

統計研究参考資料

No. 61

「統計の品質」をめぐって—翻訳と論文

1999年12月

法政大学日本統計研究所

The Japan Statistics Research Institute

Hosei University

「統計の品質」をめぐる一翻訳と論文

目次

翻訳

1. Eurostat(1998) 統計の品質についての標準報告書
Eurostat/A4/Quality/98/General/Standard Report/(Proposal) 1
2. Eurostat(1998) 統計における品質の定義
Eurostat/A4/Quality/98/General/Definition 16
3. Eurostat(1998) 統計の品質に関する用語説明
Eurostat/A4/Quality/98/General/Glossary 24
4. Raoul Depoutot(1998) 統計規範の品質 33
5. Eurostat, Sub-Group(1998) 規範の品質：最終報告 42
6. Raoul Depoutot(1998) 国際統計機関の統計生産物の品質 68

論文

7. 伊藤陽一(1999) 統計の品質 (統計の真実性と関連諸要因)
—最近の国際的論議を参考に— 80
- 訳者 (執筆者) あとがき 96

1 品質の標準報告書

Eurostat

目次

1.	適合性	2
2.	正確性	2
2.1	標本誤差	3
2.1.1	確率標本抽出	3
2.1.2	非確率的標本抽出	4
2.2	非標本誤差	5
2.2.1	枠組の誤差（フレーム・エラー）	5
2.2.2	測定誤差	6
2.2.3	作業誤差	7
2.2.4	無回答誤差	7
2.2.5	モデル仮定誤差	8
3.	速報性と時間厳守性	9
4.	アクセス可能性と明確性	10
5.	比較可能性	10
5.1	場所的比較可能性	10
5.2	時間的比較可能性	11
6.	一貫性	11
6.1	予備的統計と最終統計との間の一貫性	11
6.2	年次統計と短期統計の一貫性	12
6.3	同じ社会・経済的領域における統計の一貫性	12
6.4	統計と国民勘定体系との比較	13
7.	費用と予算	13
7.1	回答者についての費用	13
7.2	統計局の費用と予算	13

品質の標準的報告書

はじめに

この論文は統計の品質報告に関する提案をするものである。本論文は、品質についての報告書が収録すべき主題を述べ、Eurostat が、加盟国(MS:Member States)に提出することを期待する特別な情報をリストしている。この報告は、参加国が Eurostat の基準からどう乖離しているかを明確にする。以下の報告の幾つかの背景については、記録 Eurostat/A4/Quality/98/General/Definition” と“ Eurostat/A4/Quality/98/General/Glossary” でふれている。

1. 適合性(Relevance)

生産される統計が利用者にとって適合的でありつづけていることを明らかにするために、参加国は、その主な利用者の意見のサーベイを定期的に行うべきである。

多年次をベースにした報告であるべきもの：
利用者についての説明、利用者のニーズの起源と満足、そして利用者にとっての統計の適合性をふくめた要約。

2. 正確性(Accuracy)

全体的な正確性は利用者の主なニーズのひとつである。この情報は、統計とともに信頼区間を公表して与えられるべきである。信頼区間は、誤差の修正と依然として残される誤差の影響のすべて、と確率的標本抽出による分散、すなわち、以下のものを考慮に入れるべきである。

- 推定量の偏り、カットオフの境界線、無回答の帰属方法と、(再ウエイトづけ、測定誤差他による)他の系統的偏り；
- 確率的標本抽出であることによる偏り、分類の誤り、測定誤差、処理誤差、無回答に対する帰属と再ウエイトづけ、統計的モデル (例えば、季節調整)。

定義によって、分散は推定量と結びついている。修正はしばしば、統計的技法を用いて偏りを減らすために行われる。それらは、推定量にふくまれるかも知れない。データの収集時あるいはデータの入力時に生じる他の誤差は、エディティングの過程によっても、推

定量によっても修正はされないが、信頼区間において考慮されるべきである。

加盟国間で比較可能な信頼区間を得るためには、同じ構成要因が分散の推定過程に含まれるべきである。最初のステップで、われわれは、これらの構成要因を、(使われる推定量と結びついて) 確率抽出に由来する分散、分類の誤り、そして無回答に由来する補定(imputation)、とに限定すべきことを提案する。たとえ、幾つかの非標本誤差は除外できず、信頼区間に寄与するとしてもである。

加盟国は異なる種類の誤差についての情報を提供すべきである。

2.1 標本誤差

2.1.1 確率的標本抽出

ランダムな確率標本から推定するものとし、そして、他のすべての誤差の源泉(フレーム誤差、測定誤差、その他)を無視するなら、一つの推定値の正確性の尺度は、平均平方誤差の平方根である。推定量の平方平均誤差は、推定量の分散プラス偏差の平方に等しい。標本抽出と測定方法が不偏推定量をもたらすときにはいつも、平方平均誤差は推定量の分散であることは明らかである。

確率的標本抽出は、登録簿を使用できて、正確性を統計単位に関連する変数について推定できるところで使われる。幾つもの段階での標本抽出では、多段階での設計のうちの一段階で単位を取り出すために非確率的標本抽出を使うことがありうる。

例えば、第一段階の単位が店舗であり、第二段階の単位が店舗内のある商品であるという消費者物価に関するサーベイの場合がしばしばそうである。

2.1.1.1 偏り

総計の推定量はしばしば不偏である。例えば、Horvitz-Thompson 推定量を用いて作られる推定値は不偏である。しかし推定値はひんばんに、比率あるいは回帰推定量を使って計算されるが、それらは偏りを持っているが一貫性をもっている。ときとして、よりロバストであるが、偏りを持つ推定量が使われる。

多年にわたる基準で報告されるべきこと：
測定可能な場合には、推定方法による年次的偏り。
四半期および月次の統計についての偏りの年次の中位数および範囲。

2.1.1.2 分散

標準的推定法(例えば、Horvitz-Thompson 推定量あるいは標本推定量)を用いた標準的

標本設計のほとんど（例えば、単純抽出あるいは階層別抽出）については、推定量と結びついた分散や標準誤差を推定のための標準的公式がある。選出手続きのある段階が確率的標本でない場合には、ある仮定の下に分散を概算することは可能である。しかし、これはしばしば推定されるべき分散が過小推定される。

報告されるべきこと	統計の頻度		
	月(各四半期)	年ごと	数年ごと
以下の標準誤差			
総計（指数ごとのレベルに関して）	Y	Y	Y
成長率 連続的期間	Y	Y	
連続した年の同じ時期	Y		
Eurostat への提出の頻度	月次（各四半期）	年ごと	数年ごと

分散を推定する調和された方法が必要である。

この推定は無回答（項目と単位の）の補定に由来する変異を考慮しなければならない。この推定はまた分類の誤りの影響を考えなければならない。

この情報は、Eur15 レベルで統計を合計することによって、信頼できる情報を配布することができるかどうかを判断する上で不可欠である。たとえ、加盟国レベルへの配布でない場合でも不可欠なのである。

2.1.2 非確率的抽出

ある事情の下では、非確率的標本抽出が使われる。国家統計局は関連する特性（例えば、生産、就業、離職…）についての高いカバレッジを基礎にして、専門家による標本（また分別ある<judicious>と呼ばれる）を使うこともあろう。

これらの事情の下では、推定値の正確性について客観的評価をうることは不可能である。それにもかかわらず、品質についての何らかの粗い正確性指標は、感度分析(sensitivity analyses)を使って企画することができよう。単位が、年付加価値といