1992)。第一局面の標本は、補助的情報が、調査や他の出所から合併されるマスターサンプルと呼ばれる大きな標本であることが多い。マスターサン

プルは、さらに分析されて、最終的には最終標本を選出するための第二局面の標本選出フレームとして使われる。多局面抽出は、望まれる現象に向けた標本と、推計におけるその情報からの多様な効果を狙って、補助的情報の効果的な利用を可能にする。

割り当て標本抽出 (Quota sampling methods)

割り当て抽出とは、多様な非確率的標本選出法である。これらは、ある数の事前に規定しておいたタイプの回答者を獲得する、すなわち割り当てを満たす狙いで、特に市場調査に適用される。割り当て標本選出は、連絡を取ることや参加することの判断が調査トピックへの関心に影響を与えるので、選出バイアスに陥りやすい。

切り捨て (Cut-off) 標本選出は、ビジネス調査においてはかなり普通である。ここでは境界線、最も多いのは低い境界線である、を定めなければならない。これによって境界線より小さな単位はフレームからはずされ、したがって選出されることはない。このように、切り捨て標本選出は、推定結果は過小カバレッジを持つ点で問題をもつ。古い標本選出フレームの場合には、母集団の変化は大きく、切り捨てた標本はバイアスを持つ結果をもたらすことがある。したがって、切り捨て標本選出の利用では、境界線の選択の理由の説明とともに、カバレッジの分析と問題の統計の一般化が常に叙述されるべきである。

標本の大きさの決定

標本の大きさは様々の要因に影響される。すなわち、パラメータの推定値に要求される正確性、小さな領域や副次領域がカバーされる仕方（すなわち、分類）、および母集団の異質性である。当然のことながら、利用可能な資源もまた標本の大きさに影響する。正確性 (accuracy) とともに精度 (precision) は、標本の大きさが増加したときに改善されるということができる。というのは、推計値の標準誤差が減少するからである。しかし、この改善は標本のサイズの増加に比例するのではなく、より遅く起きる。

信頼性 (Reliability)

標本の信頼性は、統計的な品質指標で測定される。それぞれの確率的標本選出法は、標本選出設計や標本のサイズによる点推定値の変異性を基本的に叙述する標本分散の自身の指揮を持っている。標本分散はさらに、標準誤差、変異係数および信頼区間（あるいは誤差限界）のような幾つかの普通に使用されている品質尺度を計算するために使われている。品質尺度は、結果の発表基準を検討する場合とともに、統計的推論で使われる。

統計的品質指標に加えて、標本調査における他の誤差源泉にもまた注意を払うべきである。それらには、例えば、フレームでのカバレッジ誤差、無回答および処理誤差がある。品質指標の提示は、フィンランド政府統計の品質基準を叙述する節で更に論じられる (1.2.3節参照)。

標本の調整

回答者負担を減らすためには、定期的に繰り返される調査における連続的標本もまた調整されるべきである。積極的調整手段とは、連続する標本において可能な限り多くの同じ単位を含めることを意味する。これは、パネル調査において、可能なだけ信頼性を以て変化を推定する必要がある場合に典型的である。他方で、回答者負担を減らすために、連続する標本への同じ単位の選出を避けることを試みるマイナスの調整もある。調整は、結果にバイアスを生じないように行われるべきである。これが不可能なら、結果は、適切な統計方法によって調整されるべきである。

ある企業調査の狙いは、母集団における変化を調べることにあり、そういった場合には、十分な数の同じ単位が連続的標本に含まれることが重要である。同じ単位への行き過ぎた負担を避けるためには、単位は含められたり除外されたりローテートされるべきであり、同じ単位が連続した、あるいは現在の調査に含まれないように、様々の調査が調整されるべきである。この調整は、回答者負担が全体の企業母集団に可能な限り等しく分配されることを保証するだろう。しかし、標本の調整は、最大で最も支配的な単位に適用することはできない。それらは、推定値のカバレージを保証するためにすべての標本に対して選択されることが多い(Teikari, 2001)。

勧告

- 常に確率的標本抽出法を適用するべきである。単純ランダム標本抽出と系統抽出の両方が、標本がある副次母集団を狙うことができないときには、多目的調査の適切な基本的方法である。
- 系統標本選出は、標本がレジスターあるいは他の行政データ出所から選出されるときには、実用的な方法である。この場合、フレームの分類順序は、正しい母集団分布をもった標本をもたらすことができる。
- 階層別抽出は、標本内のある副次母集団から要素を得る必要があるときに勧告される。
- 単純任意選出は、ビジネス調査のためには、ビジネス母集団の歪みが理由で、適切な標本選出設計ではない。ビジネス調査において最も多く使われている標本選出法は、産業と規模階級別に階層化が行われる階層別標本選出である。これに加えて、PPS 標本選出法が使われる。最大の要素は、確実な単位として選出されることが多い。
- 複合的標本選出設計の便益を、その障害や技術的困難に照らして注意深く検討するべきである。2あるいは多段階の設計は費用と統計的正確性の両方の点で効率的でありうる。他方で、多段階集落標本設計は、統計的見地からは非常に非効率でありうる。
- 繰り返しの標本調査を扱うときには、特に、パネルおよびビジネス調査においては、回答者負担を減らすために、連続的標本は調整されるべきである。個人と世帯の回答者を、ある期間にわたって、類似の調査から解放する実践がある。
- 標本選出の変異性と他の誤差源泉を測定し、分析し、文書化しなければならない。最も重要な調査パラメータは、その各々の標準誤差、設計効果の推定値、回答あるいは無回答率を他の適切な品質指標とともに報告されるべきである。
- 標本選出のウェイトが標本調査データに付与されるべきである。ウェイト手続きは2.11節で

叙述されており、推定と調査データの統計分析法は、2.12節に示されている。

文書化

標本調査の目標母集団の定義、それに適用される制約やそれらの根拠、および標本選出および標本選出フレームの正確な叙述を報告することが重要である。品質報告書は、最も重要な推定量の式、変異性と誤差—特に無回答誤差—の様々な誤差、およびそれによって引き起こされるウェイトの調整の必要、を含むべきである。品質と方法報告書の詳細な方法論的叙述は、正しい設計に基づく標本変異の計算式、それとともに無回答誤差と他の誤差源の分析をふくめて、標準誤差や他の品質指標を可能な限り正確な計算を促進するべきである。

文献

Kish, L. 1965. *Survey Sampling*. New York: John Wiley & Sons.

Lehtonen, R. & Pahkinen, E. 2004. *Practical Methods for Design and Analysis of Complex Surveys*.

2nd ed. Chichester: John Wiley & Sons.

Lohr, S. 1999. *Sampling: Design and analysis*. Pacific Grove: Duxbury Press.

Särndal, C.-E., Swensson, B. & Wretman, J. 1992. *Model Assisted Survey Sampling*. New York:

Springer-Verlag.

Teikari, I. 2001. Controlling the distribution of response burden in longitudinal and cross section

business surveys. *Statistics Finland. Research Reports 232*. Helsinki.

追加情報

Cochran, W. G. 1977. *Sampling Techniques*. 3rd ed. New York: John Wiley & Sons.

Groves, Robert M., Fowler, F.J., Couper, M.P., Lepkowski, J.M., Singer, E., & Tourangeau, R. 2004.

Survey Methodology. New York: John Wiley & Sons.

Pahkinen, E. & Lehtonen, R. 1989. *Otanta-asetelmat ja tilastollinen analyysi. (Sampling Designs and Statistical*

Analysis; in Finnish only). Helsinki: Gaudeamus.

2.7 データ収集法

視角と目的

この節は異なるデータ収集法の特徴と適切な方法の選択に影響する諸要因を検討する。データ収集方法とは、標本単位の選出とデータをそこから収集する方法の両方である。すべてのデータ収集プロジェクトは、この節では調査と呼ばれる。ビジネス調査と個人あるいは世帯の調査は、何よりも目標母集団の相違によって相互に異なる。しかし、方法的な違いは大きくはないので、ここでの関心の焦点は、センサスあるいはビジネス調査ではなく、主として世帯あるいは個人の

標本調査となる。

データの収集法の選択は、調査の設計において孤立した判断ではない。というのは、それは、調査の設計、標本単位の選出、無回答率やデータ収集の費用といった統計調査過程の全体に影響を与えるからである。データ収集法はまた、調査票の設計と用意についてとともに、調査の予定と最終結果の品質に関係する。異なるデータ収集法は非常に異なる標本の大きさを可能にする。回答者の数は、結果を種々の母集団の構成部分に一般化できる信頼性の度合いを決定する（2.6節参照）。

調査が異なるタイプに区分されるときに、3つの要因、すなわち、一般的調査設計、データ収集法、データ獲得のための技術、が通常検討される。

調査はロンジチュージナル調査とクロスセクション調査にわけることができる。ロンジチュージナル調査においては、データは多かれ少なかれ定期的に、標本抽出単位から1度以上収集される。他方で、クロス・セクション調査においては、データは1度だけ収集される。ロンジチュージナル調査においては、データは、各回同じ仕方でか収集されるか、方法が変わる。パネル調査はロンジチュージナル調査の特殊なタイプである。フィンランド統計局の最も知られている個人パネル調査は労働力調査であり、最も知られているビジネスパネル調査は、消費者物価指数調査である。

表2.1 管理方式とデータ捕捉装置別調査タイプ

管理方式	データ捕捉装置	
	紙の調査票	電子的調査票
自己記入	自己管理調査票、日記	インターネット調査票（CAWI） コンピュータ支援の自己面接（CASI）
面接	面接者が管理する調査票	コンピュータ支援個人面接（CAPI） コンピュータ支援電話面接（CATI）

データ収集の管理様式とデータ獲得に適用される技術を区別することが重要である。管理様式においては、面接者が使われるか（面接者管理）、回答者が自ら答えるか（自己管理）が区分される。他方で、データ収集法が収集されたデータの文書化の様式、言い換えれば、紙の調査票と何らかの種類の電子的データ捕捉装置のどちらを使用するかを決定する。

管理方式は、自己管理—回答者自身が質問を読み、調査票に記入する—と、インタビュー者管理方式—インタビュー者が質問を読み上げ、回答を書き込む—にわけることができる。インタビューは、電話によるか、面談によって行われる。両方の管理式とも、データ獲得の2つの技術、すなわち、回答が紙の調査票への記入か、電子媒体への記入か、が適用される（表2-1参照）。

原則

—調査の目標は、調査の目標は管理の様式の選択をガイドするべきであり、逆ではない。選択

基準は異なる種々の方式の意味であるべきである。調査設計の出発点は、結果となる推定値の必要な正確性、必要とされる標本の大きさの評価、あるいは最終標本の回答者の数、に基づくべきである。ほとんどの場合、調査の信頼性 (reliability) は標本の大きさが増えるとともに高まる。もうひとつの重要な要因は、各回答者から得られる情報の量と質である。例えば、長い面接は電話で行われるべきではない。同じように、面接インタビューは、目に見える長さあるいは絵画などの対象を回答者に示す必要があるときの唯一の選択である。

調査設計における重要な要因は、一層の利用のためにデータが入手可能な日付けであることが多い。もし最終締め切り日が分かっているなら、中間段階の締め切りもまた決定される。調査が他の関係者との協力を必要とすることもまた心に留められるべきである。したがって、日程の計画づくりは通常は同意や適応を求めている。

生産の日程

生産の日程を計画する際には、調査のあらゆる段階が考慮されるべきである。必要な段階の数は調査によって異なる。例えば、電子的データ収集装置の仕上げとテストは、紙装置の設計と印刷よりも多くの時間を必要とする。他方で、データは、うまく設計された電子的データ収集の後には、一層の利用のためにより速やかに利用可能になる。というのは、一層のデータ入力是不要になり、電子的調査票にプログラムされた規則によって、自動的にチェックされるからである。

他の関係者の日程はまた、調査の生産日程に影響を与える。特に面接調査においては、現場の面接単位あるいはCATIセンターの総作業量が、データ収集の日付と継続期間に顕著な影響を及ぼす。これに加えて、面接員が十分に訓練されてきたこと、彼らへの作業指令は注意深く作成されたことは非常に重要である。データ・エディティングとデータ入力といった他の資源の利用可能性もまた生産日程に影響する。

データ収集の期間もまた管理方式に依存する。一般に、CATIによって行われる調査は、対面インタビューあるいは郵送調査よりも速やかである。郵送調査では、データ収集の期間は予想困難であろう。あらゆる事態において、現場作業の期間は、回答しない単位に郵送されるはずの残余数に依存する。

無回答

調査の期間と回答率は、逆の関係を持つことが多い。現場作業の間により多くの時間と資源を使えば使うほど、回答率はより高いだろう。受け取った無回答の率は常に妥協的である。よく引用される基本的命題は、3つ要請—低費用、低い無回答率、速やかな日程—のうちの2つだけが1つの調査で達成しうる、というものである。

管理方式は無回答率に影響し、郵便調査で無回答が最も高い。したがって、インタビューは今日では社会調査でより行きわたっている。インターネット調査では、無回答とフレームの過小カバレッジは大きな問題である。

他の多くの要因も無回答率に影響する (2.9節も参照)。例えば、インタビューの長さ、調査

票の複雑性、ならびに休日の季節の調査の実施、が無回答率を高める。これに加えて、調査のトピックが回答率に影響を持ちうる。例えば、健康問題に関わる調査は、他の調査より高い回答率をもたらす傾向がある(Lynn et al., 2002 と2.8参照)。

収集したデータ量と品質

収集されたデータの全体量、すなわちデータ量は、抽出単位数に1単位から収集されたデータ項目数を乗じて測定することができる。両方の要因とも調査の全体期間の長さに影響を与える。特に、紙の調査票を使うときには、収集されたデータ量は、必要とされる調整とデータ処理、そして結果として調査の全体期間の長さに直接的に影響する。

データ量は可能なデータ収集方法を限定する。一般に電話インタビューは30分を超えるべきでないとされる。面談インタビューはかなり長いことがありうる。他方で、回答者が容易にその気になる更に長い電話インタビューからの優れた経験がある。郵送調査においては、無回答率は、質問数とともに急に増加する。

インタビュー調査においては、質問者が回答者のために調査票に記入する。郵送調査においては、回答者は質問と指示を彼/彼女自身で読み、彼/彼女がそれらを理解したとおり、それらに回答する。インタビュー方式の長所は、質問者は、質問内容と回答の記入の仕方を正確に知っており、すべての質問に受け入れうる回答を得る点を保証することである。さらに、質問者は、調査票を良く知っているので、流暢にまた速やかに1つの質問から他の質問に移ることができる。郵送の調査票は一般により多くの誤差をふくむ。集中的な電話インタビューの特別な長所は、調査の進行とその品質を常に監視できることである。面接者の訓練、監視そして補足的指示を徹底することもまた、分散化したインタビューシステムよりかなり簡単である。

異なるデータ収集方法への賛成と反対

ロンジチュージナル調査は、繰り返されるクロス・セクション調査よりも変化—あるいは社会の動学—を叙述する上でより信頼できる。これに加えて、ロンジチュージナル調査は、同じ標本、すくなくとも部分的に同じデータの収集装置を使うことができるので、資源を節約できよう。実践的な通常作業はより効率的になる。いくつかの場合には短所は、回答者がパネルに属していることを自覚するようになり、その行動や回答を変えるパネル効果である。面接インタビューにおいては、現場作業とその組織はより大きな資源を必要とし、電話あるいは郵送調査の場合よりも大きな費用になる。他方で、面談インタビューは質問と回答の両方を明確にすることを可能にするので、面接者の存在は、インタビューの状況を管理可能にする。面談インタビューは最も高価なデータ収集法である。

電話インタビューにおいては、インタビューを助ける絵あるいは視角的な目盛りを使うことは、面接インタビューや郵送調査と違って、排除される。他方で、面接インタビューにおいては、面接者の影響が電話インタビューよりもより大きい。それに加えて、面接者の訪問は、彼あるいは彼女に、電話インタビューよりも大きな準備を求める。

郵送調査の1つの長所は費用が少ないことである。紙の調査票のセンシティブな質問に答えることは、電話調査あるいは面談インタビューよりも容易である。幾つもの回答選択肢（6つ以上）から選ぶ方式は、郵送調査、あるいは面接インタビューだけで可能となる視角化を必要とする。

コンピュータ支援のデータ収集は、紙の調査票と比較して多くの新しい可能性をもたらす（コンピュータ支援のインタビュー方法は表2.2にリストされている）。例えば、それは多様な質問と回答の組み合わせの使用を促進する。異なる回答者に向けた異なる質問を実施することが容易になり、各標本単位に対して、例えば、レジスターからの背景情報を添えることが可能になる。例えば職業による回答のコード化もインタビューの間に可能である。コンピュータに支援されたデータ収集はデータ入力誤差を減らし、収集した材料の完成を早める。それは、面接者あるいは回答者がインタビュー中にデータを直接コンピュータに入れ、正しさと論理的一貫性のチェックが調査票にプログラムすることができるからである。コンピュータに支援された方法は、フィンランド統計局のために実施されるほぼすべての電話および面接インタビューで採用されている。

表2.2 コンピュータ支援のインタビュー法の分類 (de Leeuw and Nicholls,1995)

具体的方法	コンピュータ支援形態
面談インタビュー法	コンピュータ支援個人インタビュー (CAPI)
電話インタビュー法	コンピュータ支援電話インタビュー (CATI)
自己管理形態	コンピュータ支援の自己インタビュー (CASI)
郵送調査	ウェブインタビューCAWI,メールによるディスク (DBM) ,電子メール調査 (EMS)
パネルあるいは日記調査	コンピュータ支援パネル調査 (CAPAR) ,電話インタビュー, 電子日記
面接者立会い	面接者立会いのコンピュータ支援インタビュー (CASIIP) ,スクリーン上の質問文:スクリーン上に見える質問 (CASI-V) およびオーディオ上の質問 (CASI-A)

注意：ウェブインタビューCAWIが表に加えられた。これは原表にはない。一般的な名称：コンピュータ支援のデータ収集), CASIC(Computer Assisted Survey Information Collection-コンピュータ支援調査情報収集), CAI(Computer Assisted Interviewing, コンピュータ支援インタビュー)

自己管理の電子的調査票が最近導入された。この道具はインターネット調査票かe-メールに添付される電子的調査票かである。個人あるいは世帯の調査における電子的データ収集の主な問題点は、カバレッジと十分な標本抽出フレームの欠如である。誰もがインターネットにアクセスを持つわけではなく、e-メールアドレスの完全なリストはない。これはビジネス調査の問題であることは少ない。

センシティブな質問に関しては、自己管理調査票は、面接よりも信頼できる回答をもたらし、また無回答を減らしうる。

面接インタビューにおいても、選択肢として、回答者自身が質問の幾らかあるいは総てに答えを記入させることがときどきある。その根拠は、面接者による介入が回答に影響することである。

しかし、影響の仕組みやその効果は通常十分には知られていない。

電子的自己管理のデータ収集道具はビジネス調査にだけ適用可能であり、そういった環境の下でも、総てのビジネス企業が十分に回答する能力を持っているかどうかという問題は、注意深く検討されるべきである。個人や世帯に関する調査に関して、インターネットの調査票は、追加的データ獲得装置として用いることができる。

ガイドライン

- フィンランド統計局の要請は、個人あるいは世帯の調査における回答率は、実際の標本集団から計算されなければならないということである。
- 通常フィンランド統計局が実施する調査における標本抽出単位の90%以上が面接インタビューで接触することができ、電話インタビューに関しては、電話番号は回答者のほぼ90%について見つけることができる。通常、最終的な回答率は、面接インタビューでは80%を越え、電話調査では80%に近づく。郵送調査の回答率は通常70%以下にとどまる。フィンランド統計局は、無回答を最小化するガイドラインを集めた (Japec et al., 2000参照)。
- フィンランド統計局は、2つの質問者組織、すなわち、実地面接者向けと、CATI面接者向けを持っている。データ収集の日程は常に、前もってそれらの機関の同意をえなければならない。

文献

de Leeuw, E. D. & Nicholls, W. L. II. 1996. Technological Innovations in Data Collection: Acceptance, Data Quality and Costs. *Sociological Research Online*. Vol. 1. No. 4.
<http://www.socresonline.org.uk/1/4/leeuw.html>

Japec, L., Ahtiainen, A., Hörngren, J., Lindén, H., Lyberg, L. & Nilson, P. 2000. Miska bortfallet (Minimise non-response; in Swedish only). Statistics Sweden, Örebro.

追加的情報

Ahola, A. 1993. *Tiedonkeruumenetelmä: konteksti: haastatteluvastaukset*. (Data collection method: context: survey responses; in Finnish only). Statistics Finland. Discussion memorandums 4/1993. Helsinki.

Biemer, P. & Lyberg, L. 2003. *Introduction to Survey Quality*. New York: John Wiley & Sons.

Canoune, H. L. & Leyhe, E. W. 1985. Human Versus Computer Interviewing. *Journal of Personality Assessment*. Vol. 49. 103–106.

Catlin, G. & Ingram, S. 1988. The Effects of CATI on Costs and Data Quality: A Comparison of CATI and Paper Methods in Centralized Interviewing. In Groves, R. M., Biemer, P. P., Lyberg, L. E., Massey, J. T., Nicholls, W. L. II & Waksberg, J. (eds.). *Telephone Survey Methodology*. New York: John Wiley & Sons. 437–450.

Couper, M. P., Hansen, S. E. & Sadosky, S. 1997. Evaluating Interviewer Use of CAPI Technology. In Lyberg, L. E., Biemer, P., Collins, M., Dippo, C., Schwarz, N. & Trewin, D. (eds.). *Survey Measurement and Process Quality*. New York: John Wiley & Sons. 267–287.

de Leeuw, E. D. 1993. Data quality in mail, telephone and face-to-face surveys. Amsterdam: TT-Publikaties.

- de Leeuw, E. D. & Collins, M. 1997. Data Collection Method and Data Quality: An Overview. In Lyberg, L. E., Biemer, P., Collins, M., Dippo, C., Schwarz, N. & Trewin, D. (eds.). *Survey Measurement and Process Quality*. New York: John Wiley & Sons. 199–221.
- Dillman, D. A. 2000. *Mail and Internet Surveys: The Total Design Method*. New York: John Wiley & Sons. (Second edition, 2006)
- Groves, R. M. 1989. *Survey Errors and Survey Costs*. New York: John Wiley & Sons.
- Kasprzyk, D., Duncan, G. J. & Kalton, G. 1989. *Panel Surveys*. New York: John Wiley & Sons.
- Lynn, P., Clarke, P., Martin, J. & Sturgis, P. 2002. The Effects of Extended Interviewer Efforts on Nonresponse Bias. In Groves, R. M., Dillman, D. A., Eltinge, J. L. & Little, R. J. A. (eds.). *Survey Nonresponse*. New York: John Wiley & Sons. 135–149.
- Martin, J. & Manners, T. 1995. Computer Assisted Personal Interviewing in Survey Research. In Lee, R. M. (ed.). *Information Technology for the Social Scientist*. London: UCL Press.
- Nicholls, W. L. II, Baker, R. P. & Martin, J. 1997. The Effect of New Data Collection Technologies on Survey Data Quality. In Lyberg, L., Biemer, P., Collins, M., Dippo, C., Schwarz, N. & Trewin, D. (eds.). *Survey Measurement and Process Quality*. New York: John Wiley & Sons. 221–249.
- Vehovar, V., Batagelj, Z., Lozar Manfreda, K. & Zaletel, M. 2002. Nonresponse in Web Surveys. In Groves, R. M., Dillman, D. A., Eltinge, J. L. & Little, R. J. A. (eds.). *Survey Nonresponse*. New York: John Wiley & Sons. 229–243.

2.8 調査票の設計とテスト

視角と目的

収集されたデータは標準化された装置に記録されるべきである。装置 (instrument) という用語をしばしば使うのは、書式(form)あるいは調査票 (questionnaire) という用語は通常は紙媒体の調査票という観念を伝えるが、大部分の調査票は今日では電子的だからである。それにもかかわらず、本節では調査票という用語を使う。調査票上では、要求される項目は、質問の形、例えば、「あなたの売り上げは何ですか」あるいは、データ項目の名前、すなわち「売り上げ」の形で示すことが出来る。

作業原則と調査票のデザインに関する次のような指示が、主として個人あるいは世帯について書かれてきたが、それらはまた大部分がビジネス調査にも適用される。事業所向けの調査票と、世帯あるいは個人向けの調査票は、そのデザインの原則で大きく異なることはない。違いは主として、調査の内容、回答者と回答の過程の性格から生じる。ビジネス調査では、質問は通常事実をとりあげ、回答者は使われている概念に通じていると考えられる。これに対して、個人あるいは世帯の調査においては、質問は非常に多様なトピックに関わり、回答者が最もありふれた概念

にすら通じていると前提することはできない。したがって、周知の事項についての回答者の知識は、個人あるいは世帯調査におけるよりもビジネス調査においてより優れていると考えることができよう。したがって、限られた数の指示を伴う調査票はビジネス調査で使えることが多い。

それらの設計に関しては、調査票あるいは装置は2つの要因、すなわち、管理方式（誰がデータを記入するか）および装置が電子的かあるいは紙版か（適用される技術）にしたがって分類される。結果としての4つのタイプは、そのデザインに関して異なる要求を提出する（表2.1参照）。

この脈絡で、最も重要な区別は、自己管理調査票と面接者管理の調査票の間のそれである。

原則

自己管理の調査票は回答者を考慮しながら設計され、面接者管理の調査票は面接者と回答者の両方を考慮して設計される。いかなる質問を設計する際にも中心的原則は、回答者が

- 彼あるいは彼女が質問されていることを明確に理解するべきであり、
- 質問に対する答えを知っていることを期待でき、
- 回答をどのようにするかを理解している、ことである。

面接者がもたらす変異を最小にするために、すべての面接者が同じ仕方で質問を示すことを保証するよう標準化されたインタビュー方法が強く推薦される。したがって、面接者の行動と質問の提示と調べをガイドする提示基準が面接において使用される。長く複雑なインタビューによって、面接者は通常は調査の前に訓練されるべきであり、これは調査の全体期間を大きく延長するかもしれない。

研究者は個人的には常に調査票の機能性に責任を持つ。例えば、フィンランド統計局の調査研究ユニット（Survey Research Uni）と調査実験室（Survey Laboratory）からの調査票についての助言と設計が利用可能である。他の単位もまたデータ収集装置を設計する際の経験を持っている。

研究倫理が要求することは、回答者は調査票の最初あるいは表紙の文字で、何故調査が行われるか、誰が資金を出しているか、標本のデータ源を知らされるべきである。調査票のはじまりはまた、正しい人が調査票に答えていることをチェックする装置を含むべきである。フィンランド統計局のすべての調査では、やがて行われる接触と調査の内容について各回答者に伝える事前の手紙が送られるべきである。

フィンランド統計局の1つの作業原則は、回答者の負担は可能な限り低く保たれるべきというものである。フィンランドの統計法は、以下のことを要求している（1.1.1節参照）。

統計目的のためにデータが収集されるときに、まず最初に利用されるべき出所は、一般政府の仕事を進める際に蓄積されたデータ、雇用者、自営業者、会社や基金の正規の活動の結果として作成されたデータであるべきである（統計法、第4節）。

統計を生産する機関は、回答者は統計の生産に必要なデータだけを提供するように求められるよう配慮する（統計法、第5節）。

そのデータは、経済的で、回答者に最小限の不便と費用をもたらす仕方で収集されるべきである（統計法、第4節）。

ガイドライン

調査票のデザイン

- 郵便調査とインタビュー調査の両方での質問のデザインについては多数の書物が書かれてきた。幾つかは本節末の文献にリストされている。これらのは調査票を設計する総ての者にとって不可欠な読み物を提供している。
- 概念、分類および質問の標準的定義が、フィンランド統計局の調査に適用されている。それらは、「統計的インタビュー」と「インタビュー調査票に向けた設計の基準」に説明されている。
- 通常は、事前にテストされ、良いことがわかった質問あるいは質問装置を使うよう試みるのがより良い。調査の比較可能性の前提は同じ質問の使用である。
- 電子のおよび紙の調査票の設計は大きく異なる。例えば、電子的調査票がプログラム化されなければならない、したがって、プログラミングに責任を持つ人の名前と、その課題のためのタイムリミットを設定しなければならない。
- もし調査が幾多の言語を使って行われるなら、十分な時間と資源が調査票の翻訳、作業指令、および他の材料に当てられるべきである。電子的装置の翻訳は、フィンランド語版が出来上がり、テストされた後にだけ行われるべきである。
- 指定された回答者は自己管理の調査票で明白に指示されるべきである。しかし、ビジネス調査では、要求されている情報はビジネスにかかわる。回答する企業の確認は不可欠である。しかし、情報を提供する個人の確認は不可欠ではない。回答についての明確な指示と返送アドレスがまた表紙ページあるいは表紙の文字で提供されなければならない。
- 紙の調査票は記入を終えるのが容易で、便利でなければならない。第一に、データ入力の見地から言って、余りに読みにくいものであってはならない。第二に、それは、明確で、スムーズであり、論理的でなければならない。データ入力、回答のための場所が1つの欄、例えば右端に置かれるなら、やりやすい。
- 調査名と連絡先の人物の名前と電話番号は、すべての表紙の文字に明白に示されて、質問が直接ただちに担当者に向けられるようにするべきである。積極的調査に向けた表紙の文字のコピーもまた電話交換に送られるべきである。

調査票のテスト

- 調査票をテストし評価するために広範な方法を使うことが出来る。デザイナーは、例えば彼あるいは彼女の同僚の助けを借りてテストインタビューや調査票を評価することができよう。電子調査票のテストは、技術的見地と内容的見地という2つの見地に注意を要するので、特に必要で、時間消耗的な作業である。
- 試験調査を行うことは、可能であれば何時でも、大いに行うに値する。試験調査はデータ収集の過程のすべての段階をテストできるという長所をもつ。
- 新しいかあるいは再度設計された調査票をテストする際には、フィンランド統計局の

調査実験室のサービスに相談し、利用することが勧告される。

- 焦点の集団は、統計調査の主題事項についての事前の情報を提供し、回答者が使う概念や表現の全体像をつくる助けとなる。焦点の集団の論議は記録されるべきである。
- 認識のための1対1のインタビューは、調査の質問を受けている過程で回答者の考えを調べるために使われる。この狙いは、回答者が質問を理解し、情報を思い出すことができ、受け入れ可能な回答を提供できるかを見出すことにある。
- 専門のパネルは人間の認識過程についての知識に照らして、調査票の草案を討論することに従事すべきである。調査票のコード化マニュアルもまた使用できる。
- 2つ以上の選択できる調査票版を試験的に比較する試験パネルテストを、用語、質問の順序およびデータ収集法の影響を評価するために使うことができる。

文献

フィンランド国内法規

Statistics Act (280/2004).

フィンランド統計局のガイドライン

Statistics Finland. 1980. *Tilastohaastattelijan opas*. (Guide to statistical interviewers; in Finnish only). Statistics Finland. Handbooks 7. Helsinki.

Statistics Finland. 1999. *Tilastohaastattelijan opas*. (Guide to statistical interviewers; in Finnish only). Statistics Finland. Internal memorandum. Helsinki.

Statistics Finland. 2001. *Haastattelulomakkeiden suunnittelustandardit*. (Designing standards for interview questionnaires; in Finnish only). Statistics Finland. Internal memorandum. Helsinki.

追加的情報

Ahola, A. 1992. *Haastattelukysymysten ymmärtämisestä*. (Understanding survey questions; in Finnish only). Statistics Finland. Discussion memorandums 1/92. Helsinki.

Biemer, P. P., Groves, R. M., Lyberg, L. E., Mathiowetz, N. A. & Sudman, S. (eds.). 1991. *Measurement errors in surveys*. New York: John Wiley & Sons.

Biemer, P., Lyberg L. 2003. *An Introduction to Survey Quality*. New York: John Wiley & Sons.

Converse, J. M. & Presser, S. 1986. *Survey Questions. Handcrafting the Standardized Questionnaire. Quantitative Applications in the Social Sciences*. No. 63. Newbury Park: SAGE Publications Ltd.

Fowler, F. J. Jr. 1993. *Survey Research Methods*. 2nd ed. Newbury Park: SAGE Publications Ltd.

Groves, R. M., Biemer, P. P., Lyberg, L. E., Massey, J. T., Nicholls, W. L. II & Waksberg, J. (eds.). 1988. *Telephone Survey Methodology*. New York: John Wiley & Sons.

Holstein, J. A. & Gubrium, J. F. 1995. *The Active Interview*. Thousand Oaks, CA: SAGE Publications Ltd.

Lehto, A.-M. 1998. Laatusurveyitutkimukseen. (Quality for survey research). In Paananen, S., Juntto, A. & Sauli, H. (eds.). *Faktajuttu. Tilastollisen sosiaalitutkimuksen käytännöt*. (Factual facts. Practices of statistical social research; in Finnish only). Tampere: Vastapaino. 207–233.

- Lyberg, L., Biemer, P., Collins, M., de Leeuw, E. D., Dippo, C., Schwarz, N. & Trewin, D. (eds.) 1997. *Survey Measurement and Process Quality*. New York: John Wiley & Sons.
- Presser, S., Rothgeb, J. M., Couper, M. P., Lessler, J. T., Martin, E., Martin, J., & Singer, E. (eds.) 2004. *Methods for testing and evaluating survey questionnaires*. New Jersey: John Wiley & Sons.
- Sudman, S., Bradburn, N. M. & Schwarz, N. 1996. Thinking about answers. *The application of cognitive processes to survey methodology*. San Francisco: Jossey-Bass Publishers.
- Tanur, J. M. (ed.). 1992. *Questions about questions. Inquiries into the cognitive bases of surveys*. New York: Russel Sage Foundation.

2.9 調査への回答

視角と目的

統計調査において 妥当な観察の数は、データ収集の様々の段階で、異なる理由によって減らされることがしばしばある。直接的なデータ収集におけるデータの損失は、必要とされるデータが選択された単位からは獲得できないという事実から生じる。これはまた、調査の未回答と呼ばれる。代表的には、任意的な郵便およびインタビュー調査において起こる問題であるが、データ収集の実施が最大限のものである人口センサスや企業センサスといった、回答者が回答する法的義務を持っている調査においても避けることは出来ない。データが、行政レジスターあるいは記録から間接的な方法で獲得できる何らかの調査においては、データの欠損は小さな問題であるが、一方では、多様な問題がフレーム誤差あるいは測定誤差が生じ、これらは大きな問題でありうる。

調査の無回答は一般的には2つの主な区分、項目無回答と単位無回答に区分される(Groves and Couper 1998)。項目無回答は、その他では妥当な調査票のある項目あるいは部分が不完全、すなわちデータのベクトルが不完全なときに、生じる。それは、回答する能力の欠如、知識の欠如、回答する意思がないこと、ある回答がデータエディティングの規則をパスする点での失敗、あるいは論理的チェックの欠陥、といった様々な理由から生じる。単位無回答の場合には、データは総ての変数に関して欠如している。

無回答は調査結果に影響する。回答者と無回答者の集団が背景の変数と調査変数の点で類似しているとしても、観察数の減少によって、標本分散の増加という影響を受ける。しかし、ほとんどの場合に、無回答者と回答者は類似ではなく、無回答による追加的な誤差源をもたらす。すると、無回答は調査結果に偏りを生み出しうるので、特に有害である。全数のセンサスもまた無回答によって影響を受けることが指摘されるべきである。

原則

最も重要な原則は、可能な最良の品質の、偏りのない統計を生産することである。回答率は高く標本分散は小さくあるべきである。これらの目標は、データ収集の注意深い設計、調査票をふくむ測定装置の注意深いテスト、および現場作業とその後のデータ処理の優れた実行と監視を通じて達成できる。資金、締め切り、利用可能な人的資源といった無数の制約が考慮されなければならない。現場活動の組織には特別な考慮が払われなければならない。というのは、これは回答者とデータ収集組織との間での主な接点をなすからである。

単位無回答の理由は、通常は主として3つのに区分される。すなわち、連絡をとれなかったこと、拒否、その他の理由、である。一般原則は全体的な回答率を高く保つことであり、これら3つの区分すべてが小さく保たれることを意味する。しかし、データ収集組織は、連絡をとれないこととその他の理由は、無回答をかなり減らすことが出来る理由であることを考慮すべきである。任意的な調査における拒否は、調査のトピックに影響されることが多い。選ばれた個人、世帯あるいはビジネスの意思決定に影響を及ぼすためにできることは少ないが、面接者のスキルは、もちろん、拒否を転換するテクニックや人々の調査への参加モチベーションを高める方法で面接者を訓練することによって改善できる。法律で定められた調査においては、回答の歪み、したがって偏った調査結果という危険が特に明白なので、データ収集組織は拒否率を特に低く保つよう努力をする。現場活動の後に、データをチェックし、エディットし、調査している変数と背景の変数の両方の分布を母集団と比較することが重要である。これは、結果の信頼性の評価、および（いくつもの）誤った結果を調整することを促し、偏りを減らす助けとなる。データの利用者は、データ収集、結果データの特徴、データと無回答の詳細分析を、その結果への予想される影響とともに、説明している注意深い品質報告を必要とする。偶然的な偏りをふくめて、無回答誤差は常に、適切な統計方法によって調整されるべきである(2.11節参照)。

ガイドライン

- フレーム母集団と潜在的標本フレームは注意深く検査され、目標母集団と比較されるべきである。
- 現場活動と他の調査方法は適切であり、母集団を正しく目標とするべきである。同じように、タイミング、前もっての手紙と調査小冊子、指示、データ収集方法と道具、最終的技術的工夫その他の道具は、注意深く設計され、回答者負担の減少を狙うべきである。
- データ収集方法と道具は、実際の現場活動に先立って注意深くテストされるべきである。
- 標本抽出と推定方法は可能な限り偏りの無い結果を獲得できるものであるべきである。
- 無接触単位をたどる継続的な努力が払われるべきであり、拒否の転換テクニックが適用されるべきである。拒否は法律上の調査において絶対的に最小に保たれるべきである。
- 例えばビジネス調査のように、歪んだ母集団が調査されているときには、最も影響のある単位は説得して回答してもらうべきである。
- 回答集団と無回答集団の特徴が比較されるべきであり、無回答の分析が行われるべきである。

最終データ集合は、ウェイト調整あるいは補定といった技法によって可能な限りフレーム母集団に対応するよう調整されるべきである（2.10と2.11節参照）。

- 無回答は測定され、文書化され、国際的に受け入れられる基準を使って、相対的見地とともに絶対的および構造的に報告されるべきである。データの利用者は、その起こりうる影響、歪みや他の誤差源などについて知らされるべきである。継続調査においては分類別の無回答率を監視し、世論あるいは現場作業における変化を考慮に入れることが出来るようにすることが重要である。

文献

Groves, R. M. & Couper, M. P. 1998. *Nonresponse in Household Interview Surveys*. New York: John Wiley & Sons.

Lessler, J. & Kalsbeek, W. (eds.). 1992. *Nonsampling Error in Surveys*. New York: John Wiley & Sons.

Särndal, C.-E., Swensson, B. & Wretman, J. 1992. *Model Assisted Survey Sampling*. New York: Springer-Verlag.

Särndal, C.-E. & Lundström, S. 2005. *Estimation in Surveys with Nonresponse*. Chichester: John Wiley & Sons.

追加的情報

Groves, R. M., Dillman, D. A., Eltinge, J. L. & Little, R. J. A. (eds.). 2002. *Survey Nonresponse*. New York: John Wiley & Sons.

Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., and Tourangeau, R. 2004. *Survey Methodology*. New York: John Wiley & Sons.

Japac, L., Ahtiainen, A., Hörngren, J., Lindén, H., Lyberg, L. & Nilsson, P. 1997. *Minska bortfallet*. (Decrease nonresponse; in Swedish only). Örebro: Statistics Sweden.

Laaksonen, S. 1988. *Katovirheen korjaus kotitalousaineistossa*. (Correcting for Nonresponse in Household Data; English summary). Statistics Finland. Studies 147. Helsinki.

Laaksonen, S. (ed.). 1996. *International Perspectives on Nonresponse*. Statistics Finland. Research Reports No. 219. Helsinki.

2.10 統計的エディティングと補定

視角と目的

統計的エディティングと補定は、誤差を持つデータを探り当て、訂正するために使われている。これらの局面は、通常データ収集の後に進められるが、電子的統計データ収集では、組み込まれたチェックを通じて、そして電子的データ収集での定常過程として実施することができる。エディティングと補定に使われる方法と適用はなお発展中である。現在、文献と実践の両方で種々の用語が使われている。意義があり、効率的な方法の発展を続けるためには、用語に共通の理解をもたらすように用語の標準化が明らかに必要である。ここで論じられた2つの概念は、以下のよ

うに定義される (Eurostat,2005)。

- **統計的エディティング**とは、潜在的に誤りをもつデータ記録に対して、欠損している、不適切な、あるいは一貫していない入力あるいは点を確認するチェックの適用である。
- **補定**とは、エディティングの過程で確認された欠損している、不適切なあるいは一貫しない回答の問題を解決するために使われる過程である。補定は、もっともらしい、内的に整合的な記録を創り出すことを確かにするために、エディティングの過程にある記録で何らかの回答あるいは欠測値を置き換えることで実行される。

国連統計委員会とヨーロッパ経済委員会(UN/ECE,2000a)はまた、エディティングと補定についての概念、原則、技術、および方法の包括的文書を提供している。このようにデータのエディティングは、補定と密接に関連している。データのエディティング法は、主として誤りのある、一貫しないデータ値を確認するために、補定はそれらと欠測値を正すために使われる。

より正確に言うと、**統計的エディティング**は、統計値を個別の値と、異なる変数の値間相互の比較可能性の両方に関してチェックする活動のことである。この実施は、データの最下位レベルから、すなわちマイクロ・レベル (マイクロ・エディティング) から出発し、より合計されたレベル (マクロ・エディティング) に進むべきものである。最終結果におけるレベル間の相互比較可能性が狙われるべきである。すべてのデータは調査結果と分布の大きな歪みを避けるために注意深く検査されるべきである。

補定は、変数の欠測値あるいは誤った (例えば、エディットの失敗) 値を、真の、但し未知の値に関して可能な限り正しくなければならぬ補定値で置き換えることを意味する。補定法は、データ集合のタイプ、その範囲およびデータの欠損のタイプによって大きく異なる。

実践においては、エディティングと補定は連続する局面で行われる。データのエディティングは、不適切な値あるいは論理的に一貫しない値が存在するかどうかを検査するために必要とされる。補定は次に欠測値を代置するために必要とされる。統計的エディティングは、データ収集の計画からはじまって、データファイルの形成、データ処理と分析に至る各局面で必要である。データの一層のエディティングは、データ収集が終わった数年後においてすら、例えば、ロングitudinalデータが構成されるときに、必要となるかも知れない。

原則

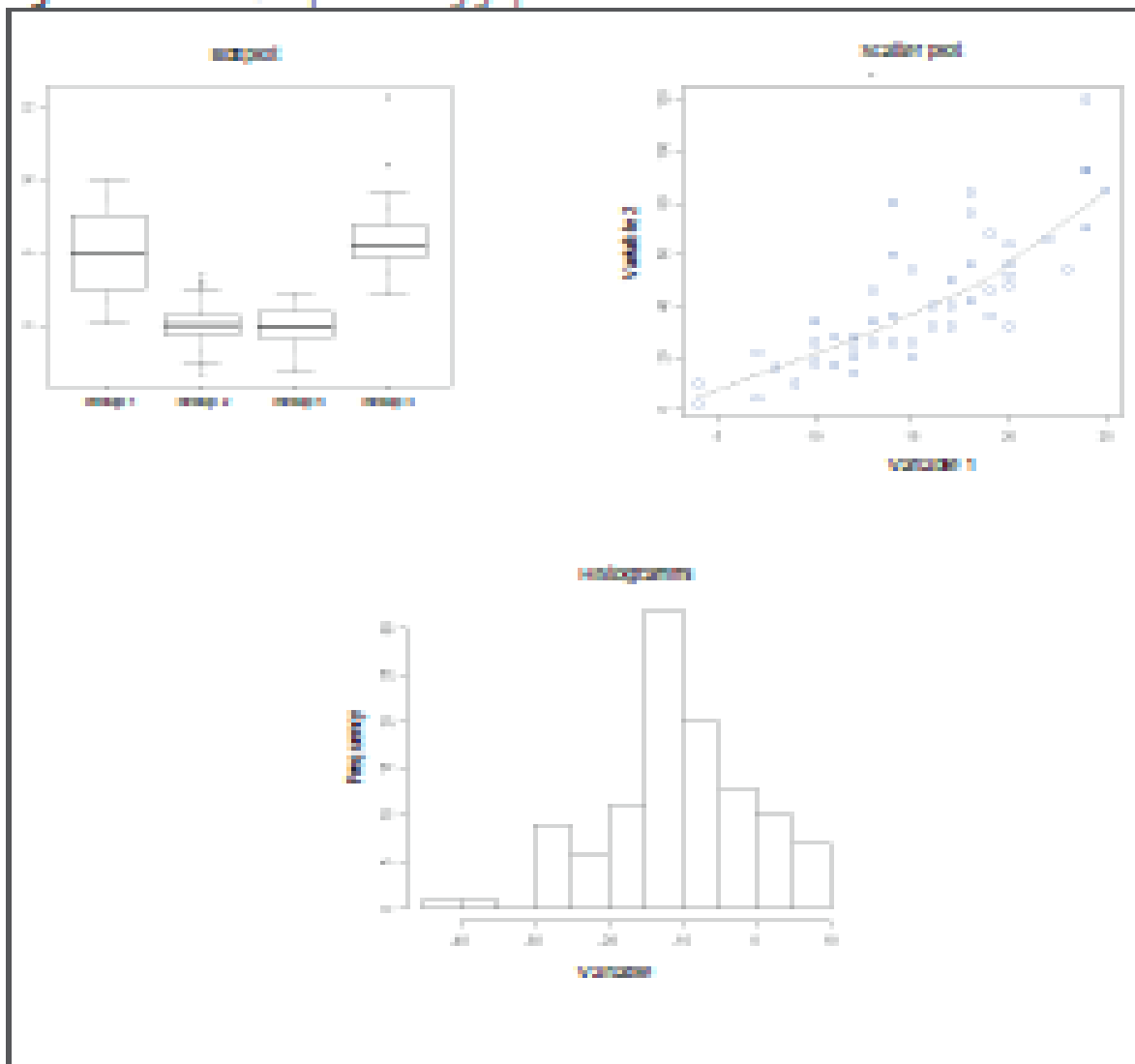
エディティングと補定の目的は三重である。すなわち、データをきれいにし、データの品質についての情報を生産し、将来における調査過程の品質を改善するための基礎を提供することである。誤差と欠陥の検出の全体的影響と重要性は、エディティングと補定の後に分析されるべきである。大きな資源が、大きな問題を防ぎ、分析することに投入されるべきであり、これに対して小さな欠陥は自動エディティングと補定による解決で処理されるべきである。

統計的エディティング

エディティング規則の設計は、統計的データ・エディティングにおいて要めの役割を果たす。エディティング規則は、多変量の分布への影響を無くし、探索された誤差の数を最大にするように設計されるべきではない。全く反対に、エディットの集合を具体化するとき、多くの要求が考慮に入れなければならない(UN/ECE, 1997), 例えば

- エディットの集合の完全性：誤差を確認するためのすべての可能なルールが定義されるべきである。
- エディットの集合の正しさ：総てのエディットは、領域について我々が持つ知識に対応しているべきである。
- エディットは宣言され、明白で読みやすい形で文書化されるべきである。すなわち、エディットは、追跡でき、それらの影響が評価されるように文書化されるべきである。

Figure 2.5 Three examples of using graphical visualization of data.



データは、情報として示される前にエディットされるべきである。この活動は、提供される情報は、正確で、完全で、一貫していることを保証することを狙っている。エディティングの品質効果の測定は、生データと最終データの間の比較に基づく。エディティング率は、実施された管理作業の相対的程度を示す。この割合は、拒絶した誤りのある観察数を、総観察数で割った比率として測定される。

エディティング割合ないし比率は データの処理の効率を評価するために使用できる。小さなエディティング比率は、疑問のある値は、すべての観察数の合計数のわずかの部分であることを示す(UN/ECE, 2000b)。高いエディティング比率と結びつくと、過程のレビューは、拒絶した値の検討に費やされた資源は、正当化されないと結論するかも知れない。しかし、指標の解釈は注意をもって検討されなければならない。

エディティングの段階は、調査の推定値のバイアスが、加重エディティングなしに資源の効率的な使用をしながら、最小限であることを保証すべきである。エディティングには、4つの共通に規定されている段階がある。すなわち、最初のエディット、伝統的エディット、選択的エディット、マクロ・エディットであり、これらは自然的な連鎖を持つ。エディティングの各段階の後に、探り当てたデータを訂正するために通常は補定が必要とされる。エディティングの段階は、異なるタイプの誤りを探索する強さと能力のために、重複することがある(Scarrott, 2006)。Fellegi と Holtの方法 (1976)が、統計的データのエディティングで使用されることが多い。それは、多くのデータ・エディティングソフトウェアに導入されている(例えば Banff)。

グラフによるエディティング

グラフによるエディティングは、分析的、非地理的手段によっては探索するのに多くの時間と努力を要するデータの異常、パターン、あるいは包含する関係を確認する、人間の視角的な知覚能力の使用を意味する。地理的エディティングは、推定値での特定のデータの値の影響を常には示さない、より伝統的エディティング法を支援する(UN/ECE, 1997)。地理的な手段によるデータの予備的活用は、エディティングと補定の過程のための境界を設定する際に大きな助けとなりうる。地理的エディティングの使用、例えば、箱型図 (boxplot) , 点の散布、回帰線および比率は、純粋に分析的手段によっては見つけるのが困難なデータのはずれ値、傾向、あるいは他の特性を確認するのに使うことができる。データの地理的外挿の通常3つのタイプを図2.5に示している。

- 箱型図は、利用者がカテゴリカルな集団によるデータの分布を探求することを可能にする。図2.5は、各集団における値の範囲を示し、外れ値を確認する。
- 散布図 (scatter plot) は2つの変数間のつながりを明らかにする。図2.5の回帰線から離れている値は、外れ値を探り当てるために注意深く検討されるべきである。
- ヒストグラムはデータにおける分布の形、ひろがり外れ値の存在を図的に示す。図2.5は、分布の尖度と外れ値を視角化する。

補定

補定の主な役割は、データ集合における欠測値あるいは誤った値を、知られていて受け入れ可能な値で代置することである。補定が行われるのは、代表的には、データシステムは完全なデータ集合とともに機能するように設計されているからである(Chambers, 2002)。Kalton and Kasprzyk (1986)は異なる種類の補定方法をふりかえったが、それは3つの主なグループに区分される(Laaksonen, 2000)。すなわち、

- 論理的補定は、統計的エディティングとの関連で起こる。補定値は、専門家の助けによるか、強い補助変数あるいは初期値によって到達する可能な最善の推定に基づいて作り出される。
- モデルドナー補定においては、補定値は、データに合致したモデルによって獲得される。
- 現実のドナーモデル補定においては、補定値は、観測値から借用される。したがって、値は現実の状況の下で可能な値であることは確かである。

この後者2つの集団のうち、欠測値の置き換えには、ドナーと呼ばれる単位から借用される。集団は、ドナーの性質によって決定される。現実の状況下では、幾多の方法を比較することを勧告することができる。これは、例えば現実のドナー補定法によっては現実の値を発見できない状況に、モデルドナー補定法がうまく適合するという理由で、ただ1つの値にかかわるところでも言えることである。

補定法は、さらに決定論的方法と確率論的方法に区分することが出来る。後者は確率的要因をふくんでいる。原則として、上記の各グループは、確率的と決定論的モデルの両方を含むことが出来る。それに加えて、補定法は、単一補定法と多元補定法に区分することが出来る。単一補定においては、欠測値は単一の補定値で置き換えられ、多元補定においては(Rubin, 1987)、2つ以上の値が使用される。いまのところ、多元補定法は、フィンランド統計局の統計生産では利用されていない。

補定は、結果に追加的不正確性をもたらすので、これは推定されるべきである。多元補定はときとして追加的分散を評価するために使われる((Rubin, 1987; Rubin 1996; Schafer, 1997)。他の方法もまた、補定法から生じる分散を調べるために開発されている(例えば、Rao and Shao, 1992; Särndal, 1996; Lundström and Särndal, 2001, Lee et al., 2002参照)。

比率の形での指標は、補定過程を監視するために計算できる。エディティングにおいて、指標の解釈に注意を払われるべきである。補定割合と補定比率は単一の変数である。補定割合は、補定された記録の、記録総数に占める割合として測定される。補定比率は、最終推定値に対する補定値への寄与の尺度であり、すべての最終値の総量で割った補定値の量として計算される(Eurostat, 2005)。

ガイドライン

- 項目無回答によって生じる欠測値は、すべての統計データにおいて必要なものとして補定されるべきである。

- 単位無回答によって生じる欠測値は、世帯と個人データをウェイト調整によって正されるものである（2.11節参照）。他の統計データにおいては、単位無回答はまた補定によって訂正される。
- 識別項の正しさは、もし調査データがウェイト作成、推定目的のために、あるいは他のデータと結合されるために使われるなら（2.9 と 2.10節参照）、不可欠である。クロスセクションの結びつきに適切な識別項は、データ集合のロンジチュージナルな結びつきには十分ではないので、特定のロンジチュージナル識別システムもまた設計されるべきである。
- エディティング過程は、理論的および実際の説明と一緒に文書化されなければならない。それに加えて、最初のオリジナルなデータと自動的エディティングに使われたシンタックスが文書化され保存されるべきである。
- 品質指標が作られ、検査されるべきである。しかし、その解釈は注意をもって研究されるべきである。エディティングのための主な指標は、エディティング率とエディティング比率であり、補定のためには補定割合と補定比率である（Eurostat, 2005; UN/ECE, 2000b）。

統計的エディティング

- 数値の範囲は、これがデータにおいてすべての受け入れられない特徴を示しているかどうかを見るためにチェックされるべきである。これは、すべてのデータ集合および適切な部分集団について行われるべきである。すべてのコンピュータに支援されたデータ収集テクニック（CAPIあるいはCATI、あるいは自己管理の電子的調査票といった、2.7節参照）を使うなら、前もってプログラムされたチェックシステムは可能なきにはいつでも実施されるべきである。
- 外れ値の探索が、値の範囲チェックの後に直接的に適用される。この活動は、データ集合における変数の分布を詳細に検査することである。すべての識別された極端な値あるいは逸脱した値について妥当性が注意深く検討されるべきである（補定法参照）。
- 3つの概念、すなわち、欠測値、ゼロ値、論理的にありえない情報、が区分されるべきである。
- データの異なる項目は、個々の項目が正しいものであるべきことと同様に、両立可能であるべきである。データの下位領域を検査するために、外部データをも活用して、クロスチェックを使うことができる。多変量分析に基づくチェックは、現実の調査デザインに適切なモデルを使用して行われるべきである。

補定

- 適切な補定モデルは幾つかの選択肢を試み、テストすることによって、構築されるべきである。
- 補定値は、補定が最初の値と区別できるように、指標変数によって伝達されるべきである。
- 補定の結果は、図と表の両方を使って検査されるべきである。もしモデル・ドナー法が使われるなら、補定値の分布は、ありうる不適切なあるいははずれ値について検討されるべきである。現実的ドナー法が使われるときには、ドナーが十分な数で使われたかを考慮すべきである。

- もし可能なら，補定から生じる追加的分散の推定値を計算するべきである。

文献

- Chambers, R. 2002. *Evaluation Criteria for Statistical Editing and Imputation*. National Statistics Methodology Series 28.
- Eurostat. 2005. *European statistics code of practice, Glossary of quality terms*.
- Fellegi, I.P. & Holt, D. 1976. *A systematic approach to automatic edit and imputation*. Journal of the American Statistical Association 71 (1976) 17–35.
- Kalton, G. & Kasprzyk, D. 1986. *The Treatment of Missing Survey Data*. *Survey Methodology*. Vol.12. No. 1. 1–16.
- Laaksonen, S. 2000. *Regression-based Nearest Neighbour Hot Decking*. *Computational Statistics*. Vol.15. No.1. 65–71.
- Lee, H., Rancourt, E. & Särndal, C.-E. 2002. *Variance Estimation from Survey Data under Single Imputation*. In Groves, R. M., Dillman, D. A., Eltinge, J. L. & Little, R. J. A. (eds.). *Survey Nonresponse*. New York: John Wiley & Sons. 315–329.
- Little, R. & Rubin, D. 1987. *Statistical Analysis with Missing Data*. New York: John Wiley & Sons.
- Rao, J. N. K. & Shao, J. 1992. *Jackknife Variance Estimation With Survey Data Under Hot Deck Imputation*. *Biometrika*. Vol. 79. No. 4. 811–822.
- Rubin, D. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sons.
- Rubin, D. 1996. *Multiple Imputation After 18+ Years*. *Journal of the American Statistical Association*. Vol. 91. No. 434. 473–520.
- Scarrott, C. 2006. *Feasibility study: A review of selective editing*. Mathematics and statistics department, University of Canterbury.
- Schafer, J. L. 1997. *Analysis of Incomplete Multivariate Data*. London: Chapman & Hall.
- Särndal, C.-E. 1996. *For a Better Understanding of Imputation*. In Laaksonen, S. (ed.). *International Perspectives on Non-response*. Statistics Finland. Research Reports 219. Helsinki. 7–22.
- Särndal, C.-E. & Lundström, S. 2005. *Estimation in Surveys with Nonresponse*. Chichester: John Wiley & Sons.
- United Nations statistical commission and economic commission for Europe (UN/ECE), 1997. *Statistical data editing: Methods and techniques Vol No. 2*. United Nations, New York and Geneva.
- United Nations Statistical Commission and Economic Commission for Europe (UN/ECE), 2000a. *Glossary of terms on statistical data editing*.
- United Nations Statistical Commission and Economic Commission for Europe (UN/ECE), 2000b. *Evaluating efficiency of statistical data editing: General framework*.

追加情報

- Laaksonen, S. 1991. *Adjustment for Non-response in Two-year Panel Data: Applications to Problems of Household Income Distribution*. *The Statistician*. Vol. 40. No. 2. 153–168.
- Laaksonen, S. 1996. *Statistical Methodology and the Principle of Subsidiarity: A Challenging Optimization*

Task for Eurostat and European Statistical Institutes. Proceedings of the 1st International Conference on Methodological Issues in Official Statistics. Stockholm: Statistics Sweden.

Longford, N.T. 2005. *Missing Data and Small-Area Estimation: Modern Analytical Equipment for the Survey Statistician*. New York: Springer-Verlag.

Statistics Canada. 2006. *Banff. Functional Description of the Banff System for Edit and Imputation. Generalised System Method Section & Business Survey Methods Division.*

2.11 ウェイト作成と無回答の調整

視角と目的

ほとんどの標本調査において、ウェイト変数は、分析のためにデータ集合に含まれるべきである。ウェイトづけの主たる目的は、選ばれた標本抽出と推定手続きによって可能なように、正確なパラメータの推定値を獲得することである。標本抽出のウェイトは、使用中の標本設計についての情報、ある」場合には、正確性を改善し/あるいは無回答を調整するために使われた補助的情報、をふくむ(以下の「原則」の小節参照)。

ウェイトの基本的形態は、選択された要素の算入確率の逆として定義される、すなわち、標本と母集団の大きさとともに標本設計の情報をふくむ設計ウェイトである。しかし、設計ウェイトは、標本抽出とデータ収集の後に、標本フレームにおける誤差、無回答および測定誤差によって起こされた歪度の可能な観察値について、調整される必要が多い。さらに、効率の改善を狙った推定手続きは、最初のウェイトのさらなる修正をふくみだろ。調整あるいは修正されたウェイトの誘導はウェイト調整 (re-weighting) と呼ばれる。

原則

設計ウェイト

設計ウェイトは、標本設計から直接的にひきだされる。ほとんどの標本設計について、設計ウェイトは、包含確率の逆である。このようにして、標本抽出手続きで使われる包含確率はウェイト変数の構築のために、ひきだされるべきである。

典型的には、標本抽出ウェイトは、拡大タイプである。拡大ウェイトは、母集団の総計の推定に必要であり、与えられた標本の要素についてのウェイト変数の値は、その標本要素によって「示される」母集団要素数であることを意味すると解釈できる。これは、標本抽出ウェイトの形式による。

原則として、調査データは標本ウェイトなしに推定されるべきである。ウェイトは、自己ウェイトづけ設計 (例えば、単純ランダム抽出あるいはランダムな順にある母集団に適応される系統抽出) であるなら無視できるし、無回答は、調整される必要のない無視できるものと判断できる。

そのときですら、推定されるべきパラメーターは、母集団総計ではなく、例えば平均、相関係数、あるいは2つの平均値比率であるべきである。それらの条件の1つですら満たされなかったなら、標本ウェイトは適切な推定のために使われるべきである。

ときとして、再設計(re-scaled)されるウェイトは、拡大ウェイトの代わりに使われる。再設計されるウェイトは、標本レベルまで合計されるウェイトおよびウェイトの平均が1に等しいように作成される。これらのウェイトは、分析ウェイトである。分析ウェイトは、社会その他の研究におけるある伝統、標本レベルでの結果の推定の必要だけ、あるいは、利用可能な設計基礎の分析の特殊な版はないという事実によって支えられている。標本レベルへの拡大ウェイトの格下げは、普通は、使用している設計の標本ウェイトを、単純任意抽出からの対応するウェイトに分割することで遂行される。ある場合には、それらの修正されたウェイトはまた、標準化された分布あるいは統計のために使われる。

ウェイト調整

ウェイト調整とは、最初の設計ウェイトを、ウェイトづけ手続きに補助的情報を導入することによって調整する狙いをもった広く使用されている方法のことである。ウェイト調整で使われる補助的データは通常、母集団あるいは標本からひきだされる。最も単純な方法は、事後階層化と比率推計である。例えば、事後階層化は、既知の母集団総計（例えば、ジェンダー別、年齢別、地域別の母集団の頻度）を考慮して、標本単位をウェイト調整することであり、一方で、比率推定は、連続変数の母集団総計（例えば、特定産業の売上高）の利用を意味する。比率推定の正しい利用は、調査変数と補助的情報の間の線形の正の関係が存在することを求める（Särndal et al., 1992）。

より複雑なウェイト調整手続きは、線形あるいはロジットモデルといった明白なあるいは暗黙の統計的モデルの使用に基づいている。次に、設計ウェイトは、補助的情報を利用するモデルで調整され、方法論は、モデルに支援される推定の範囲内におさまる。最もひんばんに使われるウェイト調整法の1つは、ウェイトのカリブレーション(calibration)である（Deville and Särndal, 1992）。カリブレーションにおいては、当初の設定ウェイトは、使用中の補助的情報のための正しい母集団分布を提供する条件をもった統計的モデルによって、可能な限りわずかの量だけ修正される。こうして、ウェイトのカリブレーションは、基本的標本設計の特性を維持するよう努め、同時に母集団分布と一貫する推定値を提供する。

殆どの場合に、ウェイト調整付けは、抽出ウェイトにその他の変異を持ち込む。したがって、標本誤差と他の品質指標へのその影響は評価されるべきである。ウェイトの変異性は、PPS標本設計あるいは重い階層化が標本選択に適用されるなら、大きく増加する。そこで、ウェイト調整によって生じる新たな変異を限定することが必要となる。

無回答の調整

モデルに支援された推定に基礎をおくウェイト調整は普通、単位無回答が生み出す可能性のあ

るバイアスについての調整で使用される。無回答は大きく、これを無視することは推計における過小カバレッジをもたらす。さらに、無回答は完全にランダムであったことはほとんどなく、われわれは、それと補助的情報あるいは調査変数の間につながりがあると仮定できることが多い。これらの仮定が正しいなら、オリジナルの設計ウェイトは、バイアスのある結果をもたらし、正しくない推論を導くかもしれない。

単純な無回答調整は、オリジナルの標本の大きさを回答数で置き換えること、すなわち、対応してウェイトを拡大することによって得られる。しかし、ウェイト調整を標本フレームで指定されている変数と調査変数を関連付ける何らかの関連情報に基づかせることがより推薦できる。関係性が存在し、回答確率によって、標本を同質的な副次母集団に分割できるなら、これは、無回答バイアスを減少させるだろう。したがって、ウェイトの構築と無回答調整のための強力な補助的データの探索、あるいは使用は、優れた統計実践の一部である。ときとして、それは、標本抽出ウェイトを既知の母集団分布に拡大することも有効であろう (Rosenbaum, 1987; Ekholm and Laaksonen, 1991; Särndal and Lundström, 2005)。

ウェイトのカリブレーション

無回答の調整は、階層後に類似する手続きで遂行されることが多い。ここでは、補助的情報の分布は、回答セットと母集団の両方を階層のような副次母集団に分割し、その後のウェイト調整の構築のために使用される (Smith 1991; Särndal et al. 1992; Deville, Särndal and Sautory, 1993)。ウェイトのカリブレーションはより先進的な理論に基づく。それは、補助的情報分布の数がただちにウェイト調整に、最も多くは限界的分布として使うことができるというものである (Deville and Särndal, 1992; Lundström, 1997)。補助的情報が母集団あるいは階層から来ることは非常に多く、ときとして標本レベルの情報が使用されうる。カリブレーションのためのソフトウェア生産物は、例えば、CALMAR (Sautory, 2003) や CLAN97 (Andersson and Nordberg, 1998) をふくむ。

勧告

- 標本ウェイトは、標本設計、参入確率、母集団の大きさ、標本の大きさ、および回答者数についての情報に基づいて作成される。それらの存在は、ほとんど常に、階層といった母集団の副次集団に基づいて計算される。もし母集団がグループ化されるなら、各要素は分類の一つに属していなければならない。母集団レベルの情報は、標本抽出のウェイトがひきだされるときには、母集団レベルの情報が利用可能なはずである。
- もし不等参入確率が使われるなら、これらの確率は各標本要素についてひきだされるべきであり、ウェイトを導き出すために使うことができる。モデルに支援された推定が適用されるときには、可能な限り信頼できるウェイトづけ式を見出すために、暫定的推定値を導くことが有効であることが多い。ウェイトの正しさが確保されるべきである。標本データ集合にわたるウェイト変数の合計を計算し、拡大ウェイトの場合には、この合計は母集団の大きさと一致しなければならない。

- 階層別標本抽出設計および/あるいは階層後が適用されるときには、各階層内のウェイトの合計は対応する階層の大きさに等しくなるべきである。カリブレーションウェイトについては、ウェイトつき推定値はまた限界分布に厳密に等しく、また、情報が使われるレベルでの連続的補助変数の合計に等しいものであるべきである(Deville and Särndal, 1992; Särndal and Lundström, 2005)。
- 種々の標本設計に対する設計ウェイトは、単位無回答が無い仮定からひきだされる。乖離について正す最初の努力は、単純調整として、当初の標本の大きさを回答者数で置き換える単純な調整でありうる。次に、われわれは、すべての要素は等しい経験的応答確率を持つことを仮定する。
- ウェイトの変動が、可能なら標本設計の計画局面ですでに検討されるべきである。ウェイトの大きな変動、そして相対的に小さなウェイトの間で特に少ない非常に大きなウェイトは、推定量の分散に強い影響を持つことがありうる。
- もしウェイトの分散が深い階層化によるなら、階層をまとめること、あるいはそれら階層内の標本の大きさを増やすことが考えられるべきである。ウェイトの大きな分散が階層後によるなら、その他の分散は、階層のまとめによって減らすことが出来る。同じようにして、ウェイトのカリブレーションが適用されるときには、限界分布 (marginal distributions) が、より少ない階級を使って再計算することができる。

文書化

方法を報告する際の重要な要素は、使われた標本設計、推定と使われたウェイト調整法の詳細な説明である。報告は、最も重要な構造、偏りの種々の源泉、特に無回答のその分析、の説明、使用された調整とウェイト調整法の説明をふくむべきである。品質と方法の説明は、最終的ウェイトの作成によって重要な推定値について得られた結果をふくめるべきである(1.2.3節参照)。方法の説明はまた、基本的設計に基づいた分散推定量に対するウェイト調整によって起こったいかなる変化をも明確に報告し、適切な標準誤差や他の品質指標が可能な限り計算できるようにするべきである。

文献

- Andersson, C. & Nordberg, L. 1998. A User's Guide to CLAN. Örebro: Statistics Sweden.
- Deville, J.-C. & Särndal C.-E. 1992. Calibration Estimators in Survey Sampling. *Journal of the American Statistical Association*. Vol. 87. No. 418. 376 – 382.
- Deville, J.-C., Särndal, C.-E. & Sautory, O. 1993. Generalized Raking Procedures in Survey Sampling. *Journal of American Statistical Association*. Vol. 88. No. 423. 1013 – 1020.
- Ekholm, A. & Laaksonen, S. 1991. Weighting via Response Modelling in the Finnish Household Budget Survey. *Journal of Official Statistics*. Vol. 7. No. 3. 325 – 337.
- Lundström, S. 1997. Calibration as a Standard Method for Treatment of Nonresponse. Doctoral

- Dissertation. University of Stockholm. Department of Statistics. Stockholm.
- Oh, J. L. & Scheuren, F. 1983. Weighting Adjustment for Unit Nonresponse. In: Madow, W. G., Olkin, I. & Rubin, D. B. (eds.). *Incomplete Data in Sample Surveys. Theory and bibliographies.* Vol 2. New York: Academic Press. 143 – 184.
- Rosenbaum, P. R. 1987. Model-Based Direct Adjustment. *Journal of the American Statistical Association.* Vol. 82. No. 398. 387 – 394.
- Sautory, O. 2003. CALMAR2: A new version of the CALMAR calibration adjustment program. *Proceedings of Statistics Canada's Symposium 2003.* <http://www.statcan.ca/english/freepub/11-522-XIE/2003001/session13/sautory.pdf>
- Smith, T. M. F. 1991. Post-stratification. *The Statistician.* Vol. 40. No. 3. 315 – 323.
- Särndal, C.-E., & Lundström, S. 2005. *Estimation in Surveys with Nonresponse.* Chichester: John Wiley & Sons.
- Särndal, C.-E., Swensson, B. & Wretman, J. 1992. *Model Assisted Survey Sampling.* New York: Springer-Verlag.
- 追加情報**
- Bethlehem, J. G. & Keller, W. J. 1987. Linear Weighting of Sample Survey Data. *Journal of Official Statistics.* Vol. 3. No. 2. 141 – 153.
- Djerf, K. 2000. Properties of Some Estimators under Unit Nonresponse. *Statistics Finland. Research Reports 231.* Helsinki.
- Kalton, G., & Maligalig, D. S. 1991. A Comparison of Methods of Weighting Adjustment for Nonresponse (With discussion). The US: Bureau of the Census. *Proceedings of the Annual Research Conference.* 409 – 447.
- Laaksonen, S. 1988. Katovirheen korjaus kotitalousaineistossa. (Correcting for Nonresponse in Household Data; English summary). *Statistics Finland. Studies 147.* Helsinki.
- Sautory, O. & Le Guennec, J. 2005. La macro CALMAR2. Redressement d'un échantillon par calage sur marges. I.N.S.E.E. Paris: I.N.S.E.E.

2.12 統計的推定と分析

視角と目的

この節は、叙述的および分析的調査の両方にあてられる。記述的統計調査における関心の焦点であるパラメータは、通常は、総母集団レベルあるいはその部分との関連での総数、平均値ある

いは割合である。分析が強調される調査においては、関心は、現象間のつながりと相互依存に向けられる。パラメータは線形モデルといった統計的モデルと結びついており、例えば、相関あるいは回帰係数によって表される。両方のタイプの調査における重要な段階は、未知のパラメータを可能な限り信頼できる形で推定することである。分析的調査は記述的調査よりも結果についてのより広い解釈をふくむことが多い。センサスをふくむ叙述的調査は、統計機関においては代表的である。しかし、統計的モデルの利用は、例えば、回帰モデルを用いた推定技術の広い適用によってなど、拡大してきている。

この節は、統計分析の品質的側面とともに、推定量と其の品質属性を検討する。推定は未知のパラメータについての数的推定値を作成するために、利用可能なデータあるいはレジスターおよび他の行政データを使う。標本調査データとレジスターデータの結合の利用が、ここでは増加している。推定に必要な計算方法あるいはアルゴリズムは、推定量と呼ばれ、その数値は推定値と呼ばれる。総計は、叙述的推定の典型例である。

統計の利用者は、統計の品質についての必要な情報を提供されるべきである。推定値の信頼性は、標準誤差あるいは使われた推定量の分散係数の計算によって評価される。パラメータに対する信頼区間(例えば、95%信頼区間)は推定された標準誤差を使って計算できる。例えば、労働力調査の関心である主なパラメータは、失業率とともに、雇用者と失業者の総数である。これらの推定値はその標準誤差と信頼区間を伴っていないなければならない。

統計分析は、調査データがさらに、適切な統計方法によって研究される1つの段階である。例として、労働力調査データは、失業期間が回答者の年齢、ジェンダーあるいは職業とどう関連するかの統計的モデルによる研究に使われるかもしれない。統計分析はまた、調査過程の内部的品質を監視するために利用できる。その目的は、誤差の源泉の確認を通じて改善をし、データの信頼性へのその影響を評価し、それらの調整のための手段を提案することである。

推定と分析の方法は、経験的データの特性が問題とされている方法を支援できるように、注意深く選択されるべきである。したがって、当初の調査あるいは標本設計は、調査から信頼できる結果を得ることを考慮しなければならない。

i

原則

信頼性と結びついた推定量の最も重要な理論的属性は不偏性、精度および正確性である(Cochran, 1977; Lehtonen & Pahkinen, 2004; Lohr, 1999)。推定量は、与えられた大きさのすべてのありうる標本に対応して、すべてのありうる推定値の平均は、未知のパラメータの値と同じであるとき、不偏である。言い換えれば、推定量の期待値が母集団の値と等しいことである。より弱い、またより広く使われている属性は、推定量の一致性 (consistency) である。推定量は、もしその期待値が、標本の大きさが増すとともにパラメータの値に近づき、標本の大きさが母集団の大きさと同じときにパラメータと一致するならば、一致性をもつ。精度は、それらの期待値をめぐる推定値の変動のことである。この変動、すなわち分散が小さいほど、推定量の精度 (precision) は大きい。この特性は有効性 (efficiency) とも呼ばれる。正確な推定量は不偏 (ある

いは少なくとも一貫性をもつ) であり、かつ精度である。フィンランド統計局は、設計を基礎にした推定量— Horvitz-Thompsonタイプのそれ、あるいは一般化された回帰推定量あるいはカリブレーション推定量といったもの—の使用を促進する(Särndal et al., 1992)。

特に、推定量の有効率性特性は標本抽出あるいは推定設計に、補助的情報を取り入れることによって改善できる。そういった補助的データは、様々な統計的あるいはレジスターの源泉から得ることが出来る。補助的情報の挿入の例は階層別標本抽出やPPS標本抽出の異なる版がある。回帰推定、比推定あるいはカリブレーション推定量といったモデルに支援された推定方法は、推定設計への補助的データの参入の例である(Särndal et al., 1992; Lehtonen & Pahkinen, 2004)。改善された有効性は、補助データが調査変数と密接に関連しているなら、改善されている。これは、より小さな分散推定をもたらす。有効性は設計効果で測定される。この設計効果は、現実の標本抽出と推定設計の分散推定値と、対応する大きさの単純任意標本に基づいた分散推定値の比の計算によって得られる(2.6節参照) (Kish, 1965)。有効性は、設計効果が1以下であれば、改善される。

無回答によって生み出される偏りを正すためにまた、補助的情報を推定設計に導入できる。調査データは、例えば、単位無回答を含みうるし、ありうる無回答の偏りを正すために、ウェイト調整される。次に、通常はHorvitz-Thompson推定量である当初の推定量のウェイト構成が調整されるべきである。そうでなければ、階層後あるいはカリブレーションに基づいた推定量が使われるべきである(Djerf, 2000)。これは、様々な副次母集団あるいは領域を調査するとき、特に重要である。領域の標本の大きさが大きければ、標準化された推定量(例えば、Horvitz-Thompson推定量あるいはカリブレーション推定値)は十分に正確な推定値をもたらすことが多い(Lehtonen & Djerf, 2001)。他方で、小さな標本によって部分母集団を検査するとき、基本的な設計に基礎をおく推定量は、十分に正確であることに失敗することが多く、特別な小地域推定量が使用されるべきである(Rao, 2003; Longford, 2005)。特に、モデルに基礎を置く小地域推定量は、標本設計に関して偏りを持つかもしれない。それにもかかわらず、小さな分散を持ちながら、それらを使用することは、なお正当化されうる。

変化の推定は、(ローテーション・パネルをふくめて) 種々の形のパネル設計、あるいは標本調整(Kasprzyk et al., 1989; Diggle et al., 2002)を適用することによって改善されるかも知れない。パネルの摩損(長期のパネルにおける単位無回答)への調整においてとともに、そういったパラメータの推定において、補助的情報を利用することもできる。

分析的調査においては、モデルのパラメータに対して応用された推定量は、不偏的であるか、一貫性があるかである。データが複雑な標本設計を使って収集されるときには(2.6節参照)、モデルとの比較での不偏性の特性は、十分ではなく、標本設計に関しては、不偏性あるいは一貫性もまた必要とされる(Chambers & Skinner, 2003)。

ガイドライン

一般的アドバイスは、調査データの特性を可能なかぎり反映する推定および分析法を選ぶことである。信頼できる結果を生産するためには、分析は当初の標本抽出あるいは調査設計を考慮に

入れるべきである。これは、様々な設計の複雑性、例えば、標本ウェイト、階層化と集落化を考慮する設計を基礎にした方法を利用することによって達成できる。

標本調査

- 推定と分析のためには、調査データは、オリジナルな標本設計の適切な情報、すなわち、階層化および集落化を叙述する指標とともに、設計ウェイト、分析ウェイト、を含まなければならない。
- 当初の標本設計は、後の推定設計のために維持しなければならない。このようにして、階層化、集落化、多重の段階と局面に関する当初の情報、およびまた、1の確率をもって選択された要素に関する情報は、後に続くデータ利用者のために、そのまま維持され、文書化しなければならない。このデータは、階層化構造をこわすことなしに、調査状況に合致する形で推定されなければならない。
- 補助的情報は、信頼を以て利用可能なときには常に、活用されるべきである。補助的情報は何よりも、(事後階層化、比率推定や回帰推定に対応する推定量をふくめて)異なるカリブレーション推定量によるウェイト調整のために使うことができる。ウェイト調整は、母集団分布により近い標本分布をもたらす傾向をもつ。推定方法の選択については、方法論の専門家に協議することが勧告される。
- 点推定値の信頼性は、標準誤差の助けによって、評価されるべきである。それに代って、分散係数(相対標準誤差)が、当初の標本設計と推計設計を考慮して、可能な限り綿密に検討されるべきである。可能な場合には、設計の影響を推定することが望ましい。
- 複雑な標本設計と非線形推定量(例えば、回帰係数)の場合、推定量の標本分散は通常は分析的に得ることはできず、近似に訴えるべきである。統計的なソフトウェアのプログラムにおいて最も一般的な近似法は、線形化であり、他に、ジャックナイフ、均衡化された半標本、あるいはブートストラップ法といった標本の再利用に基づく方法がある(Wolter, 1985; Lohr, 1999; Lehtonen & Pahkinen, 2004)。近似法を選択するときには、方法の専門家に相談するべきである。
- 標本データは、はずれ観測値、すなわち、観測値の大量の集団から明確に除外される値、をふくむことが多い。これらの観測値が適切ならば、それらはそのままに置かれなければならない。すると、点推定は、調整された平均値あるいは分位数といった頑健な方法によって行うことができる。この偏差がウェイトづけによって生じているなら、ウェイトの分散を制限する手段がとられるべきである。
- 推定は、無回答を考慮するべきである。単位無回答の影響は、カリブレーションといったウェイト調整によって調整されることが多い(2.1.1節参照)。項目無回答の場合には、欠測観測値は統計的補定技法(2.10節参照)によって、ひんぱんに補定されている。補定値は、指標とともに知らせられるべきである。無回答の調整作業の影響は、分散推定において高炉されるべきである。
- 標本設計の複雑性に正しく反応する十分に設定されたソフトウェア・プログラムがある。それ

らは、点推定値とその標本分散を計算するために使われるべきである（例えば、SUDAAN ソフトウェア、SAS手続き、SurveyregとSurveylogisticおよびいくつかのStata手続き）。分析ウェイト（2.6節参照）が、調査データの分析に導入されるべきである。

レジスター・データと時系列

- 全体調査（センサス）。レジスターあるいは他の行政データ~のデータも統計的分析は2.5節でとりあげられる（行政記録と統計レジスター）（またWallgren & Wallgren, 2007参照）。
- 時系列データの推定と分析は、2.14節（時系列と季節調整）でとりあげられる。

文書化

優れた統計的実践は、データの処理に関する全生産の連鎖が明確に記述されることを求める。推定に関して、このことは、使用された推定量、モデル、推定技法、ソフトウェアの選択の正当化を意味する。

優れた統計実践はまた、統計的推定値の信頼性を利用者にして十分に報告することを求める。統計の信頼性の分析とその結果は、文書化するとき、特に将来のニーズに応えやすくするために、適切なカバレッジを持つべきである。データ利用者に提供される品質と方法論の報告書はその目的を満たす優れた手段を提供する。

標本調査データについては、主な調査変数の点推定値とその統計的信頼性と、全体母集団と最も重要な母集団の下位集団に関する情報が提供されなければならない。点推定値、標準誤差と信頼区間に加えて、設計の影響の報告が勧告される。標本調査データとともに、信頼性は、点推定値のありうる偏りと正確性の両方の評価を示す。

文献

統計的推定

- Cochran, W. G. 1977. *Sampling Techniques*. 3rd ed. New York: John Wiley & Sons.
- Djerf, K. 2000. *Properties of Some Estimators under Unit Nonresponse*. Research Reports 231. Helsinki: Statistics Finland.
- Kasprzyk, D., Duncan, G., Kalton, G. & Singh, M. P. (eds.). 1989. *Panel Surveys*. New York: John Wiley & Sons.
- Kish, L. 1965. *Survey Sampling*. New York: John Wiley & Sons.
- Lehtonen, R. & Djerf, K. (eds.) 2001. *Lecture Notes on Estimation for Population Domains and Small Areas*. Reports 2001/5. Helsinki: Statistics Finland.
- Lehtonen, R. & Pahkinen, E. 2004. *Practical Methods for Design and Analysis of Complex Surveys 2nd edition*. Chichester: John Wiley & Sons.
- Lohr, S. 1999. *Sampling: Design and analysis*. Pacific Grove: Duxbury Press.
- Longford, N.T. 2005. *Missing Data and Small-Area Estimation: Modern Analytical Equipment for the Survey Statistician*. New York: Springer-Verlag.

- Rao, J.N.K. 2003. *Small Area Estimation*. Hoboken: John Wiley & Sons.
- Särndal, C.-E., Swensson, B. & Wretman, J. 1992. *Model Assisted Survey Sampling*. New York: Springer-Verlag.
- Wallgren, A. & Wallgren, B. 2007. *Register-based Statistics: Administrative Data for Statistical Purposes*. John Wiley & Sons.
- Wolter, K. 1985. *Introduction to Variance Estimation*. New York: Springer-Verlag.

統計的分析

- Chambers, R.L. & Skinner, C.J. (eds.) 2003. *Analysis of Survey Data*. Chichester: John Wiley & Sons.
- Diggle, P. J., Heagerty, P., Liang, K.-Y., Zeger, S. L. 2002. *Analysis of Longitudinal Data*. 2nd edition. Oxford: Oxford University Press.
- Lehtonen, R. & Pahkinen, E. 2004. *Practical Methods for Design and Analysis of Complex Surveys*. 2nd edition. Chichester: John Wiley & Sons.
- Lohr, S. 1999. *Sampling: Design and Analysis*. Pacific Grove: Duxbury.

追加的情報

- Goldstein, H. 2003. *Multilevel Statistical Models*. 3rd edition. New York: John Wiley & Sons. Cochran,

2.13 統計データの提示

視角と目的

統計データの提示、出版および配布は、貯蔵とともに統計調査の過程の最後の段階である。提示は、それ以前の段階の結果が広く有効に利用できる方法を決定するのであるから、非常に重要な段階である。この節のトピックは、電子的（すなわち、インターネット）と紙の出版物の両方に適用できる統計データに関する通常の提示法である。しかし、提示方法はまたレジスターの表を扱うときに、注意深く考え抜かれなければならない。

統計データというものは、数量的データは、3つの方法、すなわち、文章の提示、表提示、あるいは図あるいはグラフ提示、によって描くことができる。表は、はるかに、最もひんぱんに使われてきた提示方法であるが、統計図がますます一般的になりつつある。うまく構成された図は情報を、表あるいは一片の文章よりも、わかりやすくかつ有効に伝え、多くの場合に、表あるいは文章形式の提示からひきださるものよりも、より完全な描写を提供するからである。他方で、人は、表における正確性は要求どおり高めることができるのに対して、図の正確性が常に限られていることを念頭に入れなければならない。図は、メッセージを速やかにまた説明的に伝えるのに対して、表からひきだされる情報は、遅く、そして表は説明的である。

図、表および文章の役割は、同じ事項が3つすべての方法によって提示されるので、少しだけ重複する。したがって、著者はそれらの異なる役割を考えなければならない。図は、中心的結果と伝えるべき最も重要な結果をとりあげるのに適切である。さらに、幾つかの現象の特徴は、グ

ラフによってだけ描写できる。図は、規則性と相互依存関係を叙述するはるかに最善の方法である。文章は通常、表や図よりもより主観的な表現の形態である。しかし、それは、結果の広い分析を可能にし、またその分析をより広い脈絡との関連におくことが可能である。他方で、表に既に見出される数字の単なる繰り返しは推薦できない。

一般的狙いは、図と表の両方は、そういったものとして追加的知識なしに理解できる不可欠な情報を伝えるように、自己充足的なものであるべきである。

原則

表についての原則

表には2つの基本的形態、すなわち、配列(array)と表がある。配列は通常文章の中の小規模の表であり、そのかなり自由な形式で、表とは違う。統計表は、相対的に厳密に定式化された提示であるが、ここではまた、読みやすさを高めるために異なる解決法を選ぶことができる。

表のセルは異なるタイプの情報をふくむかもしれないが、それが数字からなっているのが大部分である。セルの項目は、可能な限り同質的であるべきであり、理想的には、1変数の値が同じように測定されている形であるべきである。表の狙いの1つは、データを簡単に比較可能な形に置くことである。比較ができないなら、表全体の目標は疑問視されるかもしれない。

表の諸要素

表の題は、表領域の外部に置かれ、何を？どこの？何時の？という質問に答えるよう設計されるべきである。この題は、説明的で、明確で、簡潔であるべきである。もしすべての項目が同じ属性を示しているなら、その変数は題に含められるべきである。

列の見出しは位階的で、上のレベルの見出しが下のレベルの見出しを一緒にくくるものになる。変数と測定単位は、もしそれらが列のセルのすべてで同じなら、最低レベルの列の題に与えられる（ことが多い）。ときとして、すべての列の題がともに表頭の題とされることがある。

行の題は列の題よりも長いことがある。行の題もまた位階的で、左手の題が右手の題を合わせてくくることがある。変数と測定単位は、それらが行のすべてのセルに当てはまるなら、右端の題に与えられる（ことが多い）。

さらに格子線が表の読みを容易にするために使われて良い。

統計図の原則

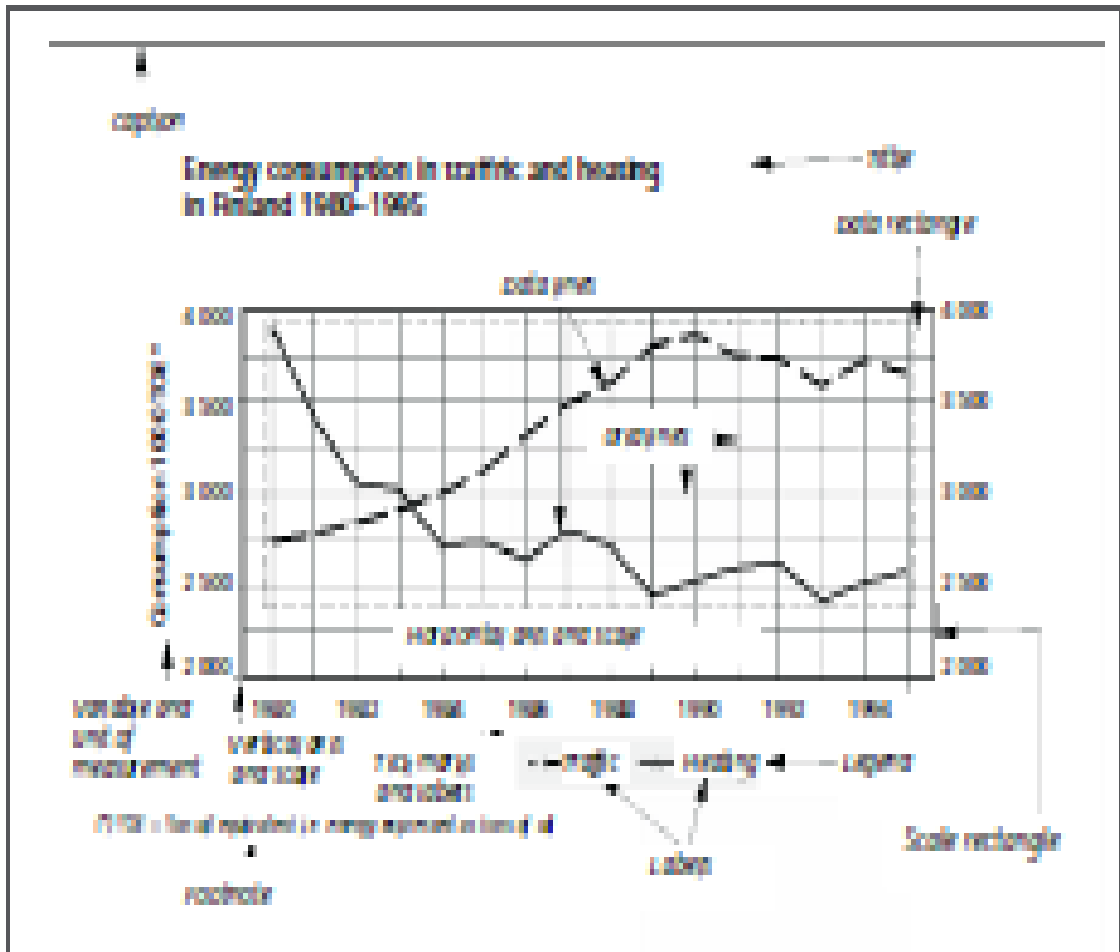
表と異なって、統計図の最も重要な機能は情報を視覚的に伝えることである。図のデザインにおいて、もしデザインの原則が完全に知られていないなら、意図しない真実をゆがめる危険がある。彼あるいは彼女の図が、とりあげている現象の真の性格の正確で明確な像を与えること、すなわち、それが正しい情報を伝えること、を確かにすることが統計家の責任である。

統計図は、データの構造の側面を示し、大量のデータを要約し、事象がどう結びついているかを示し、考えと結論を伝え、状況あるいは感情を提供することに関して、表よりも優れている。

図の諸要素

- 図の題は何が？どこで？いつ？という質問に答えるべきである。ある状況の下では、話の筋の題が十分な解決であり、他の場合（例えば、研究報告において）には、キャプション(caption)が題に置き換わることも起こる。他方で、題と見出しは相互に排他的ではない。
- 座標軸には十分にラベがつけられ、変数と測定単位が示されるべきである。
- 座標軸のきざみ(tick mark)とラベルは明確で十分なフォントの大きさをもって示されるべきである。
- すべての図の要素(線、部分、区切り、その他)にラベルが付けられるべきである。
- 解釈や結論を促進する援助は、十分に、しかし行き過ぎずに使われるべきである。
- 潜在的な食い違いが指摘され、説明されるべきである。

Chart 2.6 Parts and terms of a statistical chart



最も一般的な図のタイプは、線図、列と棒図、円グラフや主題別の地図である。これらのタイプの多くは幾多の副次的タイプを持つ。これに加えて、非常に知識を与えるが、それほどは使用

されない幾つかの図のタイプもある。

線グラフと垂直な欄図は選択肢になる。これらに共通して、連続変数—通常は時間変数—は、横軸上におかれ、それらは時系列を示すのに最も適している。横軸におかれる変数への最低限の要請は、その尺度が序数であるべきことである。

線グラフが変化と傾向を強調するのに対して、柱状図は大きさと大きさにおける変動を強調する。この違いによって、縦目盛り、あるいは数目盛りの波線は、線グラフにおいて重要性は小さく、柱状図においては、数目盛りの波線は視覚的声明を台無しにする。

横棒図は、階級や集団に付与される数量を描くのに適合している。縦軸上の変数は連続的であること必要としない。棒グラフは横棒図に代替するものではない。

円から成り立つ円グラフは、部分に分けられて、その領域が数量情報を表している。これは、円グラフが隣り合わせに置かれるとき、それらの面積が総量に対して比例する（円の面積は半径の二乗に比例して増加する）ことを、考慮に入れるべきである。

円グラフは比例的分布（%）を描写できるだけであり、情報を、線グラフあるいは棒グラフよりも不正確に伝える。円グラフは、データを正確に示すときには使われるべきでないし、同じ理由でパーセンテージ分布の数値を与えることは正当化される。

統計的な主題地図は地図上に置かれる統計図である。それらは、空間的分布を叙述するために使われる。最も知られているタイプは、おそらく脈絡膜（choropleth）図であり、そこでは1つの変数の比例分布を描写するための異なる方法において、境界を定められた地域単位が陰になってしまう。

図の装飾

図は、装飾をつける表示を作り出す必要を生み出すことが多いようにみえる。最も典型的には、人工的な第3次元を図につけ加えるものである。もうひとつの例は、絵画図の使用である。色の誇張された使用もこの部類に入る。しかし、統計的出版物でのそういった装飾についての適切な論議は見つけることは難しい。

絵画的表象図においては、例えば、人々を示す記号的なキャラクターが柱の代わりに使われる。それらは、それらがアピールし、劇的な表現であるので、人気のあるプレゼンテーションで使われる。

絵画的表象図をマスターするのが難しいので、その使用は全くあきらめるべきである。統計グラフの原則にかなう表象図(symbol chart)を作成するのは殆ど不可能である。

外見的な3次元構造は、絵画的表象と同じ理由で一控えめなトピックスを印象的なプレゼンテーションとするために—おそらく図において使われるだろう。追加される3次元はいかなる情報を提供することもめったにない。通常、図へのそういった装飾のモチーフは情報を伝えるその能力を傷つけるだけである。すなわち、第3次元は「安物図」であることが非常に多い。

3次元構造は、第3次元が本当に情報を掲載する場合にだけ使われるべきである。

ガイドライン

I 表のガイドライン

- 数はあまりに多くの桁まで示されるべきでない。まるめ (rounding) は数学的に正しくなければならない。
- 列や行の平均や総計は、これが必要とされるときと、比較を容易にするときはいつでも示されなければならない。
- 列と行の順序は論理的でなければならない。ときとして、最善の解決として、行と列の両方に関する変数の上降順を使用することがある。
- 列には比較の意味を持つ数を置く。というのは、数字は横に向けてよりも、下に向かっての方が読みやすいからである。
- 表のレイアウトは明確でなければならない。
- 通例の注釈は、特別な状況 (例えば、データの欠如) について存在し、注目されるべきである。説明を伴うそういった注釈は、例えば、フィンランド統計年報に見られる。

II 図のガイドライン

図のデザイン

- 最善の統計図の特徴は以下のとおりである。すなわち、それらは
- それに付された数値なしで、視覚的に伝達し機能する。
- 小さな空間で大量の情報を伝えている。最善の図は、大きなデータを整合的にする。逆に、わずかの数字から図をつくる意味はない。
- データのメッセージを破壊しない。図においては、大きさ (あるいは割合) に関する情報は数字と合致するべき視覚的象徴として示される。ということは、2倍の大きさを持つ量は、2倍の柱で示されるべきである。この場合、横になっている要因の数値は1である。
- 脈絡外のデータをグラフ化しない。題やラベルといった十分な追加的情報は現実の表現とつながりを持つ。
- 広い全体像から細かい構成まで、詳細さの幾つかのレベルでデータを示す。一瞥して何かが観察でき、更に調べると追加的情報を与え、より詳細な吟味が、それ以上のものを示す。
- 見る者が、グラフ提示の方法、デザインあるいは技術ではなく、図の内容に関心をもつように努力する。見る者がプレゼンテーション技術ではなく、示されているものの内容により関心を置くようにすることを考慮すべきである。

線グラフ

- 測定の連続的の尺度を持つ変数、例えば、時間、が横軸におかれるべきである。
- 縦軸上の数量は、出発点 (0) を必要とはしない。
- 横軸上の尺度の縦軸上の尺度との関係、アスペクト比率は図の表現に決定的な影響を与える。このアスペクト比率は、45度の傾きの線が均一的な傾向を示すように定められるべきである。

- 尺度は両軸上で等しい間隔を持つべきである、対数尺度だけは例外として許される。
- 線グラフでは余り多くのデータ線を取り入れるべきでなく、各線が判別されるべきである。

柱状図

- 測定の継続的な尺度を持つ変数、例えば時間、は横軸におかれるべきである。
- 縦軸の数量目盛りは中断してはならない。それは出発点(0)から出発するべきである。
- 尺度は両軸上で等しい間隔を持つべきである、対数尺度だけが許される例外である。
- 柱の間の空間は柱の広さの25~50%であるべきである。
- 柱の集団は最大で3つの隣接した柱をふくむべきである。
- アスペクト比率はこの図の表現に影響を与える。

横棒グラフ

- 横軸の数量目盛りは中断しない。他方で長すぎる棒は端でカットされうる。
- 横棒の集団は最大で3つの棒をふくむべきである。
- 柱の間の空間は柱の広さの25~50%であるべきである。
- 横棒は大きさの順、通常は最も長い棒をトップにおいて配列されるべきである。

円グラフ

- 1つの円グラフには最大で6つの区分が描かれて良い。
- 区分は、最大のものから最小のものへ大きさ順に配列されるべきである。
- 区分は12時から出発する時計まわりか、3時から出発する時計と逆まわりに置かれる。
- 情報は異なる大きさの隣り合わせの円グラフによって描写されうるが、同じ大きさであってもそれらと比較することは非常に難しい。

あなたの聴衆とメディアを知る

図をデザインするときには、聴衆と聴衆のプレゼンテーションを理解する能力の違いを考慮すべきである。例えば、広く配布される新聞向けの図と小さな範囲の専門家向けの図は同じ原則でデザインされるべきでない。プレゼンテーションの媒体(TV、スライド、OHP、新聞、インターネット、その他)も、各媒体が図を異なる形で作るので考慮されるべきである。これに加えて、プレゼンテーションのフォーラムが何らかの影響を持つだろう。

あなたの能力を知る

最悪の失敗はおそらく、図のデザイナーが、統計グラフの基礎についての知識すら持たずに、誰でもが統計図を作成できるという考えに陥ってしまうことから生じる。

ソフトウェアが提供するものだけに限定しない

統計図のプレゼンテーションには、様々の品質の特別なソフトウェアが使用可能である。残念なことに完全なものはない。結果が良くはない場合でも、既存のソフトウェアが提供できるも

のに限ってしまう誘惑は大きい。すべての可能性を捨ててしまう他の極端もまた結果として失敗となろう。少なくとも、ソフトウェアの不履行に自身を委ねない。

最もひんぱんに使われる図は、統計グラフ向けにデザインされたソフトウェアによって作成できる。しかし、それらの不履行は主として販売目的によって規定されている。したがって、不履行に自らを限ってしまうと、結果として通常は劣ったプレゼンテーションになる。より優れた結果は他の選択肢を試すことによって得られることが多い。統計グラフに通じているソフトウェアの技術者は少ないこと、そしてソフトウェアは、大部分が市場性の見地から開発されていることを留意すべきである。

3次元構造を使う前に考える

3次元構造は、第3次元が、他の手段では描くことが難しい情報を掲載するときだけに正当化される。

提示において文章、表、図の役割のバランスをとる

総てのコミュニケーション手段は、その役割と課題を持つ。図は、中心的な最も興味深い問題と検討を要する問題を描くことに優れている。表は、大量のデータを示すのに優れている。

文献

フィンランド統計局のガイドライン

Kuusela, V. 1994. Hyvin laadittu tilastokuvio. (Well-composed statistical chart). In Melkas, J. *Tilastokeskuksen ulkoisen tiedottamisen käsikirja*. (Statistics Finland's guidelines for releasing information; in Finnish only). Helsinki: Statistics Finland. 23–34.

追加的情報

Kuusela, V. 2000. *Tilastografikan perusteet*. (Basics of statistical graphics; in Finnish only) Helsinki: Statistics Finland & Oy Edita Ab.

Statistics Finland. 2006. *Statistical Yearbook of Finland 2006*. Official Statistics of Finland (OSF). Helsinki.

Tufte, E. R. 1983. *Visual Display of Quantitative Information*. Cheshire: Graphics Press.

Wainer, H. 1997. *Visual Revelations: Graphical tales of fate and deception from Napoleon Bonaparte to Ross Perot*. New York: Copernicus

2.14 時系列と季節調整

視角と目的

時系列は、一定の時間間隔をもって測定されることが多い。失業率と住宅価格がそういった定

期的に発表される統計の例である。発表される統計の絶対値に加えて、一般に以前の観察値からの変化の比較に関心がおかれる。

時系列は各観察の時点でインデックス化された観測値の集合として定義される。例えば、失業を叙述する時系列は、失業の数字を時間順に置くことで作り出される。時間次元で観測値の比較可能性を確保するためには、失業が、同じ方法論を使用して作成されることが必要である。

失業の長期的展開は時系列を視覚的に検討して観察できる。時間にわたる展開は、系列のトレンドとしても言及される。人はまた失業率は季節によって異なることを知ることができる。季節的失業による波に似た動きは、時系列の季節的変動の1例である。

季節調整は、系列の長期的展開を明らかにするために適用される。他の言葉でいえば、時期列の季節的要因を除去するのである。季節的要因は、年次的にほぼ同じ形で繰り返される年内の周期的な変動のことである(Gomez & Maravall, 1996)。周期的変動は、年のリズムに関連する他の現象(例えば、休日と消費習慣)によるとともに、季節によって起こされるである。これによって、季節変動は、年ごとに同じ形を持つ(Kokkinen and Alshail, 2005; Bell and Hillmer, 1984)。フィンランド統計局は、季節調整を、例えば、労働費用指数、産業生産の指標、および国民勘定の時系列に適用する。

季節調整は幾多の異なる方法によって行なうことができる。いわゆる事実上のフィルターは、主として移動平均法に基づいている。X-12ARIMAは、そういった事実上のフィルターの例である(Findlay et al. 1998)。TRAMO/SEATSはモデルに基礎をおく方法である ((Gomez and Maravall, 1996)。それらの方法の最善の特性を結びつけることは、研究の問題分野である(Findley, 2005)。

原則

季節調整は、時系列は以下の構成部分の合計として書くことができるという考えに基づいている(2.7図参照)。すなわち、

1. トrendは時系列の長期的変動を描く。
2. 循環変動は、例えば、経済循環による運動をとらえる。
3. 季節変動は、ほぼ同じ形で毎年再び起こる定期的変動を描く。
4. 不規則な構成部分は、前述の3つの要因には帰すことができない、系列の予測できない部分である。例えば、ストライキによって生じるある特定のはずれ値も、この構成要因に属している。

トレンドと循環変動を簡単でユニークな仕方で区分することは難しい。したがって、それらは、簡単化のために、トレンドと呼ばれる1要因とみなされる(Kokkinen and Alshail, 2005; Kaiser and Maravall, 2001)。季節調整済み系列として、季節要因が除去された系列を意味する。ということは、季節調整済み系列は、トレンドと不規則変動をふくむということである。トレンドと季節調整済み系列の両方は、フィンランド統計局の出版物に示されている。

Chart 2.7 Time series and their components

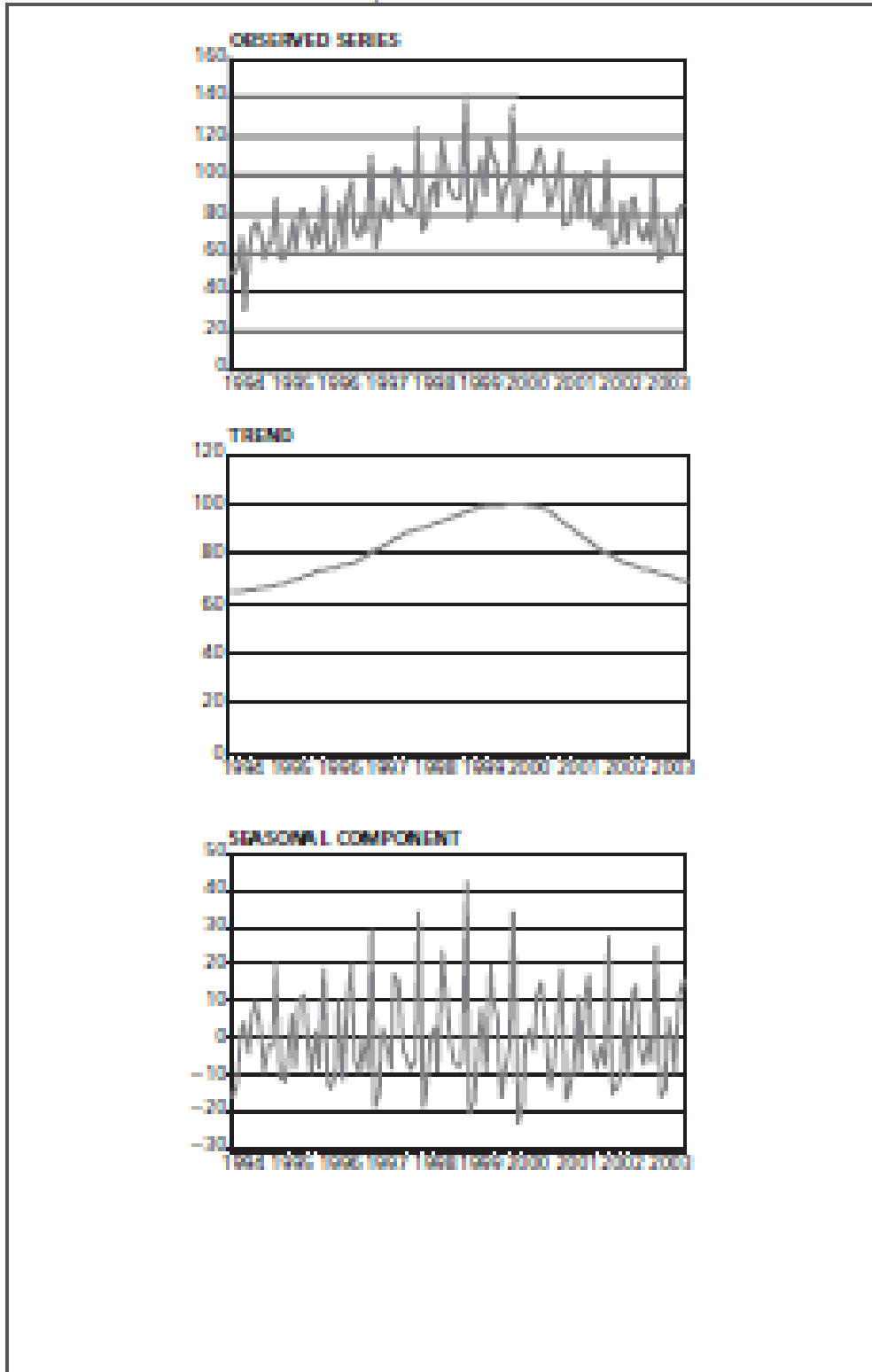
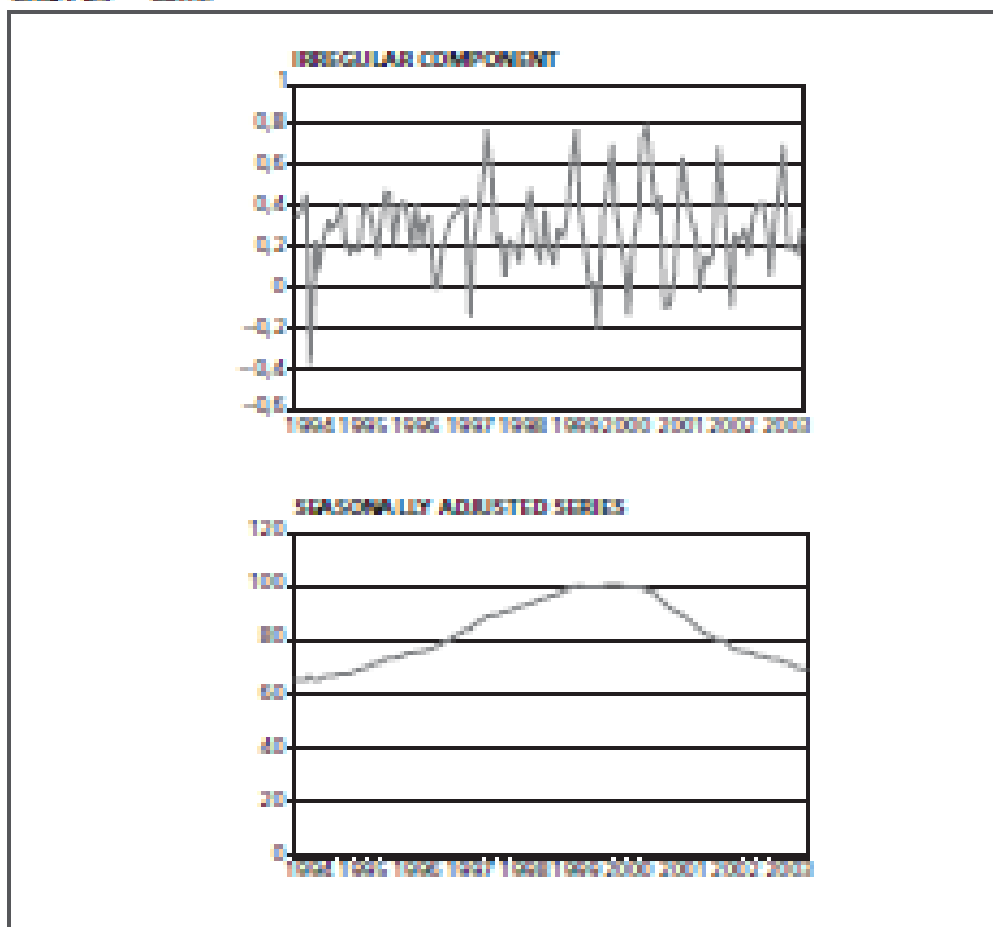


Chart 2.7 Cont.



モデルに基づく季節調整法であるTRAMO/SEATSはフィンランド統計局では2005年以来使われてきた。系列がTRAMO/SEATSで季節調整される前は、取引日修正が実施されている。これに加えて、ストライキや他の予期できない事象によるはずれ値が探り当てられ修正されるべきである。

その後、ARIMAモデル(Box and Jenkins, 1970; Box, Jenkins and Reinsel, 1994) をデータに当てはめることができる。

時系列のモデル化で、われわれは時間次元での観測地の依存関係を数式で叙述することを意味している。この式は時系列モデルと言われる。時系列モデルを当てはめることは、モデルの選択とパラメータの推定を意味する。その狙いは、データを単純だが正確な形で叙述することにある。時系列モデルによって、人は系列の周期性を検討し、将来の展開の予測を創り出すことができる(Hamilton 1994)。トレンドと季節要因の推定はこれらのモデルに基づく。

TRAMO/SEATSは、時系列の構成要素は、線形ARIMAモデルから引き出すことができるという仮定に基づいている(Box and Jenkins, 1970; Gomez and Maravall, 1996)。データが非線形的特徴を持っているなら、取引日の調整とはずれ値の探索は十分直感的には見えない。ときとして、視角

的には満足はいく季節調整が、診断テストに従うと最適ではなく、逆も同様であることに注意を払うかも知れない(Alsu hail, 2005)。それらの状況下では、最終的な季節調整済みの時系列を選択する前に、主な目的に近付き、時系列を使用すべきである。

以下の図2.7の図は、観察された時系列がどのように異なる構成要因の合計とみることができるとかを描いている。トレンド、季節変動と残差の合計である観察された時系列は、最初の図に示されている。トレンドは系列の長期的展開が描かれるのに対して、季節的要因の形は毎年ほぼ同じである。季節調整済み系列は、トレンドと残差からなり、それらは系列をよりジグザグにする。

ガイドライン

- TRAMO/SEATS法がフィンランド統計局の時系列の季節調整に使用されるべきである。
- 系列の季節調整の前に、はずれ値の探索と修正とともに、取引日調整が行われるべきである。
- 時系列が新しい観測値で更新されるとき、トレンドと季節調整済み時系列の最近値は大きく改訂されることになる。というのは、トレンドと季節調整済み時系列の最近値は部分的には予測に基づいている。この不確実性は、フィンランドと統計局の出版物で指摘されなければならない。
- 時系列の比較可能性を制約する定義あるいは方法の変更は明確に叙述されなければならない。
- 季節調整の成功は視角的にも、診断テストの両方によって、評価されなければならない。
- フィンランド統計局は、その値が出版物の表にふくまれていないグラフを公表するべきではない。

文献

- Alsu hail F. 2005. *Kausivaihtelun poistaminen taloudellisista aikasarjoista*. (Removing Seasonal Variation from Economic Time Series; in Finnish only). Statistics Finland. Internal memorandum. Helsinki.
- Box, G. E. P. & Jenkins, G. M. 1970. *Time Series Analysis Forecasting and Control*. San Francisco:Holden-Day.
- Box, G. E. P., Jenkins, G. M. & Reinsel, G. C. 1994. *Time series analysis: forecasting and control*. 3rd edition. New Jersey: Prentice-Hall.
- Bell, W. R. & Hillmer, S. C. 1984. Issues Involved with the Seasonal Adjustment of Economic Time Series. *Journal of Business & Economic Statistics*. Vol. 2. No. 4. 98 – 127.
- Findley, D. F., Monsel, B. C., Bell, W. R., Otto, M. C. & Chen, B. 1998. New Capabilities and Methods of X-12-ARIMA Seasonal Adjustment Program. *Journal of Business & Economic Statistics* Vol. 16. No. 2. 127–152.
- Findley, D. F. 2005. Some Recent Developments and Directions in Seasonal Adjustment. *Journal of Official Statistics*. Vol. 21. No. 2. 343 – 365.
- Gomez, V. & Maravall, A. 1996. *Programs TRAMO and SEATS. Instructions for the User*. Banco de Espana – Servicio de Estudios. <http://www.bde.es/servicio/software/tramo/guide.pdf>

Hamilton, J. D. 1994. *Time Series Analysis*. Princeton: Princeton University Press.

Kaiser, R. & Maravall, A. 2000. *Notes on Time Series Analysis, ARIMA Models and Signal Extraction*. Working Papers 0012. Banco de Espana. (Research report accessible on the Internet).
<http://www.bde.es/informes/be/docs/dt0012e.pdf>

Kokkinen, A. & Alshail, F. 2005. ARIMA-mallipohjaisesta kausitasoituksesta. *Kansantaloudellinen aikakauskirja*. (About ARIMA-Model-Based Seasonal Adjustment. The Finnish Economic Journal; in Finnish only) Vol. 101. No. 4. 469 – 483.

追加情報

Statistical Office of the European Communities. 2002. *Final Report. Task Force on Seasonal Adjustment of Quarterly National Accounts*. European Commission, Eurostat, European Central Bank.

2.15 指標と指数

視角と目的

ある期間にわたる同じ現象, 例えば経済的現象の測定は時系列として知られる統計データを作りだす。代表的な時系列の例は, 物価と商品量, 雇用者の賃金・俸給, 労働時間数, 有価証券の価格と数量についてのデータである。単純化して表わせば, 時系列は一定の期間にわたって同じく管理された現象について取られた正確な測定から得られた数字の連続である。このようにして, 時系列は, 関心を置く現象の一時的な発展を記述する(2.1.4節参照)。

うまく定義され抑制された意思決定においてすら, トピックは, 互いに同一基準では測定できない異なる種類の時系列によって描くことができるという事実によって, 時系列の利用は困難になる。「情報を管理できる時系列は単純にはあまりにも多くあるので, 時系列はすべてを直ちに含む。可能な限り情報の損失を少なくするように, それらを結合する必要がある」(Törnquist, 1974, p.21)。経済的な行為者は一般的には, 個別の時系列ではなく, 統計的一般化に関心を持つ。いくつもの時系列に含まれている統計情報をわずかのパラメータに要約する方法は, 時系列の総合化として知られる。時系列の総合化の最も簡単な例は, 与えられた一時点の観察値の算術平均の計算であろう。価格の総合化においては, これらの算術平均は, 場合によって単位価格あるいは単位値と呼ばれる。

単位価格あるいは値の測定における決定的問題は, 各時系列の観測値が合計全体に影響を与える相対的割合とウェイトである。もし相対的ウェイトづけが絶えず変化するなら, 総合化された時系列は一時的に比較可能ではない。価格と賃金の総合化の際には, 時間にわたって消費や労働力構成が変化するときには, この状況は優勢である。この問題は指数理論によって解決できる。

指数理論は経済学の一分野であり, 関心の焦点は, 品質で同じであり測定単位が自由な比率の総合にある。指数理論は, 情報を商品グループの物価および/あるいは数量の時間にわたる相対

的变化（発展）を測定する指数系列に要約することによって、質的に異なる商品の物価と数量の時系列の総合化（正規化した合計）を解決する。

例えば、時間にわたる消費構造の変化は、指数の質的变化を生みはしない。というのは、使われた指数算式次第で、比較されている消費構造は、基準およびあるいは比較期間の状況にしたがって標準化される。指数計算が答えるために努力する問題は、基準およびあるいは比較時の消費者の癖を認めつつ、比較時（基準時）の商品に対して基準時（比較時）にどれだけ多く/少なく支払われるべきか？ である。

一般的には、典型的指数は、価格と数量比率のいずれかの加重平均である。ウェイト構成は、基準時およびあるいは比較時のいずれかの状況を表わすように決定される。価格変化の評価に際しては、これは、物価と数量が同時に変化しても、数量的変化が価格変化の意図した測定を妨げることを許さないことを意味する。

簡単にいえば、指数理論は、記述的および分析的（経済的-理論的）研究法に分けられる。分析的研究方法は、消費者と企業といった経済的行為者は、数学的な行動モデルに従って、合理的に振るまうという仮説に基づいている。分析的研究における問題は、経済的行為者は、例えば、その社会・経済的地位あるいは個人的な選好の変化の結果として、時間にわたってその行動を変えるかもしれないことである。消費市場の商品—それらの例としては、コンピュータ、携帯電話、タイプライター、および蒸気機関、がある—の参入と離脱といった経済的現実には、経済的行為者の行動モデルに同化することは難しい。経済理論に基づく指数は、理論にしばられ、数学的には難しいことが多い。記述的方法是Vartia (1976), Samuelson and Swamy (1974), Allen (1975) and Diewert (1974, 1975)で論じられている。

指数計算に関しては、記述的方法是、平均値(means), 幾何平均値, 平均 (averages) および分散といった最も一般的な統計的パラメータを使う。それは、経済的行為者の母集団についての理論的参照枠組みを指定することを要求はしない。簡単にいえば、記述的方法是、指数を価格あるいは支払い水準あるいは個別価格あるいは賃金の平均的变化を測定する道具と規定している。記述的方法是Vartia (1976,1983, 1995), Vartia and Vartia (1984)で論じられている。

原則

指数は、与えられた商品あるいは商品集団（マーケットバスケット：商品籠）の物価およびあるいは数量が、基準時から比較時にかけて、相対的表現でどれだけ変化したかについて示す相対変化の指標である。指数計算の出発点は、同質の商品集団と個別価格、数量と価値関係—いずれも、これらの商品に適用された標本抽出によって確定されている—である。同質の商品（例えばエメンタルチーズ）あるいは同質商品（例えばチーズ）からなる商品集団のグループは、通常は個別価格あるいは数量の指数から計算される加重平均値である。比率のウェイトは、基準時およびあるいは比較時の価格および数量データから決定される。

単位価格指数 (Unit value index)

単位価格指数は時間にわたる平均物価の変化を測定する。例えば、それは、基準時と比較時の間の2つの構造的に異なる商品バスケットについて、価格比率を叙述する。基準時と比較時の商品バスケットは構造的に異なるので、平均物価の変化のどの部分が真の価格変化に関係し、そのどれだけ部分が消費構造の変化によって生じたかを判断することは不可能である。単位価格指数は、基準時と比較時の価格およびあるいは数量指数は相互に相関していないという条件の下で、価格変化の唯一の正確な指標である(Vartia, 1995, p.6)。

平均値の比較である単位価格指数は、基準時と比較時の平均値が正確に測定されたときですら、相対的価格変化の優れた尺度ではない。言い換えれば、基準時と比較時で計算された平均値の標準誤差が0—それらが母集団の平均と同じであろうことを意味する—であったとしても、平均値の比較は、優れた指数を構成しない。というのは、基準時と比較時の商品バスケットがなお異なった構成になっているだろうからである。単位価格指数の誤差は、それが含むインプリシットな数量指数の弱点から生じる。構造的に異質な数量の合計は、数量の発展を正しくは記述していない。客観的比較は常に商品バスケットの構造的均一性を必要とする。

ラスパイレス、パーシェ、エッジワース、フィッシャー指数

典型的には、基準時と比較時に物価と数量の両方が同時に変化する。この結果、比較されるべき商品バスケットはおのずから構造的に異なっており、したがって、例えば平均価格の直接的比較は、価格変化のどれだけ部分が真の変化を表わているか、どれだけ部分が構造変化によって引き起こされたかを明らかにしない。指数計算において、この問題は、ラスパイレス、パーシェ、エッジワースやフィッシャーの指数のように、商品バスケットが構造的に同一と仮定される仮説的な比較状況を構成することで解決される。

ラスパイレス指数では、消費構造を基準時の消費構造が出发点として選択され、仮説的消費支出は、基準時のバスケットの商品について比較時の価格で構成される。この指数は、消費構造は、基準時のそれに対応させて固定するという条件の下で、基準時と比較時の相対的価格変化を与える。ラスパイレス指数は物価変化による代替効果を許さないで、それは真の価格変化を過にか過小に推定することがある。真の価格変化に対する指数の正確性は、消費の量や物価が基準時から比較時にどう変化したかに依存する。もし物価と数量の変化が相関していなければ、指数の偏りはなくなる。

もう一つの良く使われるパーシェ指数と呼ばれる指数においては、消費は比較時の構造に標準化される。この指数は、比較時の商品バスケットが、比較時の価格での費用と比較して基準時の費用がどれだけ大きいかわかを語る。パーシェ価格指数の問題点は、ラスパイレスのそれと類似しており、価格変化によって生じる消費の再配分をふくめることができず、真の価格変化を過大にか過小にか推定することである。

ラスパイレスとパーシェ指数の基本的問題点は、それらが消費構造を基準時か比較時かのいずれかから持ち込むことにある。フィッシャー式はラスパイレスとパーシェ指数の幾何平均として

計算されるので、先の2指数よりも価格変化のより正確な尺度となる。それは、消費構造の変化を許すので、価格変化の測定における消費の変化から生じる代替の偏りを除去する。フィッシャーの物価指数の問題点は、それが商品バスケットの明確な解釈を欠いていることである。

さらにラスパイレス指数とパーシェ指数の平均として推定される他の指数がエッジワース指数であり、ここでは指数は、指数のウェイトが古い商品バスケットと新しい商品バスケットの平均として決定される。それは変化した消費構造を認めるので、それはラスパイレスやパーシェ指数よりも価格変化のより正確な尺度になる。

さらに普通に使われる価格および数量変化の尺度には、Törnqvist, Vartia I およびSato-Vartia指数がある。それらにおいては物価、数量および価値は、一時的なスライディング尺度上で指定され、ここでは、それらは継続的に変化し、指数は時の関数を示す。これらの指数は、消費者の癖の変化を許し、これは、固定ウェイト（ラスパイレス指数とパーシェ指数）に基づくものよりも正確であることを意味する。

品質変化と指数計算

消費者物価指数においては、商品の品質変化は標本抽出を通じて評価できる。このことは、基準時と比較時の間の価格変化の測定は、同じ品質の商品上でだけ行われることを意味する。例えば、住宅会社のシェア、宅地、一戸建ての価格、賃金・俸給の場合には、品質変化から生じる指数計算問題は、標本抽出法だけでは解決できない。指数の計算は少なくとも以下の点によって複雑になるだろう。

1. 統計単位、例えば住居あるいは雇用者は、基準時と比較時の間に、部分的あるいは全体的に変化するかもしれない。
2. 構造変化は、基準時と比較時の間の測定の目標で生じるかも知れない。
3. 統計単位の特徴の評価は、基準時と比較時間に変化するかも知れない。

統計単位が基準時と比較時間に変化しないままであるとすれば、基準時と比較時の比較グループにふくまれている統計単位は、それらの品質特性の点で異なることがありうる。価格変化の推定は価格変化のどれだけ部分が真一ということは、品質に関して調整された価格変化一か、どれが品質変化によるか、を確実に検証することはできない。

上のリストの第2項は、例えば、相対的見地から検討されるとき、分類された統計データの構成の一時的変化とむすびついている。例えば、職業、ジェンダー、教育レベルで分類された労働力は、基準時と比較時の間の比例的な労働投入との関連で変化することがある。対応して、住宅市場の相対的焦点は、基準時と比較時に地理的に変化するかもしれない。

ヘドニック指数法は上の3つの指数計算問題を解決するために使われることがある。この方法の狙いは、価格変化の測定における品質変化の影響を除去することである。この方法は、対数変化で測定された真の価格変化が以下のように、すなわち、真の価格変化=品質で調整された価格変化+変化した品質特性から生じる価格変化、という構成要因に分割される状況を導く。

方法論的には、真の価格変化のその構成要因への分解は、2つの大まかに区分されたヘドニック品質標準化法、すなわち、Griliches タイプのヘドニック法(Griliches, 1971)、あるいはヘドニック補定 (Koev,1997)で行うことができる。Griliches タイプのヘドニック法は、上記の項目1と項目2に示される指数計算問題の除去に集中する。Griliches 法で生産される典型的な方法は、住宅会社のシェア指数と中古車価格の指数である。ヘドニック補定法は、指数問題を2段階で解決する。すなわち、第1段階では、品質について標準化された推定価格は、統計単位—例えば、販売された資産あるいは、ある職業集団の雇用者—について計算され、第2段階では、それらの価格推定値が、典型的な指数計算とかなり同じ仕方で品質調整された価格指数を作り出すために使われる。ヘドニック補定が使われてきたのは、金属産業の賃金・俸給の測定、戸建て住宅あるいは宅地の価格の監視、品質について標準化された労働費用指数、および中央政府、地方政府および教会といった部門間の賃金格差の推定、においてである。

価格決定のモデル化は、ヘドニック法のこれら両方の典型である。価格は、条件付き期待値、統計単位の特性は条件として決まると仮定される。Griliches タイプの方法では、価格への特性の影響は推定期間中には標準的であり、したがって、それは個人的選好に起こりうる変化、すなわち、特性の市場評価の時間的な変化を受け入れない。価格モデルは典型的に、(標準的係数)分散分析タイプで、ここでは推定期間は、時間指標変数によって互いに分離されている。時間指標変数の推定係数は、基準に比較した価格の相対変化を直接的に明らかにする。この方法は、品質変化と指数計算の問題を同時に解決する。推定された価格変化は、すべての統計単位について、2つの時点間で比例的である。

ヘドニック補定においては、品質の標準化と指数計算が2つの方法論的段階において行われる。第1段階では、基準時と比較時の価格モデルは、機能的指定との関係で同じとして具体化される。基準時と比較時の価格モデルは、互いに独立に推定されるので、特性変数の係数は推定の間、制約されない。

この後に、品質に関する標準化が次のように行われる。基準時と比較時の価格モデルを使って、特性と構造(例えば、地域的、商品あるいは作業のインプット分類)を伴う基準時の商品が、比較時の特性と構造的要因の市場評価をもって、比較時にどれだけの費用になるか、を推定する。そして、比較される商品の特性とそれらの分類から生じる構造的要因は、基準時の状況に従うので基準時と比較時の商品は質的には同じである。基準時の価格と、比較時の品質について標準化された価格の間の変化は、例えば典型的なラスパイレズ式を使って計算することができる。

対応する手続きを、その特性と構造(例えば、地域分類)を持つ比較時の商品が、基準時の特性と構造の市場評価をもって比較時にどれだけの費用になるかを推定するために使うことができる。比較される商品の特性とそれらの分類から生じる構造的要因は、比較時の状況に従うので比較時の数量データが与えられるなら、価格変化は例えばパーシェ式を使って計算できる。ヘドニック的補定は、指数算式の選択を制約せず、すべての代表的な指数算式を使うことができる。

品質に関して標準化された指数に適用される見地は、あらゆる指数に対してと同じように、それらは指数計算でコントロールされた要因を条件として、条件付きの統計であるということ

ある。特徴としては、品質に関して標準化された指数値は、価格モデルの説明変数が変化すると変化することである。したがって、価格形成のどの説明要因が適合的であるかの検討が行われるべきであり、説明要因の選択は一般的には、承認された統計的方法によってテストできる。現実において価格に影響する要因が指数計算で許されないなら、指数系列の構成と比較可能性は疑問のあるものになる。

ガイドライン

指数作成戦略のガイドライン (Vartia, 1976):

1. 指数を使用する目的は、以下を具体化することによって決定されるべきである。すなわち、
 - 比較される商品の一般的性格、
 - 相対的変化を測定する視角に立つ経済的行為者
 - 相対的変化を推定する期間の長さ。
2. 指数計算の技術的問題は以下を具体化することで解決する。すなわち、
 - 比較される商品に適用されるべき分類
 - それらの商品の価格データの収集の方法
 - 適切なウェイト構成。
3. 指数計算方法は以下を指定することで決定される。すなわち、
 - 指数算式、
 - 指数系列作成の戦略。

指数は、基準指数 (base index) か連鎖指数 (chain index) かで作成できる。基準指数では、基準期間は、時間につれて比較時が柔軟に移動する間に不変のままである。連鎖指数では、例えば、基準時と比較時は連続する時期であるか、連続する年次の同じ四半期であり、期間はときとともに柔軟に移動する。

指数構成戦略の第4の問題領域は、商品の品質変化および、指数計算における新しい商品と消滅する商品に関する。

新商品と消える商品の問題を解決する際の決定的判断は、基準時と比較時の間に許されるべき時間間隔に関わる。時間間隔が長いほど、古い商品は新しい商品で置き換えられることはより確かである。これは、おのずから指数系列は一時的に比較可能ではないことを意味する。新商品と消滅する商品の問題は、連鎖指数よりも基準指数において大きいジレンマであろう。基準指数では、この問題は基準時と比較時の時間間隔を短くすることで軽減できる。

文献

Allen, R. G. D. 1975. *Index numbers in theory and practise*. London: Palgrave Macmillian.

Diewert, W. E. 1974. Intertemporal consumer theory and the demand for durables. *Econometrica*. Vol. 42. No. 3. 497-516.

- Diewert, W. E. 1975. *Ideal log-change index numbers and consistency in aggregation*. Discussion papers. No. 75 – 12. University of British Columbia. Department of Economics.
- Griliches, Z. (ed.). 1971. *Price indexes and quality change*. Cambridge: Harvard University Press.
- Koev, E. 1997. *Constructing a Hedonic Wage Index: Pilot Study for the Finnish Metal Industry*. Handout Statistics Finland. Helsinki.
- Samuelson, P. A. & Swamy, S. 1974. Invariant economic index numbers and canonical duality: survey and synthesis. *American Economic Review*. Vol. 64. No. 4. 566–593.
- Törnqvist, L. 1974. *Aikasarjojen konstruointi*. (Construction of time series; in Finnish only). Kautio Toikka, A. (ed.). Helsinki: Gaudeamus.
- Vartia, Y. 1976. *Relative changes and index numbers*. Research Institute of the Finnish Economy. Series A. Helsinki.
- Vartia, Y. 1983. *Aikasarjojen konstruointi ja indeksiluvut*. (Construction of time series and index numbers; in Finnish only). University of Helsinki, Department of Statistics handouts. No. 3. Helsinki.
- Vartia, Y. & Vartia, P. 1984. Descriptive index number theory and the Bank of Finland currency index. *Scandinavian Journal of Economics*. Vol. 86. No. 3. 352–364.
- Vartia, Y. & Vartia, P. 1995. *Quality and price of labour*. Discussion Papers. No. 366. University of Helsinki. Department of Economics. Helsinki.

3 文書化

3.1 統計調査の文書化

視角と目的

文書化とは、使用された概念、定義と方法、およびその情報システムと作業指令を伴う生産過程自体をふくむ統計活動の叙述をさす。これに加えて、結果の要約、情報の品質と比較可能性の説明が提供されなければならない。統計的文書化は、利用者を意図した文書化と統計の生産者を意図した文書化がある。

文書化によって、統計の利用者は次が可能であるべきである。すなわち、

- どの統計と統計データ(ファイル)があるかを見出す、
- 利用者が必要とする統計を見出す、
- 統計を解釈し分析する、
- データを処理する。

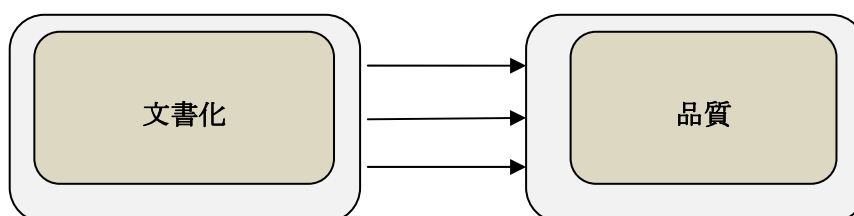
統計の生産者を意図した文書化は、彼らを次の点で助ける。すなわち、

- 信頼できる形で効率的に統計を生産する、

- それらの統計の質を改善する,
- 作業方法, 生産過程と情報システムを維持し発展させる,
- 他の類似の過程のために情報を収集する,
- 新しい雇用者に指示し訓練する。

文書化は統計的過程と統計生産物の文書化にわたることができる。過程本位の文書化は生産物がどのように生産されているかを叙述し, 他方で統計の利用者は統計とその品質の説明を必要とする。

図 3.1 統計調査の文書化の総合的品質とのつながり(Sundgren 2001)



原則

文書化は, 包括的であいまいさのない, 多様な利用者に適していることをめざすべきであり, 少なくとも, フィンランド統計局内では, 統計と統計生産の公的説明を示すべきである。それらの説明は, 最新のものであり, 十分組織され, 容易にアクセスでき, 簡潔で正確であるべきである。新しい職員は, 自ら生産過程の段階に習熟するよう使用できるべきである。文書化のための優れたおおざっぱなやり方は, 大きな問題の原因となることが文書化されないままに残されているなら, 何であれ文書化されるべきことである。

フィンランド統計局の共同のメタデータベース, すなわち, 統一化されたファイル, 概念・分類データベース, システムレジスター・分類データベースおよびアルヒーフ形成計画 (AMS: Archive Formation Plan). は文書化において利用されている。メタデータベースの利用者インターフェースは, メタデータを記録することへ利用者を案内する構成を持つ文書化を支援するよう企画されている。さらに, 指令と説明のモデルが, 利用者支援のために準備された。文書は, 分布データベース, アルヒーフと他の電子的手段および紙出版物といった異なる目的のために, メタデータベースから作ることができる。

それらのメタデータの文書化に加えて, 普通の実践は, 例えば, 過程の説明と作業指示の文書化のために発展されるべきである。統計調査の明確な説明は, 統計生産の継続性と発展を支援する。使用可能な文書化の範囲と正確性は, 統計の生産者と維持者および顧客のニーズに依存する。

ガイドライン

文書化は, 常に関心ある統計に対して顧客がいただくすべての要求を常に満たすべきである。統

計とその生産過程の全般的説明は、フィンランド政府統計の品質基準に対応させて構成された以下のチェックリストに従うべきである（1.2.3節参照）。すなわち、

1. 統計データの適合性

利用目的:

- 統計が生産された目的,
- 歴史。

定義:

- 「調査主題,
- データの収集者と提供者,
- 概念とその定義,
- 使われた分類。

2. 正確性と信頼性

方法:

- データ収集法,
- 標本抽出と推定方法。

データの統計的品質:

- 目標母集団と調査母集団の対応,
- 標本調査の無回答と無回答誤差を修正する方法,
- 測定, 処理, 標本誤差。

3. 適時性と迅速性

- データ収集時期,
- 発表時期,
- 迅速性, ないしは約束された時期の観察。

4. 整合性と比較可能性

比較可能性

- 同じ主題についての他の国と国際的統計と,
- 以前に収集された統計および時系列と,
- 勧告および基準と。

整合性

- 同じ主題についての他の統計との均一性/一貫性,
- 概念と収集過程の違い。

5. アクセス可能性と明確性

データの配布

- データ配布の通路,
- データの物的入手可能性,

- 利用者が必要とする形式の悉皆性,
- 入手可能なデータの説明。

データの解釈可能性:

- 使用された概念と方法の説明の入手可能性,
- 結果の分析のレベル。

6 文書化

情報システム:

- 構成, 情報モデル, 活動, およびバージョン管理
- 源泉コード, およびテスト,
- 処理とチェックの規則,
- 誤差と変化の管理。

生産過程:

- 過程の異なる作業段階(2.3図参照), 組織と責任。

データウェアハウス:

- データウェアハウスの技術的および全般的説明,
- データの技術的および全般的説明,
- 推論規則。

データ管理に関連するメタデータ

- 統計の名称,
- 統計のトピック,
- 責任者,
- データ提供の義務,
- 入手可能性,
- 秘匿性規則と利用制限,
- 貯蔵とアルヒービングとバージョン管理,
- EU規制,
- 使用された資源,
- 他の情報源と補助的情報。

文献

Sundgren, B. 2001. Documentation and Quality in Official Statistics. The International Conference on Quality in Official Statistics. Stockholm 14.–15.5.2001. Proceedings. Statistics Sweden & Eurostat.

Söder, L. 1998. Ehdotus Suomen Virallisen Tilaston (SVT) laatukriteereiksi. (Proposed quality criteria for the Official Statistics of Finland; in Finnish only). Statistics Finland. (Internal memorandum).

追加情報

Advisory Committee of Official Statistics of Finland. 2006. Basic and Quality Criteria for Official Statistics of Finland. TK-41-1227-06.

Haikala, I., & Märijärvi, J. 2006. Ohjelmistotuotanto. (Software engineering; in Finnish only). Helsinki: Talentum.

Statistics Canada. 2003. Statistics Canada Quality Guidelines. Ottawa.

von Rothstein, I., & Sundgren, B. 1996. Sweden's Statistical Databases: Detailed Descriptions of the Metadata in the Macrodatabase. Statistics Sweden.

3.2 文書保存

視角と目的

アルヒービングの目的は、将来のデータ収集、ロンジチュージナルデータと時系列の作成、異なる研究、および情報サービスに向けて、データの保持とその再利用を保証することである。

フィンランド統計局の作業原則は、それは国の記憶として活動すると述べている。統計データのアルヒービングの目的は、この知識資本を蓄積することである。アルヒーブ化されたデータの説明が高品質であることは、データの活用と正しい利用にとって重要である。

アルヒーブ管理の目的は、優れた記録管理と均一の情報管理の原則に従って、データの再利用を可能にし、データプールの管理を強める形で、データの保存を守ることである。

アルヒーブ法(831/1994)は、アルヒーブ管理の中心的機能は保存時間と文書の様式を決定することと、アルヒーブの形成の計画を維持することであると述べている。アルヒーブ法にそった義務には以下がある。すなわち、

- データの数値的評価と査定,
- 価値あるデータの保存,
- 保存したデータの組織化とカタログづくり,
- 研究目的のための借り入れと貸し出し,
- アルヒーブ化されたデータの情報サービス,
- 不必要なデータの処分, そして
- 非常事態に対するデータの保護。

アルヒーブ形成計画 (The Archive Formation Plan) は、フィンランド統計局の統計データについての蓄積指示として働く。それはデータが永久に、あるいはフィンランド統計局の判断によって、定められた期間について保存されることを規定している。国立公文書館は、永久に保存されるべきデータについてのアルヒーブ形成計画を承認している。すべてのフィンランド統計局の電子的データは、標準化された形、すなわち、統一化されたデータ・ファイルとして保管されなければならない。これはデータが、オリジナルのデータのインプットプログラムと装置が旧式にな

るとしても、利用者のニーズにしたがって将来回復できることを保証する。

原則

フィンランド統計局のアルヒービングの最も重要な道具は、統一ファイルシステムであり、これはファイルとファイルの説明、あるいは保管されるデータのメタデータをとりあげている。統一ファイルシステムはまた、フィンランド統計局の他のメタデータシステム、すなわち、分類データベース、統一されたデータファイルの現場での説明においてさらに利用される概念データベース、システムデータベース、およびアルヒーブ形成計画、を構成する。後者2つはファイルの説明で使用された。原則として、保管された説明は常に1箇所だけに貯蔵される。もし、既に分類データベースに叙述されている分類が保管されるべきデータで使用されるなら、これは、別個には記述されず、分類リンクによって、統一データファイルの現場説明にリンクされる。概念リンクが同じように作られる。アルヒーブ形成プランのリンクは、統一データファイルが、どのアルヒーブ項目に属するかを示し、システムコードはどのシステムにリンクするかを示す。

フィンランド統計局のメタデータシステムの使用がアルヒービングにおいて推奨される。というのは、それは説明作業を促進し、単純化するからである。アルヒービングのABCガイド (Taivainen, 2005b) は、アルヒービングと必要な応用、その利用と支援、利用者IDと応募者のインストロールについての情報に対して指示を与える。

アルヒービングの主な段階は次のとおりである。

- 第1段階.** アルヒーブ形成計画の説明におけるアルヒービング項目についての情報をチェックする。説明を更新することによって小さな変化がoccurする。本質的変化の場合（例えば、総計データから標本データへの変化）、新しいアルヒービングの項目が作成されるべきである。
- 第2段階.** 統一ファイルシステムの統一データファイルをアルヒーブ形成計画の説明にしたがって説明する。統一データファイルの名前は、アルヒーブ形成プランの説明にはリストされておらず、保管されるデータだけがリストされている。アルヒーブする項目は、幾多の統一データファイルを含んでいる。統一されたデータファイルに名前をつけるときには、アルヒーブ形成計画に定義されているシステムコードを使って、ファイルを将来にもまた見つけることができるようにするべきである。
- 第3段階.** 統一データファイルをUnixに保管する。アルヒービングが終了したとき、統一ファイルシステムの中のファイルは、貯蔵の領域に保管されたエントリーを持つ。
- 第4段階.** アルヒーブ形成計画に応じるときには、問題の統一データファイルの名前と他の発表情報を計画に付されている発表書式に書き込む。それから、その発表書式を、何らかの他の紙の材料とともに、統計のアルヒーブに向けて送る。もしアルヒービング項目が違う時間に完成するはずの幾つかの統一データファイルを含んでいるなら、発表書式を、すべての統一データファイルが保管されたときにはじめてその発表書式に記入する。

ガイドライン

- 統一ファイルシステムをアルヒービングに使うときには、統一データファイルの叙述は、常にデータの活用の容易性と正しい将来の利用を保証するために注意深く作成されるべきである。
- データの説明は叙述されているデータファイルに対応しなければならない。アルヒーブの説明をする前に、保管されるべきデータファイルについてのすべての情報は明確にされるべきである（それは何と呼ばれるか、どの版が保管されるのか、それはどこに置かれるか、だれがそうする権利を持つか、どれはどんな変数を持つか、使われる分類と概念は、分類データベースおよび概念データベースのそれぞれに叙述されているかどうか）。さらに、アルヒービングのガイドラインにも注意が払われるべきである(Taivainen, 2005a, 2005b,2005c)。
- 説明をすることは、常に協力の結果である。優れた説明は、生産過程とそれらのデータを生産するために行われる主要な決定の両方の理解をふくむ。この理解は、利用者においても促進されるよう伝えられるべきである。
- ファイルの説明は、情報内容が完成するときに速やかに行われるべきである。このタイミングは生産過程の性格に依存する。説明が十分早くに行われるときには、同じ説明がアルヒービングの前に、他のニーズ—例えば、PX-Webあるいは StatFinオンラインサービスデータベースのメタ説明—のために使用することができる。
- 説明が終わったとき、統一データファイルに責任を持つ者は、説明をチェックし、それを承認されたとおりにレジスター化けるだろう。承認された説明だけが保管される。

文献

フィンランドの国内法

Archives Act (831/1994).

フィンランド統計局のガイドライン

Sajjets, M. & Toivonen, P. 1995a. Keskuskoneen koontitiedostojen kuvausjärjestelmä. (Description database of the unified file system in mainframe; in Finnish only). Version 1. Statistics Finland. Internal memorandum. Helsinki

Sajjets, M. & Toivonen, P. 1995b. Keskuskoneen koontitiedostojärjestelmän käyttöohje. (Manual for the unified file system in mainframe; in Finnish only). Version 2. 2nd rev. ed. Statistics Finland Internal memorandum. Helsinki.

Syvänperä, R. & Jouhki, S. 2000. Käsitietokannan käyttö. (Using the concepts database; in Finnish only).

Statistics Finland. Internal memorandum. Helsinki.

Taivainen, T. 2005a. AMS-sovelluksen käyttäjän ohje. (User's guide to the Archive Formation Plan application; in Finnish only). Statistics Finland. Internal memorandum. Helsinki.

Taivainen, T. 2005b. Arkistojen ABC – Perustietoja arkistoinnista ja metatietojen kuvaamisesta (Archivers ABC – User's guide to filing data files and describing metadata at Statistics Finland; in Finnish only) Statistics Finland. Internal memorandum. Helsinki.

Taivainen, T. 2005c. Tilastoaineistojen arkistoinnin yleisohje. (General guidelines for archiving; in Finnish only). Statistics Finland. Internal memorandum. Helsinki.

4 統計利用の促進, 出版および顧客サービス

4.1 統計利用の促進

視角と目的

統計の利用の促進の視角は, 統計リテラシーと統計利用における技能を高めることとともに, 統計情報とそれに基づく生産物とサービスの有用性の改善, それらについての公共の知識の広がり拡大する。

統計の利用の拡大における狙いは, 社会における統計の広く効率的な活用を保証することである。これを達成するためには, 統計は容易に入手可能であり, 利用可能でなければならない。統計の入手可能性は, 一般のおよび生産物に特定な, あるいは統計に特定なコミュニケーションやマーケティングによって, そして見つけ出しアクセスが容易な生産物とサービスのマーケティングによって, 改善される。統計の有用性は, 統計生産物とサービスを, 明らかで, 説明的で, 利用者のニーズに適切なものにするによって, 高められる。統計の利用はまた, その利用と読みにおけるガイダンスと訓練の提供によって促進される。

フィンランド統計局のコミュニケーション戦略 (2003) に規定されている目標と原則には, それらの発行と顧客サービスにおけるとともに, 統計利用を促進する際に従うことになる。コミュニケーション戦略に従うと, 以下の原則が, コミュニケーションと顧客サービスにおいて観察されるべきである。すなわち,

- 信頼性 (Reliability)** : 発表されたすべての情報は正確であり, 信頼性のレベルが示される。
- 公平性 (Impartiality)** : そのコミュニケーションにおいて, フィンランド統計局は, データ生産の客観性, 独立性, 平等性を強調し, 統計倫理の原則に従う。
- 公開性と平等性 (Openness and equality)** : データは公正な形で公表され, その発表の時は事前に通知され, 関心を持つ誰でもに, 公表時に同時に受け取る機会を与えられるものとする。フィンランド統計局の職員は, それらの分野での積極的専門家であり, 期間の内外で, 統計とサービスに関する情報を提供するものとする。
- 理解可能性と明瞭性 (Understandability and clarity)** : すべてのコミュニケーションにおいて, フィンランド統計局は, 過剰な専門用語を避けることと, 統計によって与えられる叙述を分析することによって, 理解可能性と明瞭性を狙う。提示される情報は完全で, その利用者は自身の解釈や結論に向ける可能性を持つものとする。

フィンランド統計局の顧客サービスの原則は、サービス約束に定式化されている(4.3節参照)。

原則

生産物とサービスについての公衆の知識と入手可能性は、統計の一般的供給についての自覚を高めることと、最終利用者が必要とする情報を見つけるのを助けることによって、**促進される**。最終利用者は、冊子、新聞発表、顧客雑誌、ニュースそしてインターネットサービス上の特色ある項目によって、また多様なイベント、機会、催しにおいて、新しい情報、生産物とサービスについて通知される。新聞発表は、新しい統計データと研究結果、統計における方法の変化、現在の公的討論に関係するトピック事項、ニュースあるいは他の関連事項について行われる。印刷されたカタログがフィンランド統計局の生産物とサービスに関して毎年作成される。

統計の一般的な供給を描く情報やリストがフィンランド統計局のインターネット・サービス上で公表される。統計とその入手可能性についての情報は、統計リスト、統計の説明、フィンランド政府統計の系列における出版物リスト、統計の発表カレンダー、およびフィンランド統計局の週間カレンダーで見つけることができる。オンラインで公表されるリストとカタログもまた、電子的に発表された統計情報へのリンクを通じて、情報探索を支援する。インターネットサービスもまた統計生産物とサービスの説明をふくむ。

統計の図書館のTILDAカタログデータベースは、図書館で入手可能な材料の詳細をふくみ、WebStatサービスはインターネット上の最も信頼できる国際的統計データについての情報を提供する。出版物の交換実践、図書館システムと統計図書館の図書館間借用ネットワークは統計出版物が他の図書館でもまた利用可能であることを保証する。

統計の有用性は、方法論および生産物の発展を通じて改善される。最終利用者から体系的に収集されるフィードバックは、生産物開発にとって事実上重要である。有用性テストもまた、ウェブサービスの生産物開発で活用される。生産物の有用性は、表と統計図、読みやすさとテキストの構成の改善によって高められる。これは、出版物の編集の体系的ガイドと訓練の提供が求められる。フィンランド統計局は、国際的協定と規制にしたがって国際機関に対して、作成する統計と統計データを供給する。国際機関からの問い合わせはフィンランド統計局のレジスター局—ここは供給したデータの情報を保持する—を通じる。

統計的リテラシーと利用者の技能を促進するためには、フィンランド統計局は、他の統計生産者および教育機関やメディアといった最終利用者と協力する。フィンランド統計局は、顧客の訓練を組織し、そこでの課程のトピックスは、統計の編集、解釈、利用および統計に基づく生産物とサービス、をふくむ。適合的な訓練がまた顧客向けに整えられる。フィンランド統計局はまた統計におけるeコースを提供し、統計の解釈と利用における導入的授業、情報検索、統計的思考と主要統計の基本を与えている。

勧告

●統計の供給についての公衆の認識は、コミュニケーションとマーケティングによって高めら

れるものとする。

- 統計の入手可能性は、インターネット・サービスの発展によって、また図書館ネットワークでの統計の可視性を支援することで改善されるものとする。
- 統計生産物とサービスの有用性は改善されるものとする。
- 情報のニーズを持つ者は、情報の関連する出所の発見法のガイドを与えられるものとする。
- 顧客ニーズが確認され、統計生産物とサービスは、それらに基づいて発展されるものとする。
- 統計リテラシーと利用者の技能は、他の統計作成者や、教育機関やメディアといった最終利用者との協力の下に発展するものとする。
- フィンランド統計局は、公共サービス品質戦略(the Association of Finnish Local and Regional Authorities, 1998)に従い、それが作成した統計情報の配布を増加させることを狙うものとする。

文献

Association of Finnish Local and Regional Authorities. 1998. *Julkisten palvelujen laatustrategia*. (Quality strategy of public services. Project on quality strategy for public services; in Finnish only). Helsinki.

Statistics Finland. 2003. *Tilastokeskuksen viestintästrategia*. (Statistics Finland's Communication Strategy; in Finnish only).

4.2 出版

視角と目標

統計の倫理原則は、統計機関の社会的責任を強調する。信頼でき、十分に悉皆的な統計は民主的社會における事項の効率的な管理の基礎である。市民はその生活条件と社會の現実的描写を獲得し、その利害を調べ、意思決定に参加するために、統計情報を必要とする。行政、EU、ビジネス界、研究活動もまた信頼できて偏りのない統計情報を必要とする。

フィンランド統計局は、国の情報ニーズとヨーロッパ統計システムの両方のために生産物とサービスを計画し、発展させつつある。フィンランド統計局の主な生産物である社會についての一般的統計は、インターネットで総ての者が無料で入手可能である。フィンランド統計局は、協定とEU規制によって要求されている統計情報を、EUと国際機関に対して無料で提供している。フィンランド統計局はまた、統計データの利用者に、その多方面にわたる情報サービスで奉仕しており、そのサービスの幾つかは顧客の注文に対して無料で提供されている。出版の目標は、フィンランド統計局が生産した統計情報を遅れることなく利用者が入手可能とすることである。

原則

出版は、一般的利用のための統計情報を、例えば、印刷および電子的出版物および統計データ

ベースの形で配布することを意味する。フィンランド統計局の出版の目標と原則は、フィンランド統計局の活動戦略(Statistics Finland, 2003a)、ウェブサービスの開発方向 (Statistics Finland 2003b)、およびコミュニケーション原則 (4.1節参照)に沿っている。出版の原則は以下のとおりである。

- **信頼性 (Reliability)** 発表されたすべての情報は正確で信頼性のレベルが示されるものとする。
- **公平性 (Impartiality)** 情報は予定どおり発表され、誰でもが同時に入手可能であるものとする。
- **迅速性 (Immediacy)** すべての情報は、それが描く対象期間ののち、可能な限り速やかに発表されるものとする。
- **明瞭性 (Clarity)** すべての情報は、最終利用者のニーズを考慮にいて明確に提示されるものとする。情報の利用者は、自らの結論をひきだすあらゆる機会を与えられるものとする。
- **中立性 (Neutrality)** 議論のある社会的問題の扱いには注意と自制をすることが重要である。
- **解釈 (Interpretation)** すべての情報は、異なる現象の大きさと割合を叙述することによって、また変化と現象の原因と結果を説明することによって、解釈され、分析されるものとする。可能なときには、与えられた統計生産物内に含まれている情報は、同じ現象に関連する他の統計データや他の関連する情報と比較されるものとする。
- **適時性 (Timeliness)** すべての発表される情報は、現在の社会的論争や問題とむすびついている。フィンランド統計局は統計情報の生産と発行においてイニシャチブをとる。
- **公開性 (Openness)** 信頼できる統計情報は隠されてはならない。
- **案内 (Guidance)** 最終利用者は情報の獲得と探索で支援されるものとする。

発表は、ある時期を描く幾つかの統計からの主要データをはじめて公開する統計情報の発表のことである。統計法の原則に対応して、完成された統計データは、遅れることなしに発表され、情報をすべての利用者に同時に入手可能にされる(Statistics Finland, 2004a)。

新しい統計情報は、各統計トピックの統計ホームページとして公表される。言語法に対応して、情報は、フィンランド語とスウェーデン語の両方で同時にすべての利用者に発表される(Statistics Finland, 2004b)。統計的発表は、英語で、国際的重要統計について発表される (例えば、短期経済統計、国民勘定および人口発展, 2006a and 2006b)。データ発表と関連するデータ利用者への通知は、フィンランドのすべての国家機関に課せられている通りのフィンランド統計局の通知の一般的義務の一部である。

発表は体系的である。発表に至ることについての広い情報は、事前に年次の発表日程とより詳細は週間ニュースレター(Statistics Finland, 2006c)で提供されている。発表日程は、最終利用者が新しい情報の発表日にしたがって事前に計画することを容易にしながら、機関の活動において予測可能性を改善する目的に仕える。発表日程は、情報サービスにおいてだけでなく、マーケティングやコミュニケーションにおいて、過程の改善と顧客作戦を改善することに貢献する。発表計画に従った統計の発表は定期的に監視される。

フィンランド統計局は、社会の多様な分野をカバーする約200の異なる統計集合からデータを生産し、発表する。統計のほとんどはフィンランド政府統計（ODF）になる。フィンランド統計局が定期的に生産している統計から、毎年約700の発表が行われている。

データの秘匿性と発表前の開示

公表に至る統計情報は、予定発表時までは開示されないままで保管されなければならない（Statistics Finland, 2008）。特に、資本や金融市場に関連する統計は、秘密におかれるべきである。秘匿性のある統計の例としては、総生産の月次指標、四半期の国民勘定や消費者調査がある。他の統計の結果についても事前の情報は与えられてはならないが、すべての者に同時に公表されなければならない。

法規あるいは特別の取り決めに基づいて、特別な場合には、情報を前もって、例えば、政府統計の他の生産者あるいはEurostatに対して、それら自身の出版物の編集のために提供することができる。これらの場合には、禁止原則には、受領者は同意された発表時前に、いかなる第三者に対しても情報を提供してはならないことが常に明確に述べられている。

印刷された出版物

フィンランド統計局は、連続ものと別個ものの両方での印刷出版物を発行している。フィンランド政府統計の範囲内の統計からの主な印刷出版物のシリーズは、**フィンランド政府統計（OSF）**と呼ばれる。それはまた、フィンランド統計局以外の政府統計の生産者の出版物もふくむ。各出版物は、それが関心持つ統計的トピックを明示する。フィンランドの政府統計の諮問委員会は、**OFSシリーズの発表ガイドライン（2005）**を発表し、このシリーズに含まれている出版物のカタログを保持している。

OFSの出版物に加えて、フィンランド統計局は、研究（Studies）、評論（Reviews）、ハンドブックとジェンダー平等と題したシリーズの出版物を発行している。**研究**は高いレベルの科学的研究の出版シリーズである。このシリーズの出版物は、発行前に同僚の審査を受けている。**評論**シリーズは、これまで述べてきたシリーズでの発行には適していない統計報告および研究報告をふくんでいる。**ハンドブック**シリーズは、主な統計的ハンドブック、分類やガイドラインからなる。**ジェンダー平等**シリーズはジェンダー平等に関わる統計報告と研究報告をふくんでいる。

フィンランド統計局は一連の出版物の他に、別個の書物あるいはモノグラフ、および他の出版物（地方評論、ポケット統計および雑誌）を発行している。別個の書物は個別の高い品質の統計報告あるいは研究報告である。地方評論は、説明的な形での市による統計情報をふくむ。一般的出版物『**数字でみるフィンランド**』とトピックに特定した出版物の両方ともポケットサイズの統計として発行されている。フィンランド統計局の雑誌は、記述的記事によって統計についての情報を提供する。

電子的出版物

フィンランド統計局が定期的に生産している全部で200ほどの統計集合は、それ自体の恒久的ホームページをフィンランド統計局のウェブサービスに持っている。一般的社会統計とそこに含まれている統計データすべてインターネット上で総ての者が入手可能である一は、発行され、個別的統計のホームページに貯蔵される(Archive of statistical releases http://tilastokeskus.fi/til/arkisto/index_en.html)。データベースで発表される統計情報はまた、ホームページを通じて見出すことができる。現実には出版される統計データに加えて、このホームページは、説明、品質と方法論の説明、使用された概念と定義および分類といった、関心を持った統計についての深い、背景になる情報をもふくむ。

情報検索を促進するために、すべての統計は、フィンランド政府統計の主題分類に応じて区分されている(1.1.4節参照)。統計についての分類、インデックス化および広い背景データは、最終利用者がテーマ項目を分析し、問題のトピックについてより多くを学ぶことを助ける。

検索した誤差を訂正する手続き

発表された統計データで見つけ出された誤差は、フィンランド政府統計の諮問委員会の勧告(Advisory Committee of Official Statistics of Finland, 2005) およびフィンランド統計局の内部的ガイドライン(Statistics Finland, 2007) に対応して直ちに訂正されなければならない。このガイドラインは、統計データの異なる配布チャンネルにおいて、誤差がどう訂正されなければならないかのガイドを提供する。出版物あるいはウェブサービスで見出された何らかの誤差の速やかな訂正は、信頼の維持にとって決定的である。

訂正手続きは、暫定数値一そこでは、変化の可能性は、発表にあたって指示される一として発行される情報には関係しない。しかし、暫定データのいかなる明らかな誤差は適切に訂正されなければならない。

勧告

- 提示と記号の均一の形態、探索した誤差の訂正手続き、および発表日程の利用と内容に関するフィンランド政府統計とフィンランド統計局のガイドラインには、統計の出版と発表において従われるものとする(Statistics Finland 2003b; 2004a; 2004b; 2006c, 2007, and 2008; Advisory Committee of Official Statistics of Finland, 2005)。
- 問題の統計の編集に関与したか、あるいはその作業のためにそこからのデータを出版前に必要とするフィンランド統計局の雇用者だけが、その発表前に統計データを扱い、それらの情報を受け取るものとする(Statistics Finland, 2008)。
- 統計の信頼性を評価するために、編集方法は公表されており、一般的に受け入れるものとする。最終利用者に向けて発行される出版物とガイドラインの目的は、統計の正しい利用と解釈を促進することになる。この狙いに向けて、統計出版物は、品質説明一ここから、データの信頼性と正確性のあらゆる可能なリスクを結論できる一をふくむものとする(Advisory Committee of

Official Statistics of Finland, 2006).

文献

国のガイドライン

Advisory Committee of Official Statistics of Finland (OSF). 2005. Suomen virallisen tilaston julkaisuohje.

(Guidelines on publishing in the Official Statistics of Finland Series; in Finnish only).

Advisory Committee of Official Statistics of Finland (OSF). 2006. Suomen virallisen tilaston perusja

laatukriteerit. (Fundamental and quality criteria for Official Statistics of Finland; in Finnish only).

フィンランド統計局のガイドライン

Statistics Finland 2003a. Strategy. Reg. No. TK-04-1225-03 http://stat.fi/org/tilastokeskus/strategia_en.html

Statistics Finland 2003b. *Statistics Finland's development lines of web services*. Reg. No. TK-04-1225-03

Statistics Finland 2004a. *Päätös tilastojen julkistamisajankohdasta Tilastokeskuksessa*. (Decision about release times of Statistics Finland statistics; in Finnish only). Reg. No. TK-00-1576-04.

Statistics Finland 2004b. *Päätös kielilain huomioonottamisesta Tilastokeskuksessa*. (Decision on observation of the Language Act at Statistics Finland; in Finnish only). Reg. No. TK-00-1425-04.

Statistics Finland 2006a. *Päätös englannin kielilinjauksesta*. (Policy decision concerning the English language; in Finnish only). Reg. No. TK-00-1100-06.

Statistics Finland 2006b. *Englannin kielen käyttö Tilastokeskuksen tuotteissa ja palveluissa*. (Use of the English language in Statistics Finland's products and services; in Finnish only). Internal guidelines.

Statistics Finland 2006c. *Tilastojen ja tiedotteiden julkistaminen Tilastokeskuksessa – periaatteet ja käytännön ohjeet*. (Principles and practical guidelines on publishing Statistics Finland statistics and press releases; in Finnish only). Internal guidelines on Statistics Finland's intranet.

Statistics Finland 2007. Ohje: *Julkistetun tilastotiedon muuttuminen ja julkistamisen virhetilantee* (Guideline: Revisions to published statistical data and publication errors). Reg. No. TK-00-861-07.

Statistics Finland 2008. *Ohje tilastotietojen salassapidosta ja luovutuksesta ennen julkistamista*. (Guidelines on the non-disclosure and release of statistical data prior to their publishing). Reg. No. TK-00-23-08.

追加的情報

The Language Act (423/2003).

Prime Minister's Office. 2002. *Recommendation by the Prime Minister's Office on the Principles and Procedure of Central Government Communication*. Government Communications Unit. Helsinki

4.3 顧客サービスにおける品質

視角と目的

フィンランド統計局の基本的義務のひとつは、情報を可能な限り広く配布することである。その目的は、フィンランド統計局が生産した一般的社会統計からの主要な統計情報は、顧客がインターネット上で無料なセルフ・サービス利用できることにある。さらに、フィンランド統計局の統計図書館のサービスはすべての利用者が利用可能である。調査と面接サービスは有料のサービスとして提供される。統計情報はまた、顧客の注文に対して有料の仕事として供給される。

顧客サービスは、顧客が、連絡をとった部門に関係なく必要とする情報やサービスを得ることができるように、原則と実践の一貫した集合に基づくことが重要である。この理由によって、そして、顧客サービスのための均一で優れた実践を創り出すために、フィンランド統計局は2004年に顧客サービス原則を規定した。

原則

顧客サービス原則

顧客サービスの品質は、フィンランド統計局の組織イメージに大きな関係を持つ。顧客に供給される情報が、信頼でき、適時的であり、分かりやすく、利用しやすく、適切な価格であり、データ安全規制に従っているべきである。顧客のニーズには正当な注意を払われ、情報は、顧客がその活動においてその情報を優れた、費用効率的利用ができるような形式で提供されるべきである。顧客のフィードバックの収集と対応は、高い品質の基準とサービスの改善を維持するために重要である。

フィンランド統計局の全スタッフは、顧客サービスの原則を良く知っていて、この原則にしたがうと、

- われわれにはアプローチしやすい
- われわれはわれわれの顧客に柔軟かつ効率的に奉仕する
- われわれは、われわれの顧客ニーズに注意深く世話をする
- われわれは速やかかつ正確に作業する、
- われわれはわれわれの顧客に最新の情報を提供する
- われわれは継続してわれわれの顧客から学ぶ。

ガイダンスと情報サービス

最終利用者が必要な情報ニーズへのアクセスを得て、それを正しく解釈する方法を知ることが確かにするためには、フィンランド統計局の生産物とサービスについてアドバイスとガイダンスが重要である。フィンランド統計局の生産物とサービスおよびその価格についてのカタログが年次的に作成されている。これに加えて、フィンランド統計局のインターネット・サービスは、入手可能な生産物とサービス、関連するメタデータと分類、データ収集、品質叙述と統計的情報の入手可能性、をふくんでいる。

フィンランド統計局の顧客サービスは単位の境を越えた協力的努力である。顧客の電話は、不必要にある者から他の者にまわされてはならない。そうではなく、機関が電話による連絡を受

けたときに、最初の接触点は顧客の連絡情報と必要な情報を記録し、問い合わせに有効に対処できる組織内の連絡先を見出す。次に顧客は、情報が検索されるや否や連絡を受けることになる。顧客の問い合わせとフィードバックは、望むらくは直ちに、そして2作業日以内に返答されるものとする。

顧客からの仕事は秘匿して扱われなければならない。さらに、フィンランド統計局の情報サービスに関する法律やデータ安全の要件に従うことを確かとするすべての必要な手段がとられる(1.1.1 と2.2節参照)。許可証を必要とする資料については 顧客はその資料へのアクセスをうるために申請書を提出する必要がある(Statistics Finland,2005)。

主要な統計情報は、フィンランド統計局のインターネット・サービスの統計サイトでの無料のセルフサービスとして、最終利用者が入手可能である。短いガイダンスと情報一通常15分以内の作業時間を必要とする一を無料で提供されている。他の場合には、時間ごとに別個の料金が課せられる。

顧客サービスの窓口は通常の事務時間の間開いている。出版物への注文は、オンラインのStatMarketサービスを通じて24時間通して認められる。統計図書館、その施設、装置、コレクションは事務時間中に顧客の利用に供されている。フィンランド統計局は、Turku, Tampere, Seinäjoki と Oulu に地方サービス事務所を持っている。

勸告

- 総ての顧客は完全な尊敬をもって対応されなければならない。その狙いは、親切な専門的サービスの提供にある。
- 顧客は常に時間内に請求した情報を受け取るものとする。何らかの遅れが生じるなら、顧客には直ちに通知があるものとする。
- 顧客の請求に応じて提供された生産物やサービスには料金が課せられる。サービスの価格は現在の価格表に基づくが、これは国のサービスについての価格基準に関する法律と政令(159/1992, と211/1992,のそれぞれ), およびフィンランド統計局のサービス料金に関する財務省の決定(1208/2006)に基づいている。
- フィンランド統計局の生産物とサービスには品質説明が付される。さもなければ、顧客は統計の作成に使われた方法、使用された分類と概念についての情報を受け取るものとする。料金のある仕事については、契約条文と配達条件を顧客が入手できるものとする。
- 狙いは、顧客に誤差のない生産物とサービスを提供することである。提供された情報と生産物は常に、顧客に送られる前に検査される。
- 顧客のフィードバックは、フィンランド統計局のホームページ上の継続的な匿名のフィードバックのチャンネルを通じて収集される。スタッフはまた受け取ったフィードバックの記録を保持するものとする。すべての要求と不満への回答は親切に速やかに行われるものとする。
- 顧客のニーズと満足度は評価され、監視されるものとする。フィンランド統計局は、毎年、顧客満足度調査を行う。これに加えて、最終利用者調査が、特定の生産物とサービスについて行

われることがある。

- フィンランド統計局の雇用者は、フィンランド統計局の顧客サービス原則をあくまで実施する。
- 顧客サービス原則の実現は、調査や他のフィードバック法によって追求される。
- 文書化
- 課題あるいは情報サービス契約は、すべての情報サービスの仕事について準備されるものとする ((Statistics Finland, 2002)。
- 進行中の仕事のすべての局面は、記録され、顧客管理システムにとり入れられる。
- 提出される引用あるいは署名された契約といったあらゆる交換に関する情報は、顧客管理システムに保存されるものとする。

文献

国内法と法令および国内のガイドライン

Act on the Charge Criteria of the State (150/1992)

Decree on the Charge Criteria of the State (211/1992).

Decree of the Ministry of Finance Concerning the Charge Criteria of the Performances of Statistics Finland (1208/2006).

フィンランド統計局のガイドライン

Statistics Finland. 2002. *Tietopalvelusopimuksen yleiset sopimusehdot*. (General terms concerning information service agreements; in Finnish only).

Statistics Finland 2004. *Statistics Finland's customer service principles*

http://www.stat.fi/org/palveluperiaatteet/index_en.html

Statistics Finland. 2005. Guidelines on the Granting of User Licenses to Statistics Finland.s Unit-level Data. Reg. No. TK-00-128-05.

追加的情報

Heino, J., Kaisio, R., Koskenkylä, T., Notkola, V., Orjala, H., Ranta, J., Rantanen, J. & Siljander, A.2000.

Hinnoitteluopas. (Pricing guide; in Finnish only). Statistics Finland. Helsinki. Reg. No.TK-50-1453-00.

(Update due in 2007).

StatFin online data service. Statistics Finland , <http://statfin.stat.fi>

Stat.fi website service. Statistics Finland. , <http://stat.fi/index.html>

索引

【以下のページ数は原本のもの。本資料のページは、ほぼ原ページマイナス3~4】

A

accessibility; アクセス可能性 : 10-12, 15, 23, 26, 29-30, 115

accuracy; 正確性 : 10-11, 15, 17, 26, 28, 30, 40-41, 49-50, 62, 65, 66, 68, 91, 93, 108, 114

administrative data; 行政データ : 58, 61, 64, 66, 90, 93-94

administrative records; 行政記録 : 5, 14, 31, 48, 53-60

administrative registers; 行政レジスター : 6, 31-32, 50, 52-53, 77, 93

aggregation; 総計 : 41, 44, 106, 112

analysis; 分析 : 5, 28, 34, 36, 41, 61, 64, 66-67, 78, 80, 85-86, 90, 92-93, 95, 105, 110, 115

analysis weight; 分析ウェイト : 86, 92-93

Archive Formation Plan (AMS); アルヒーブ形成計画 : 114, 117-119

Archives Act; アルヒーブ法 : 117, 119

archives management; アルヒーブ管理 : 117

archiving; アルヒービング : 34-36, 94, 113, 116-119

ARGUS; 40-41

ARIMA; 101-102, 105

Auditing; 監査 : 22, 24, 32

auxiliary information; 補助的情報 : 34, 50-52, 61-64, 85-88, 91-92, 116

availability; 入手(利用)可能性 : 7, 14, 29, 37, 69, 115-116, 120-122, 128

B

balanced scorecard (BSC); 均衡得点カード : 18

bar chart; 棒図 : 97, 99

base period; 基準時 : 106-108, 110-111

bias; 偏り : 62, 64-65, 72, 77-78, 82, 87, 91, 93, 108

bootstrap; ブーツトラップ : 92

building code; 建築物コード : 55

Business Register; ビジネス・レジスター : 38, 53

C

calibration; カリブレーション : 87-89, 91-93

calibration estimator; カリブレーション推定量 : 89, 91-92

CALMAR; 88-89

CAPI; 72, 84

CASI; 134

CASI-A; 134

CASI-V; 134

CATI; 69, 72

CATI centre; 69

census; センサス : 27, 31, 46, 48, 56, 59-61, 63, 67, 77, 89-90, 93

Central Population Register (CPR); 中央人口レジスター : 54

change information; 変化情報 : 54

change statistics; 変化統計 : 56

CLAN97; 88

clarity; 明瞭性 : 26, 29-30, 47, 115, 120, 123

classification accuracy; 分類の正確性 : 41

Classification Database; 分類データベース : 46, 49, 114

classification principles; 分類原則 : 45-46

classification services; 分類サービス : 49

classification standards; 分類基準 : 46, 49

classification variables; 分類変数 : 52

classifications; 分類 : 3, 6, 12, 14, 27-29, 33-34, 36, 45-50, 74, 110, 114, 118, 124-125, 128

cluster sampling; 集落抽出 : 61, 63, 66

clustering; 集落化 : 61, 64, 92

code system; コードシステム : 47, 53

coefficient of variation; 変動係数 : 65, 90

coherence; 整合性 : 10–12, 15, 23, 26, 29–30, 115

column chart; 列図 : 97, 99

comparability; 比較可能性 : 6, 9–12, 15, 23, 26, 29–30, 115

comparison period; 比較時 : 106–111

completeness; 完全性 : 50, 52, 80

Computer Assisted Personal Interview; コンピュータ支援個人インタビュー : 72, 134

Computer Assisted Telephone Interview; コンピュータ支援電話インタビュー : 134

concepts; 概念 : 5–6, 14, 26–27, 29, 33, 36, 45–50, 59, 73–75, 79, 84, 113–115, 117–119, 125, 128

Concepts Database; 概念データベース : 47, 117–119

confidence interval; 信頼区間 : 65, 90, 93

confidential data; 秘匿データ : 37–39, 43

confidentiality; 秘匿性 : 8, 10, 24, 38–39, 43–44, 116, 124

consistency; 一貫性, 一致性 : 14, 26, 29–30, 57, 70, 91–92, 112, 115

core processes; 中核的過程 : 18–20

correctness; 正しさ : 25, 52, 57, 59, 70, 80, 83, 88

costs; 費用 : 8, 14, 34, 52–53, 56, 60–61, 63, 66, 69–70, 72, 128

covariance analysis; 共分散分析 : 110

coverage; カバレッジ : 5, 7, 26, 28, 49–52, 58–59, 64–65, 69, 71, 87, 93

crime nomenclature; 犯罪用語 : 45

cross-sectional data; クロスセクション・データ : 54

customer focus; 顧客焦点 : 16

customer needs; 顧客ニーズ : 17, 36, 122, 129

customer orientation; 顧客中心 : 3

customer service outlets; 顧客サービス窓口 : 128

customer training; 顧客訓練 : 49, 121

customs tariff nomenclature; 関税用語 : 45

cut-off sample; 切り捨て (カットオフ) 標本 : 64

D

data collection method; データ収集法 : 5, 27, 33–34, 61, 67, 69–70, 72, 76, 78, 115

data collection mode; データ収集形式 : 70

data management; データ管理 : 34, 43

data processing methods; データ処理法 : 34

data protection; データ保護 : 5–7, 33, 37–41, 43–44, 128

data protection methods; データ保護の方法 : 33

data security; データの安全 : 127

definitions; 定義 : 5–7, 12, 26, 31, 34, 45–47, 49, 74, 104, 113–114, 125, 128

demographic classifications; 人口分類 : 45

derived statistical activity; 派生的統計活動 : 31

derived statistical data; 派生統計データ : 31

design effect; 設計効果 : 66, 91–93

design weight; 設計ウェイト : 86–88, 92

design-based methods; 設計に基礎を置く方法 : 92

deviant values; 逸脱値 : 84

dissemination; 配布 : 3, 12, 14–15, 17, 34, 38, 94, 115, 125

documentation; 文書化 : 3, 5, 15, 19, 32, 34, 36, 49, 59, 66, 79, 88, 93, 113–116, 129

donor imputation; ドナー補定 : 83

dwelling code; 住居コード : 55

dwelling unit; 住居単位 : 42, 48, 50, 52, 55, 56

E

Edgeworth's index; エッジワース指数 : 108

editing; エディティング : 4–5, 27–28, 34, 69, 77, 79–85

efficiency; 有効性 : 14, 28, 62–63, 81, 85–86, 91

establishment code; 事業所コード : 55

estimation; 推定 : 5, 27, 34, 51, 64, 66, 78, 83, 85–88, 90–93, 102, 109–110, 115

estimator; 推定量 : 52, 66, 88–94

ethical principles; 倫理原則 : 7, 122

- evaluation; 評価 : 15, 17–20, 22–24, 26–27, 32, 36, 60,84,106
- evaluation process; 評価過程 : 36
- exhaustiveness; 悉皆性 : 46, 58–59, 115
- expansion weights; 拡大ウェイト : 86
- F**
- face-to-face interview; 面接インタビュー : 68–71
- family; 家族 : 42, 50, 54, 56–57
- Farm Register; 農場レジスター : 52
- field interview unit; 現場インタビュー単位 : 69
- Fisher's index; フィッシャー指数 : 108
- frame error; フレーム誤差 : 60
- frame population; フレーム母集団 : 50, 52, 78
- G**
- generalisability; 一般化可能性 : 49–50
- Guide to Classifications; 分類へのガイド : 49
- H**
- harmonisation; 調和, 調整 : 3, 5, 45, 47
- hedonic imputation; ヘドニック補定 : 109–110
- hedonic index methods; ヘドニック指数法 : 109
- Horvitz-Thompson estimator; Horvitz-Thompson 推定量 : 91
- household; 世帯 : 10, 31, 48, 50–52, 56, 62–63, 66–67,71, 73, 78–79, 83, 85, 89
- I**
- immediacy; 速時性 : 123
- impartiality; 公平性 : 7, 10, 120, 123
- imputation; 補定 : 5, 27–28, 34, 57–58, 78–80, 82–85, 109–110
- inclusion probability; 包含確率 : 62, 86
- independence; 独立性 : 7–10, 13–15, 30, 43, 120
- index calculation; 指数計算 : 106–111
- industrial classification; 産業分類 : 48
- information content; 情報内容 : 27, 50, 59, 119
- information system; 情報システム : 14, 36, 53–57, 113,115
- international classifications ;国際分類 : 47
- international concepts; 国際的概念 : 14, 26
- international recommendations; 国際勧告 : 6, 20, 46–47
- international standards; 国際基準, 標準 : 13
- International Statistical Institute (ISI); 7–8, 38–43
- interpretation; 翻訳, 解釈 : 12, 14, 18, 28, 82–84, 90, 97,108, 120–121, 123, 126
- interviewer training; 面接者訓練 : 34, 70
- ISO 3166; 47
- ISO 9000; 11, 20
- item non-response; 項目無回答 : 77, 83, 93
- J**
- jack-knife; ジャック・ナイフ : 92
- L**
- leadership; リーダシップ, 指導性 : 9, 13, 16
- legal units; 法的単位 : 52
- licence; ライセンス, 認証 : 128
- line chart; 線図 : 97, 99
- logical checks; 論理チェック : 77
- logical imputation; 論理的補定 : 82
- longitudinal survey; ロンジチュージナル調査 : 67, 70
- M**
- main group level; 主要グループ・レベル : 45
- Mainland Finland; フィンランド本土 : 48
- measurement error; 測定誤差 : 76, 77, 86
- metadata; メタデータ : 11–12, 19, 26, 29, 32, 34–36, 49,114, 116–119, 128
- methodological description; 方法の説明 : 22–23, 66, 125
- microdata; ミクロデータ : 23, 38–44
- missing data; 欠損データ : 77, 85, 94, 98
- model-assisted estimation; モデル支援の推定 : 87–88, 91
- model-donor imputation; モデルドナー補定 : 83

multiple imputation; 多重補定 : 83, 85

N

neutrality; 中立性 : 123

non-response; 無回答 : 5, 28, 34, 65–67, 69,
71–72, 77–78, 83, 85–88, 91–93, 115

non-response error; 無回答誤差 : 66, 78, 115

non-response rate; 無回答率 : 66–67, 69, 78

non-sampling errors; 非標本誤差 : 27–28

NUTS regional division; NUTS地域区分 : 48

O

obligation to supply data; データ提供義務 : 6, 116

Official Statistics of Finland (OSF); フィンランド政
府統計 : 25–26, 30, 46, 49, 50, 53, 65, 100, 114,
116, 121, 124–126

openness; 公開性 : 15, 20–21, 24, 37, 43, 120, 123

operating framework; 作業枠組み : 3, 5, 6, 32–33, 45

operational process; 作業過程 : 34–35

organisation of fieldwork; 現場作業の組織 : 34

OSF producers; OSFの生産者 : 25

outlier detection; はずれ値の探索 : 84, 102

overcoverage; 過大カバレッジ : 50–52, 59

P

Paasche's price index; パーシェ物価指数 : 108

panel research; パネル研究 : 55

panel survey; パネル調査 : 65, 67, 72, 94

paper questionnaire; 紙調査票 : 68–70, 73, 75

parameter estimates; パラメータ推定値 : 61, 64, 66,
85

parameters; パラメータ : 28, 52, 86, 90, 92, 106–107

partnership; 協力 : 16–17

personal identity code; 個人識別コード : 50, 53, 55

personnel development; 個人の発展 : 16

pie chart; 円グラフ : 97, 99

pilot study; 試験調査 : 32, 75, 112

planning process; 計画過程 : 33–34

point estimation; 点推定 : 93

population; 母集団 : 13, 27–28, 31–33, 39,

45–67, 77–78, 86–88, 90–94, 107, 115, 123

Population Information System; 人口情報システム :
53–55, 57

post-stratification; 階層後 : 86–89, 91, 92

PPS sampling; 62–64, 66, 87

precision; 精度, 精確性 : 61, 64–65, 85, 91, 127

presentation; 提示 : 5, 14, 28, 45, 65, 73, 94–95, 97–100,
125

press releases; 新聞記者発表 : 121, 126

privacy; プライバシー : 7, 10, 37–38, 44

processing error; 処理誤差 : 65

PRODCOM list; 48

producers of OSF; OSFの生産者 : 25–26

product classifications; 生産物分類 : 46, 48

product description; 生産物説明 : 34, 113

product quality; 生産物品質 : 16

production process; 生産過程 : 3, 17, 19, 26–27,
29, 113–114, 116, 119

professional ethics; 専門家の倫理 : 7, 8, 14, 38, 43–44

promptness; 迅速性 : 26, 28, 30, 115

public administration recommendation; 行政的勧告 :
46

public awareness; 公衆の自覚 : 122

public responsibility; 公衆の責任 : 17

publication; 出版物 : 3, 13, 15, 25, 28, 28–30, 34, 38,
44, 47–48, 76, 94, 97, 102, 104–105, 114–115, 121,
123–126, 128

punctuality; 時間厳守性 : 11–12, 15, 115

Q

quality assurance; 品質保証 : 17, 19

quality change; 品質変化 : 109–112

quality criteria of OSF; OSFの品質基準 : 29

quality evaluation; 品質評価 : 15, 17, 20, 22–24, 26,
60

quality improvement; 品質改善 : 15, 17

quality indicator; 品質指標 : 15, 30, 65–66, 84, 87
quality management; 品質管理 : 9, 13, 16–20, 22, 24
quality planning; 品質計画 : 17
quality reports; 品質報告 : 26–27, 29, 30, 34, 66
quality standards; 品質基準 : 10, 15, 17, 127
questionnaire; 調査票 : 5, 32, 34, 56, 60, 67–71,
73–77,84

questionnaire design; 調査票設計 : 5, 73–74
questionnaire testing; 調査票のテスト : 75

R

random sampling; 任意抽出 : 62–63, 65–66, 86
ratio estimation; 比率推定 : 86–87, 91–92
real estate code; 不動産コード : 55
real-donor imputation; 真のドナー補定 : 83
reference person; 参照人 : 56–57
regional classifications; 地域区分 : 45, 48, 110
Register of Completed Education and Degrees; 学歴・学位レジスター : 53, 56
Register of Enterprises and Establishments; 企業・事業所レジスター : 52, 54–58
register of income support recipients; 所得支援受給者レジスター : 55
register of job applicants; 求職者レジスター : 55–57
regression estimation; 回帰推定 : 91–92
release calendars; 発表日程 : 124–125
relevance; 適合性 : 8, 10–11, 15, 23, 26–27, 30, 49,
114
reliability; 信頼性 : 7, 9–11, 15, 20, 26, 28, 30, 59,
61, 65, 67–68, 78, 90–93, 114, 120, 123, 126
re-scaled weights; 再目盛付けウェイト : 86
residual group; 残余グループ : 46–47
resources; 資源 : 8, 10–11, 16, 21–23, 33, 39, 65, 69–70,
75, 77, 80, 82, 166
response burden; 回答者負担 : 11, 23, 53, 56, 61,
65–67, 74, 78
results orientation; 結果志向 : 16

re-weighting; ウェイト調整 : 78, 83, 86–88, 92–93

S

sample data; 標本データ : 31, 40, 88, 93, 118
sample of households; 世帯の標本 : 52
sample selection; 標本選出 : 50, 52, 63, 87
sample size; 標本の大きさ : 62–65, 67–68, 87–88, 91
sample survey; 標本調査 : 27, 28, 31, 50, 61, 65–67,
85, 89–90, 92–93, 115
sampling design; 標本設計 : 27–28, 34, 61–67,
85–88, 90–93
sampling error; 標本誤差 : 12, 27–28, 52, 79, 87, 115
sampling frame; 標本フレーム : 34, 49, 50–52, 61,
63–64, 66, 71, 78, 86–87
sampling unit; 抽出単位 : 61, 64, 67, 69–71
sampling variance; 標本分散 : 63, 65–66, 77
sampling weight; 抽出ウェイト : 65, 85–88, 92
SAS; 93, 119
seasonal adjustment; 季節調整 : 5, 28, 93,
101–102, 104–105
seasonal variation; 季節変動 : 101–102, 105
SEATS; 101–102, 104–105
self-administered questionnaire; 自己管理調査票 : 71,
73, 75
self-evaluation; 自己評価 : 17, 18
simple random sampling (SRS); 単純任意抽出 : 62
small area estimator; 小地域推定量 : 91
sources of error; 誤差の源泉 : 17, 28, 90
standard classifications; 標準分類 : 45–47, 49
standard error; 標準誤差 : 28, 65–66, 90, 92–93, 107
standards; 標準, 基準 : 10, 12–17, 92–93, 107
Stat.Fi; 6, 8, 36, 49, 126, 129
Stata; 93
StatFin; 129
statistical analysis; 統計的分析 : 66–67, 85, 90, 93–94
statistical chart; 統計図 : 96–98, 100
statistical graphics; 統計グラフ : 97, 100, 121

- statistical grouping of municipalities; 自治体の統計的グループ化 : 48
- statistical production process; 統計生産過程 : 3, 113
- statistical survey; 統計調査 : 3, 5, 14, 19, 29, 31–36, 49–51, 61, 67, 75, 77, 90, 94, 113–114
- statistical unit; 統計単位 : 8, 45, 52, 109–110
- Statistics Act; 統計法 : 6, 8, 33, 36–38, 42–44, 53, 56, 60, 74, 76, 123
- StatMarket; 128
- stratification; 階層化 : 51, 61–64, 66, 86–89, 91–92
- stratified sampling; 階層別抽出 : 50, 63, 66, 88
- student register; 学生レジスター : 55, 57
- SUDAAN; 96
- survey data; 調査データ : 66, 72, 83, 85–86, 89–93
- survey design; 調査設計 : 27, 61, 67–68
- Survey Laboratory; 調査実験室 : 32
- survey population; 調査母集団 : 51, 115
- system work; システム作業 : 33, 36
- T**
- table; 表 : 12, 77–78, 83, 114
- table format data; 表形式のデータ : 37, 42
- tables of magnitude; 大きさの表 : 39–40
- target population; 目標母集団 : 28, 31–33, 49–51, 59, 66–67, 78, 115
- telephone interview; 電話インタビュー : 59–71
- thematic map; テーマ別地図 : 97
- time series; 時系列 : 5, 27–29, 32, 93, 97, 101–106, 112, 115, 117
- timeliness; 適時性 : 8, 10–12, 14–15, 26, 28, 30, 50–52, 58–59, 115, 123
- total count statistics; 全数統計 : 56
- total data; 総計データ : 31, 40, 43, 118
- total quality management (TQM); 総合的品質管理 : 16
- total survey; 全体調査 : 32, 49–50, 93
- Trade Register; 貿易レジスター : 56
- TRAMO; 101–102, 104–105
- transparency; 透明性 : 8, 15, 18, 21, 100
- trend; トレンド : 82, 87, 99, 101–102, 104
- two-stage sampling; 2段階抽出 : 50
- U**
- unambiguity; あいまい性 : 46
- unbiasedness; 不偏性 : 91–92
- undercoverage; 過小カバレッジ : 28, 50–52, 58, 64, 69, 87
- unified data file; 統一データ・ファイル : 49, 117–119
- uniformity; 均一性 : 20, 107, 115
- unit non-response; 単位無回答 : 77, 83, 87, 91–93
- unit specificity; 単位詳細指定 : 59
- unit value index; 単位価格指数 : 107
- unit-level data; 単位レベルデータ : 53, 129
- usability; 有用性 : 3, 41, 118, 120–122
- use restrictions; 利用制約 : 116
- user groups; 利用者グループ : 34
- user manual; 利用者マニュアル : 34
- user needs; 利用者ニーズ : 59, 120
- user rights; 利用者の権利 : 38, 42
- V**
- value range check; 数値範囲チェック : 84
- variable transformations; 変数転換 : 34
- vehicle register; 車輛レジスター : 55–57
- weight construction; ウェイト構築 : 5, 83, 85, 87
- W**
- work instructions; 活動指示 : 34, 49, 69, 113–114

略号のリスト

AMS, Archive Formation Plan	NACE, Nomenclature Générale des Activités Economiques
ARGUS, Software for statistical confidentiality	dans les Communautés Européennes
ARIMA, Autoregression integrated moving average	(Statistical classification of economic activities
BSC, Balanced scorecard	in the European Community)
CALMAR, Calibration program (calibration on margins)	NSI, National Statistical Institute
CAPI, Computer Assisted Personal Interview	NUTS, Nomenclature des Unités Territoriales Statistiques
CASI, Computer Assisted Self Interview	(Nomenclature of territorial units for statistics)
CASI-A, Text on Screen and on Audio	OSF, Official Statistics of Finland
CASIC, Computer Assisted Survey Information Collection	PPS, Probabilities Proportional to Size
CASIIP, Computer Assisted Self Interviewing with Interviewer Present	PRODCOM, List of products of the European Community
CASI-V, Question Text on Screen	RAMON, Classifications Database of Eurostat
CATI, Computer Assisted Telephone Interview	SAS, A large intelligence and analysis software package
CAWI, Computer Assisted Web Interview	SRS, Simple random sampling
CLAN, Software for computation of point and standard error estimates in sample surveys	Stakes, National Research and Development Centre for Welfare and Health
CoP, (European statistics) Code of Practise	SUDAAN, Software for analysis of complex survey data
CPA, Classification of Products by Activity	SYS, Systematic sampling
DBM, Disk by Mail	TQM, Total quality management
Deff, design effect	TRAMO/SEATS, Software for seasonal adjustment
DESAP, Development of Self-Assessment Program	UN, United Nations
EC, European Community	UNSD, United Nations' Statistical Department
ECE, Economic Commission for Europe	ALLI,
EFQM, European Foundation for Quality Management	
ESA, European system of accounts	
EU, European Union	
Eurostat, (ESTAT) Statistical Office of the European Communities	
ISI, International Statistical Institute	
ISO, International Organization for Standardization	
MSE, Mean square error	

Documentation page

Published by
Statistics Finland

Date of publication
November 2002

Editors
Johanna Laiho
Leena Hietaniemi

Type of publication
Handbook

Commissioned by
Statistics Finland

Title of publication
Quality Guidelines for Official Statistics

Parts of publication

Abstract

The statistical production process forms the core of Statistics Finland's activity. The aim of the *Quality Guidelines Handbook* is to improve the usability of the skills and competence required in the designing and implementing of statistical systems by gathering the existing principles into common knowledge capital.

The *Quality Guidelines Handbook* is intended for all who are interested in the functioning of statistical systems, as well as for the users and producers of statistics. The Handbook outlines the framework within which the field of statistics operates in Finland and describes the relevant legislation, as well as current best methods and recommendations.

Keywords

Quality of statistics, statistical survey, production of statistics

Other information

Series (key title and no.) Handbooks 43b		ISSN 0355-2063	ISBN 952-467-073-9
Pages 132	Language English	Price EUR 30	Confidentiality Public
Distributed by Statistics Finland Työpajakatu 13, Helsinki 00022 STATISTICS FINLAND		Publisher Statistics Finland	

KÄSIKIRJOJA
 HANDBÖCKER
 HANDBOOKS

Nro 1	Koulutusluokitus 2008 Utbildningsklassificering 2008		2007
Nro 4	Toimialaluokitus TOL 2002 Påipalkaten		2002
	Toimialaluokitus TOL 2002 Liite 1 Hakemisto		2002
	Toimialaluokitus TOL 2002 Näringsgrensklassificeringen TOL 2002 Standard Industrial Classification TOL 2002 Liite 2 Tiivistelmä Bilaga 2 Sammanfattning Annex 2 Summary		2002
Nro 5	Sektorilukitus 2000 Sektorklassificeringen 2000 Classification of Sectors 2000		2000
Nro 8	Rahoitusvärdeluokitus 1998 Classification of financial assets and liabilities 1998		1998
Nro 10	Yhteisöjen tehtävien luokitus Julkisvirastojen ja voittoa tavoittelemattomien yhteisöjen tehtävien luokitus Uppgiftsklassificeringar för sammanslutningar Uppgiftsklassificeringar för offentliga sammanslutningar och icke-vinstsyftande sammanslutningar Classifications of the functions of government and non-profit institutions serving households	uusittu laitos förynyd upplaga revised edition	1988
Nro 11	Pääasiallisen toiminnan luokitus Pääasiallisen toimeentulo lähteen luokitus Klassificering av befolkningen efter huvudsaklig verksamhet Klassificering av befolkningen efter huvudsaklig verksamhet Classification of the Population by Type of Activity Classification of the Population by Main Source of Livelihood		1980
Nro 12	Valtiot ja muut 2004 Stater och länder 2004 Countries 2004		2004

Nro 14	Ammatilliluokitus 2001 Förskåpning		2001
	Ammatilliluokitus 2001 Lisä 1 Hakemisto, huhtikuu 2005		2005
	Ammatilliluokitus 2001 Yrjöluokiförteckning 2001 Classification of Occupation 2001 Lisä 3 Tiivistelmä Bilaga 3 Sammanfattning Annex 3 Summary		2001
Nro 16	Rakennusluokitus 1994 Byggnadsklassificering 1994 Classification of Buildings 1994		1994
Nro 17	Sosioekonomisen aseman luokitus 1989 Classification of Socio-economic Groups	uusi painos revised edition	1989
Nr 17b	Sosioekonomisk indelning 1989	förnyad upplaga	1989
Nro 18	Demografiset ja sosiaaliset perustilastiedot Iki, naimattomuus, siviilisyys, kieli, kansallisuus, uskonto Demografiska och sociala grundläggande Ålder, kön, civilstånd, språk, nationalitet, trosförhållande Demographic and Social Basic Classifications Age, Sex, Marital Status, Language, Nationality, Religion		1983
Nro 20	Suomen ympäristötiedot		1998
Nro 21	Aineellisten varojen luokitus Elinkein, varustevarat, muut aineelliset varat Klassificering av materiella tillgångar Fast kapital, lagerkapital, övriga materiella tillgångar Classifications of Tangible Assets Fixed Capital, Inventories, Other Tangible Assets		1985
Nro 22	Ikkilokitus Objekt i klassificeringen Åldersklassificeringar Bakgrunder för användning av åldersklassificeringar		1988
Nro 27	Riksbrottskatalog Brottsnomenklatur Crime nomenclature		1993
Nro 28	Kunnat ja kuntajohdot alustet 2007 Kommuner och kommunstyrelse förvaltnings 2007 Municipalities and Regional Divisions Based on Municipalities 2007		2007
Nro 30	Tilastokeskuksen ammatilliset opas		2006
Nro 30b	Guidelines on Professional Ethics		2006
Nr 31	Tuottajaindeksi 1995=100 Indeksin käyttäjän käsikirja Producer Price Index 1995=100 Handbook for Index Users		1993
Nro 32	Määräkkönnäkömuutosten indeksi 1995=100 Käyttäjän käsikirja		1993
Nro 33	Anstöttsindex 1995=100 Käyttäjän käsikirja The Index of Wage and Salary Earnings 1995=100 Handbook for Users		1993

Nro 33	<i>Ansietusindeksi 1990=100</i> <i>Käyttäjän käsikirja</i> <i>The Index of Wage and Salary Earnings 1990=100</i> <i>Handbook for Users</i>	1994
Nro 34	<i>Tuoteluokitus</i>	1995
Nro 35	<i>Väestöluokitus 2000</i> <i>Käsikirja</i>	2001
35b	<i>Folkförtegen 2000</i> <i>Handbok</i>	2001
35c	<i>Population Census 2000</i> <i>Handbook</i>	2001
Nro 36	<i>Sivillisten nimistöä</i> <i>Nomenclatur für civilna</i> <i>Nomenclature of civil cases</i>	2002
Nro 37	<i>Jäteluokitus</i>	2005
Nr 37b	<i>Guide to Waste Classification</i>	1999
Nro 38	<i>Vuoden 1990 väestöluokituksen</i> <i>otsonniston käsikirja</i>	1997
Nro 39	<i>Kuluttajahintaindeksi 1995=100</i> <i>Käyttäjän käsikirja</i> <i>Consumer Price Index 1995=100</i> <i>Handbook for Users</i>	1998
Nro 40	<i>Maa- ja metsätalouden luokitus</i> <i>Markensbruksklassificering</i> <i>Land Use Classification</i>	2000
Nro 41	<i>Julkisyhtiöiden teollisuusluokitus</i>	2001
Nro 42	<i>Rakennuskustannusindeksi 2000 = 100</i> <i>Käyttäjän käsikirja</i> <i>Building Cost Index 2000=100</i> <i>User's Handbook</i>	2001
Nro 43	<i>Lautas tilastoissa</i>	2007
Nr 43b	<i>Quality Guidelines for Official Statistics – 2nd Revised Edition</i> revised edition	2007
Nro 44	<i>Yksityisen kulutuksen käyttöarvo otoksen mukainen luokitus (COICOP)</i>	2002
Nro 45	<i>Use of Registers and Administrative Data Sources for Statistical Purposes</i>	2004

Statistics Finland

– Combining data and expertise into statistics and information services

varied, reliable data on the society and its development

broad and active co/operation networks in Finland and abroad

statistics production built on the work of 1,000 experts in different fields

head office in Helsinki, regional service offices in Turku, Tampere,

Seinäjoki and Oulu

Statistics Finland

Street address: Työpajankatu 13, Helsinki

Postal address: FI-00022 Statistics Finland

Puhelin: +358 9 17341

www.stat.fi

訳者論評—フィンランド統計局の統計品質活動にふれながら—

1. はじめに—高い評価を受けるフィンランド統計局

本資料は、フィンランド統計局の2007年のハンドブック *Quality Guidelines for Official Statistics, 2nd Revised Edition* の翻訳である。初版は2002年に発行されており、後述するが、この改訂2版ではかなりの部分に改訂がほどこされている。

さて、フィンランド統計局に注目した理由は幾つかある。第一に、本統計参考資料 No.102 で ESS の統計品質論をとりあげたが、ESS 中の各国で、統計品質論と実践がどう進められているかを見ることが、例えば、日本の参考にするためにも必要であった。第二に、ESS 中でも、例えばその Peer Review の第一ラウンド(2006~2008年)で高い評価を受けていたのがフィンランドである(表1参照)。主要国のウェブサイトの中で、フィンランド統計局のそれは、各統計資料の画面から容易に「品質記述」を参照でき、利用者本位の見地が徹底している点で、オーストラリア、カナダ統計局とともに訳者が注目していた。1999年に ISI を主催したフィンランド統計局には、見学ツアーに参加し、再度訪問したことがある。このときの見聞もあって『統計研究参考資料』No.63(2000)で「フィンランドにおけるレジスター・ベースの統計生産」をとりあげた。北欧に属していて、レジスター・ベースは定着していることも注目した理由である。

そこで、Peer Reviewにおけるフィンランド統計局をみると表1のとおりである。この評価を全面的に受け入れて、異なる評価者の比較を引用して良いかは留意すべきだが、これらの諸国・機関の中で◎が最も多い。

表1 ESSにおける主要国について同業者評価の項目別結果

原則	指標		フィン ランド	スウェ ーデン	イギ リス	フラ ンス	Euros tat
1	1.1	独立性の法規での規定	○	○	○	○	○
統計の 独立性	1.2	統計局長の政府機関内地位	◎	◎	◎	◎	○
	1.3	統計局長の独立性保持の権限と責任	◎	◎	◎	◎	○
	1.4	統計局長の方法・手続き・発表に関する決定権	◎	○	○	◎	○
	1.5	統計政策の発行と定期的進捗度報告	◎	◎	◎	◎	◎
	1.6	統計発表の区分と独立性	◎	◎	◎	◎	◎
	1.7	批判と誤用へのコメント権	◎	◎	◎	◎	◎
2	2.1	申告義務の法的規定	◎	◎	◎	◎	◎
データ 収集権 限	2.2	行政記録の統計目的への利用	◎	◎	◎	◎	◎
	2.3	法的回答強制権限	◎	◎	○	◎	—
3	3.1	ヨーロッパのニーズに見合う十分な資源	◎	◎	○	◎	○

資源の 十分性	3.2	ヨーロッパ統計のニーズにあう視角，詳細，費用	—	—	—	—	—
	3.3	ヨーロッパ統計の新しい要請の評価手続きの存在	—	—	—	—	—
	3.4	統計活動停止可能性を評価する手続きの存在	—	—	—	—	—
4	4.1	ESS品質基準による定期的監視	◎	◎	◎	○	○
品質責 任	4.2	統計の収集・処理・配布の品質監視の実施	◎	○	◎	◎	△
	4.3	品質検討の過程の実施	○	○	◎	◎	○
	4.4	品質ガイドラインと職員訓練，公衆への周知	◎	◎	◎	○	◎
	4.5	外部専門家を使った主要統計生産物の定期的 で徹底したレビュー	○	○	◎	○	○
5	5.1	法規での統計的秘匿性	◎	◎	◎	◎	◎
統計的 秘匿性	5.2	職員の任命時の署名	◎	◎	◎	△	◎
	5.3	自覚的違反への罰則	◎	◎	◎	◎	◎
	5.4	指示とガイドラインの用意と公衆への周知	◎	◎	◎	○	◎
	5.5	統計データベースの安全性とintegrity保持の 物的・技術的装備	◎	◎	◎	◎	◎
	5.6	研究目的でのマイクロデータの外部利用者への 厳密な協定	◎	◎	◎	◎	◎
6	6.1	統計の客観的作成	◎	○	◎	◎	◎
不偏性 と客観 性	6.2	出所と技法の選択の通知	◎	◎	◎	◎	◎
	6.3	公表統計の速時の誤差修正	◎	◎	◎	◎	◎
	6.4	使用方法と手続きの公衆の入手可能性	○	◎	○	○	◎
	6.5	統計発表時の事前告知	◎	◎	◎	◎	○
	6.6	公表統計への全利用者の同時的アクセス可能 性，外部利用者の特別優先の制限と公開	○	◎	○	○	△
	6.7	新聞記者発表の客観性と無党派性	◎	◎	◎	◎	○
9	9.1	統計の範囲の絶対的必要への限定		—			
回答者 負担の 軽減	9.2	回答負担は可能な限り広げる		○			
	9.3	情報は回答者が可能な限り直ちに入手でき， 回答で電子的手段の利用		○			
	9.4	詳細が入手できない場合の最善の推定値・近 似値の受け入れ		—▽			
	9.5	行政的出所の利用		○			
	9.6	調査重複を避けるためのデータシェアリング		—△			
15.	15.1	適切な解釈と意味ある比較を促進する形式	◎	◎	○	◎	○
アクセ	15.2	現代的ICT技術と伝統的紙印刷での配布	◎	◎	◎	◎	○

ス可能 性と明 瞭性	15.3	顧客の注文による統計の提供	○	○	○	◎	◎
	15.4	研究目的のマイクロデータの使用	◎	◎	◎	◎	◎
	15.5	標準化されたメタデータシステムでの文書化	◎	○	○	○	○
	15.6	利用者への方法論と品質の告知	◎	○	○	○	○

注 1. 原則とはヨーロッパ統計実践規約にうたわれている。
2. ◎完全に合致 (fully met), ほぼ合致 (largely met), △部分的に合致 (partly met), ×合致せず (not met)

2. フィンランドにおける統計活動と統計局の歴史と現在

そこで、このガイドラインの改訂第2版では削除されており、2002年の第1版に付録として収録されていたフィンランドとフィンランド統計局の紹介を引用してみよう。もちろん、局の構成その他すべては2002年時点のものであることに留意をいただきたい。

(1) 「付録 フィンランドとフィンランド統計局概要

FINLAND	
(independence 6th December 1917)	
Land area	Population trends (2001)
Total 338 145 km ²	Population growth 2.7‰ (Per 1 000 mean population)
Cultivated 6%	Excess of births 1.5‰ (Per 1 000 mean population)
Forest 68%	
Water 10%	
Greatest length/width	Life expectancy at birth (2001)
From south to north (Hanko-Utsjoki) 1 157 km	Males 74.6 years
From west to east (Närpiö-Illomantsi) 542 km	Females 81.5 years
Water boundaries	Languages (2001)
At west Gulf of Bothnia	Finnish 92.3%
At south Gulf of Finland	Swedish 5.6%
	Sámi 0.03%
	Russian 0.6%
State boundaries	Membership of
At north with Norway 727 km	UN 1955
At north-west with Sweden 586 km	EFTA 1986
At east with Russia 1 269 km	EU 1995
Capital	Mobile telephone connections (2001)
Helsinki, population 559 718 (31 January 2001)	80/100 inhabitants
Permanent population	Internet connections (2001, July)
5 194 901 (31 January 2001)	17/100 inhabitants
Population density	Unemployment rate (2001)
17.0/km ²	9.1%

(2) フィンランドの統計の歴史 フィンランドの統計組織の出発点は、18世紀半ばの Tabellverket (Tables Office) がスウェーデンに人口統計を保管する目的で設置されたときにさかのぼる。これは、世界で最長の継続的な人口統計の開始を印した。

最初の独立したフィンランドの統計局は、フィンランドがロシア帝国の自治大公国であった1865年に創設された。最初の統計年報は1879年に出版された。統計局の活動はフィンランドが1917年に独立するまで途絶えることなく継続した。1917年の後、社会問題や経済的傾向といった分野で阿他rしい統計が作成された。国民勘定計算は、国連の勧告が採択された直後の1947年

に導入された。

ヨーロッパ連合への加盟は、国際機関内で行われる活動量を大きく増加させた。重要な目標は、統計の国際的比較可能性を改善し、国際統計の利用を促進することである。EUの統計機関であるEurostatとは特別に密接な接触がある。特にフィンランドの近隣地域では、国際競技はますます重要になっている。

1748年 スウェーデンに製表局設立。

1749年 第1回フィンランド人口統計が完成。

1865年 フィンランド統計局 (Statistical Office of Finland) が1865年10月4日の創設される。

1879年 フィンランドの最初の統計年報が発行される。

1948年 粗国民所得の計算がはじまる。

1971年 フィンランド統計局 (Statistical Office of Finland) がフィンランド統計局 (Statistics Finland)に改称される。

1990年 世界で2番目の国として、フィンランドで最初のレジスター・ベースの人口センサスが実施される。

1995年 EU加盟が、フィンランド統計における国際統計要素を増大させる。

1995年 ヨーロッパで最初の統計機関として、フィンランド統計局は2月にインターネットサービスを開始する。

(3) フィンランドの政府統計とフィンランド統計局 30ほどの機関が政府統計を作成している。それらとして、フィンランド統計局に加えて、農森業省、フィンランド地方政府協会、ヘルシンキ市情報管理センター、関税局、福祉健康研究開発センター、社会保障研究所、市がある。

フィンランド統計局は政府統計の4分の3に責任を負っている。また、他の機関の統計活動を監督し調整している。この部門は全体として統計法に支配されている。統計法はデータの収集と統計作成の一般的原則、開示と秘匿性の義務を規定している。統計情報は広くメディアに配布されている、すなわち、毎年、印刷メディアはフィンランド統計局が発表する資料に基づいて総計で3,000の新しい報告と記事を発表している。

統計局は10の単位にわかれており、そのうちの3つは、主として対内と対外のサービスに関わり、6つが一般的統計単位である。

ー人口統計課は、人口、教育、司法と犯罪、選挙および住居と地方の就業についての統計を作成する

ー社会統計課は、局の標本調査と面接活動に責任を持ち、就業、所得と富、消費と生活条件の側面に関する統計を作成する。それは更に、時間利用と文化・娯楽に関する統計を生産する。

ー物価・賃金統計課は、物価と賃金および労働争議に関する統計に責任を持つ。

ービジネス統計は2つの単位、景気動向 (Business Trends) 課とビジネス構造 (Business Structure) から成り、ビジネスからデータを収集し、その資料を様々な産業を叙述する統計に仕上げる。

一経済統計課は、いくつもの基本統計からのデータを結合して、国民勘定体系をつくり、また総生産と金融市場に関する種々の主要指標を発行する。

フィンランド統計局は、そのデータの大半を様々なレジスターのみならず、企業、世帯および個人から直接的に獲得する。行政記録の使用は増加しており、これはまた回答者負担をへ減らしている。フィンランド統計局は、国の異なる部分に居る124名の調査員をふくめて、1000任を超える職員を持つ。調査員は、電話あるいは面接による調査データの収集の際に、ラップトップと最新の自動データ移送法を使う。統計研究・開発単位は、統計的と方法論的専門の異なる分野の専門家15名を雇用している。この集中化した単位に加えて、統計方法を専門にする多数の雇用者が統計生産単位で活動しており、統計方法のスタッフの総数は40人になる。研究部長(Director of Research)に加えて、フィンランドあるいは外国からの専門家が、短期の研究員協定の下に招かれて局で仕事をしてきた。科学的助言委員会(The Scientific Advisory Board)が研究活動を監督し、将来の研究活動の主な方向を引き出す助けをする。幾つかのフィンランドの大学と研究機関が委員会に出席している。

(4) フィンランドの統計学 国の20の大学に13の統計学部門がある。統計学部門は、社会科学の部あるいは数学とコンピュータ・サイエンスの学部として活動している。多くの他の学部も統計学の基礎的教育プログラムを持っている。

統計教育の1例は、Jyvaskyla大学の統計部門の統計システムにおけるマスターのプログラムにある。プログラムの主な焦点は、大きなデータ集合の収集、分析、管理と配布におかれている。統計課学の見地からは。プログラムは、標本調査理論と複雑な調査の分析をふくむ調査方法に集中している。

諸大学での統計学の主な領域には、ベイジアン統計、バイオ統計学、人口統計学、エコノメトリックス、伝染病、一般化した線形モデル、ライフ・ヒストリー分析、線形モデル、数理統計学、多変量解析、ノンパラメトリック統計学、標本抽出、統計的コンピューティング、空間過程、品質改善のための統計方法、農林科学の統計学、確率過程、調査方法、時系列分析、がある。

他の研究機関とその統計的研究の主な分野は、農業研究所(実験計画)、ヘルシンキ市情報管理センター(都市統計)、森林研究所(森林測定のための統計的方法)、国立健康研究所(バイオ統計学、伝染病)、フィンランド経済研究所(エコノメトリックス)とフィンランド統計局(標本調査、調査分析)、である。

1920年に創設されたフィンランド統計学会は、非営利の学問的な専門家の学会をなしおり、その目的は、統計理論の発展とその応用を促進することと、統計と研究の異なる分野の統計家と専門家を一堂に会させることである。(約500人の)会員は、フィンランドの大学と統計局や研究機関の統計学部門で働いている統計家からなる。」

(5) フィンランド統計局の本部と地方サービス事務所 フィンランド統計局は以下の地方サービス事務所を持つ。以上は、引用による紹介であった。



Statistics Finland's head office and regional service offices

3. フィンランド統計局における品質活動

今日、各国の統計品質活動の基礎になる理論・規定は、①基礎的規範等―②品質評価枠組み―③品質構成要素―④品質表示の在り方であり、⑤これに基づく実際の活動と、⑥自己評価―同業者評価等を経て、次の改善に進む。品質ガイドラインなりハンドブックは、一般的には、これら①～④を説明し、さらには⑤あるいは⑥に関する叙述をふくむものである。これをフィンランド統計局についてみる。

3.1 フィンランド統計の品質各レベルと品質表示

ガイドラインの2007年版は、2006年いっばいのESSにおける品質論の諸装置をフィンランド統計局が受け入れて書かれているので、ESSの品質論とミックスした形になっている。

(1) **基礎的規範** 訳出したガイドラインの第1章に示されているが、国内の2004年統計法、専門的倫理（倫理原則のガイドライン<2006>）、品質宣言<2001>、ESSの国と共同体統計機関の独立、integrity および説明責任に関する勧告<2005>、ヨーロッパ実践規約<2005>、国連の政府統計の基本原則<1994>、ISIの専門家の倫理宣言<1985>等に基づいている。

(2) **品質評価枠組み** ガイドラインは、1.の規範の1.1.3でESSの実践規約をとりあげ、フィンランドもまたこれに従うことを示している。同時に、1.2.1では、一般の総合的品質管理を紹介し、品質計画（quality planning）、品質保証/検査（quality assurance/verification）、品質改善（quality improvement）という組織による品質保証のための手段におよぶことも指摘している。

(3) 品質評価と基準

ガイドラインは、実践規約にも基づきながら、品質評価は、品質記述、品質監査、自己評価、等があるという。フィンランドは、国連の政府統計の基本原則と、ヨーロッパ統計実践規約に基づいて、フィンランド政府統計審議会は、2006年にフィンランド政府統計として受け入れられる統計の基準を具体化した。

本訳出資料の1.2.3に示されている。以下のとおりである。

OSFの基本的基準

OSFの基本的基準1:

この統計は、フィンランド統計局が管理するOSFのリストに含まれる機関等（agency or institution）が生産する。

OSFの基本的基準2:

発行機関は、統計の内容とデータの正しさ（correctness）に責任を持ち、したがって、当該統計の生産を継続的に保証する。OSF統計はいかなる個人の名前で公表してはいけない。

OSFの基本的基準3:

OSF統計はフィンランド政府統計の諮問委員会によって承認された更新された品質記述をふくむ。

OSFの品質基準

- 1. 適合性：** OSF統計は、それが表わすトピックについて社会のニーズに対して不可欠で全国的に包括的なデータをふくむ。
- 2. 正確性と信頼性（Accuracy and reliability）：** OSF統計は、それが関わる現象の状況と変化を正確かつ信頼性をもって記述する。統計の編集において適用される方法は、明確に報告されるべきであり、統計の信頼性に関係がありうるすべての事実は、分析され報告されるべきである。OSF統計でおそらく探り当てられた誤差を修正する手段は、フィンランド政府統計の諮問委員会が承認した勧告を遵守すべきである。
- 3. 適時性と迅速性（Timeliness and promptness）：** OSF統計のデータは、可能なだけ現在のなものでなければならず、その発表時点は、前もって周知され、フィンランドの政府統計の諮問委員会が発行した支持によって定められた手続きにそって、その発表者が保持している発表日程に発表されなければならない。
- 4. 整合性、一貫性および比較可能性（Coherence, consistency and comparability）：** OSF統計は他のOSF統計と論理的に一貫しており、時間と地域単位の間で可能な限り比較可能でなければならない。OSF統計は、基本的に一般的で、定められた国際的概念や統計分類を使用しなければならない。統計の一貫性や比較可能性に貢献する諸要因は、文書化され、統計に明細に示されるべきである。異なる定義を持つデータ集合は、異なる概念によって明確化されるべきである。
- 5. アクセス可能性と明瞭性：** OSF統計は、より明瞭で、透明で、理解可能な形で示され、適切で便宜的な仕方で発表されるべきであり、その利用者への指示とともに、それらを支持するメタデータが、普通にアクセス可能でなければならない。

このガイドラインは、研究書ではないので2007年段階でのフィンランドの品質要素を結論的に示している。2002年版では、フィンランド政府統計の品質基準として、「1. 統計情報の適合性、2. データの正しさと正確性、3. データの適時性と迅速性、4. 定時性、すなわち、予定通りの配布、5. データのアクセス可能性と透明性/明瞭性、6. 統計の比較可能性、7. データの整合性と一貫性/明瞭性、8. 文書化」を掲げていた。これが、本ガイドラインでは、実践規約を経たESSの品質基準を一方に前提しながら、ESSの品質基準を取り入れ、さらに2006年のフィンランド基準を経て、ESSの基準とほぼ同じにいたっている。違いは、ESSの（実践規約では

正確性と信頼性という項目を立てた後に、品質報告関係書ではしばって)「正確性」に対して「正確性と信頼性」、ESSの「適時性と定時性」に対して「適時性と迅速性」になっている点である。

とはいえ、次項【3.2の(1),(2)】に見るように、ウェブサイトでの各統計生産物の品質説明の項目は、上記(2007年版ガイドライン)の5項目とは異なっただけである。すなわち、(i)基準2:「正確性と信頼性」、に対して「正しさと正確性」、(ii)基準4:「整合性、一貫性および比較可能性」に対して「比較可能性」と「整合性と一貫性/均一性」に分かれている。

これに関する説明文書はない。品質構成要素の概念的展開としばりこみでは、なお未整理・不整合が残っているというべきか。

3.2 品質表示

品質表示は、報告書において、またウェブサイトにおいて行われるが、フィンランドの場合には、上記の品質基準についてとともに、方法論の説明も含むべきことが指摘されている。

(1)統計品質報告書 品質報告書に何が盛り込まれるべきかに関しては、ガイドラインの1.2.3の表1.1に示されている。基準も5項目に統計調査の方法論的記述がプラスされた6項目となっている。

(2) ウェブサイトでの提示 では、トップから以下のように進む

①トップページ

ホーム	統計	メタデータ	データ収集	生産物とサービス	組織	ニュース
-----	----	-------	-------	----------	----	------

②統計

調査名	ニュース	今後の発表	発表	レビュー	表	図
新聞向け発表	説明	品質説明	方法の説明	概念と定義	分類	一層の情報

③品質説明

1	統計情報の適合性
2	統計調査の方法の説明
3	データの正しさ(correctness)と正確性 (accuracy)
4	データ公表の適時性と迅速性
5	データのアクセス可能性と透明性/明瞭性
6	統計の比較可能性
7	整合性と一貫性/均一性

④メタデータ

統計の説明	概念と定義	分類	レジスターの説明	データ保護	統計法規
国家統計業務	一層の情報	データ収集	統計倫理の原則	政府統計の品質ガイドライン	

(3) ウェブ表示の具体例として労働力調査をみよう。月次の報告のそれぞれにこの記述が与えられている。2008年1月分では「(1)統計情報の適合性、(2)統計調査の方法の説明、(3)データの正しさと正確性、(4)データ公表の速やかさと適時性、(5)アクセス可能性とデータの透明性/明瞭性、(6)統計の比較可能性、(7)整合性と一貫性/均一性」にわたって8ページの説明がある。

(2)は、例えば、適合性や標準誤差等の一般的説明をもふくんでいる。(3)データの正確性では、無回答、測定誤差、と標準誤差によるとしながら、主として労働力調査にそくした標準誤差の説明にあてられている。無回答は(2)統計調査の説明において、「フィンランド統計局の調査員は、2007年に、コンピュータ支援の電話インタビューによって約115,000人について行われた。この調査の無回答率は平均して20.2%であった」と言い、無回答はいわゆるウエイト・カリブレーションによって修正されていると説明している。各説明の中に、より詳細な説明文書へのリンクがあり、また質問先のメール・アドレスと電話が掲載されている。(5)アクセス可能性の項目では、いわゆるマイクロデータを研究目的で特別な許可のもとに開示されることも伝えている。

4. 『政府統計の品質ガイドライン第2版』へのコメント

第一に、全体構成では、1章で全般的作業枠組み、として統計品質に関わる総論的記述の後、第2章で、全体の2/3のボリュームをさいて、統計調査と統計生産として、調査の前提、と調査の諸問題、データの提示(統計表と図表論)、時系列・季節調整、指数までをとりあげ、第3章で文書化と保存、4章で統計利用、出版、顧客サービスの促進、という構成をとっている。

第二に、各章と節では、視角と範囲—原則—ガイドラインにわけて説明があり、文献としてヨーロッパおよび国際基準、フィンランドの国内法・文書、フィンランドのガイドライン、追加的情報、が示されている。

以上の章・節構成と章・節内の叙述スタイルは、カナダ統計局のガイドラインを参考にしていられると思われる。カナダのガイドラインとともに、統計調査関連に多くを当てている点は、イギリスの国の実践規約(National Code of Practice)のシリーズとして、「原則の声明(Statement of Principles)」、「品質管理のプロトコール(Protocol on Quality Management)」、「統計的品質測定ガイドライン」(*Guideline for measuring statistical quality, version 3.1, 2007*)をそれぞれ発行し、トピックに特化しているものと対照的である。

第三に、標本調査の設計、無回答の処理、エディティング、補定等、統計生産に関わる従来のトピックスと現在のトピックスの主要なものがある。短い、第3章の文書化と蔵書化、第4章の統計の利用、出版、顧客サービスなどは、重要な問題でありながら、他の国においては、脇におかれがちなトピックスを、ルーティン的活動に組み入れているフィンランドらしく、とりあげているのである。

日本での統計関係論議で欠けている点を憂いている立場からみれば、第2章の各トピックスはそれぞれ統計関係者は常識的に把握しておくべき点であり、統計データの表示(統計表・図論)は適切な内容であり、特に、行政記録とレジスター、標本調査、調査票の設計、文書化、文書保存、も注目すべきである。

これら全体によって、統計活動における統計品質への対応の方向の概略は、特に掲載されている文献等をふくめて、フィンランド政府統計の関係者と公衆に対して与えているといえる。

第四に、統計に品質論とその実践に関しては、その多くをESSの品質論、フィンランド自体の他の文書に委ねていることもあつてか、詳しくはない。3.1の(3)で述べたが、統計品質の構成

要素についての説明は一貫してはいない点がある。

第五に、2002年版と比較してみよう。

- ①第1章、特に品質構成要素の採用と分類およびその概念的説明は大幅に書き換えられた。表1.1は新たなものである。これは2002年の後、国際的論議の中で品質評価枠組みがより明確化され、品質論議の中核的位置にあるESSのヨーロッパ統計実践規約が2005年に定められ、フィンランド統計局の活動が、かなりこれを取り入れるに至ったことがある。叙述のスタイルも、例えば、旧版の第2章：統計調査では、調査過程の説明自体で終わっている部分があった（これは、統計調査過程の説明としては、それなりのものであった）が、2007年版では、それらの過程が統計品質にどのような影響を与えるかに論述が及んだ。
- ②第1章のデータ保護8ページは新設であり、第2章の「標本調査」と、「エディティングと補定」ではページ数を増やして書き換えがある。旧版の「行政レジスター」が「行政記録とレジスター」に変えられた。
- ③一方で、第4章は約半分に縮小され、節構成も変更された。
第六に、弱点に関して少々コメントする。
- ① カナダ統計局型の品質ガイドラインとイギリス型のガイドラインの選択の問題がある。もちろん、ガイドラインと関連する文献との脈絡にもよって、全体として論議や指針が、どう整備・配置されているかによる。とはいえ、フィンランド型は、全体として統計調査論の傾斜してしまっているように思える。
- ②再三繰り返すことになってしまうが、品質構成要素一基準をめぐって、説明の一部に一貫性がない部分がある。
- ③2.6-2.10節には、叙述の内容的な重複がある。これは、2002年版を受け継ぎながら新しい叙述が組み込まれたためだろうか。

5. 本資料の仮訳と訳者論評は、伊藤陽一（日本統計研究所客員研究員）が担当した。なおこのガイドラインの日本語での翻訳出版を許諾されたフィンランド統計局に感謝したい。翻訳は基本的には09年9月上旬に終了していたが、翻訳許可を受け取るのに時間を要してしまった。訳者論評も09年9月時点のままである。

この訳出の過程で、幾つかの誤植等については正した。

追記となるが、フィンランド統計局は、政府統計における品質のヨーロッパ会議を2010年5月17-18日に開催しようとしている。主催国からの発表参加は多くなるので、ここで、フィンランドにおける統計品質論の実践の成果がより豊かに示されることが期待される。

統計研究参考資料(最近刊行分)

号数	タイトル	刊行年月日
89	統計の品質(3)—国際統計機関における統計の品質	2005. 09. 30
90	韓国2000年産業別購買力平価の推計	2005. 10. 03
91	イギリス国家統計局(ONS) 世帯サテライト勘定(試験的)方法論	2005. 12. 25
92	ジェンダー予算・人々中心の予算(1)—翻訳と関連論文	2006. 03. 25
93	統計の品質(4)—IMF・品質サイトとQ2004を中心に	2006.07. 25
94	中国国民経済計算体系2002	2006. 08. 01
95	韓国「統計法」改正	2007. 02. 01
96	日中韓2000年産業別購買力平価の推計	2007. 04. 01
97	統計の品質論(5)—Q2006と2006サテライト会議から(翻訳と関係論文)	2007. .05. 31
98	Eurostat:世帯生産と消費—世帯サテライト勘定の方法と提案	2008. 01. 31
99	中国国家統計局「都市家計調査」の家計収支項目分類の変遷に関する研究	2008. 10. 20
100	中国産業連関表のデフレーターと実質化	2008. 11. 01
101	ロシア人口センサスの調査環境	2009. 01. 31
102	統計の品質論(6):論文と翻訳—ESSの統計品質論と実践	2009.08. 31
103	第18回国際労働統計家会議における「労働時間測定決議」	2009. 09. 05
104	ビジネス・レジスター勧告マニュアル	2009. 09. 06

統計研究参考資料 No. 105

統計の品質(7)
翻訳 フィンランド統計局
政府統計の品質ガイドライン

2010年1月20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, 2326
Fax 042-783-2332
Email jsri@mt.tama.hosei.ac.jp

発行人 森 博美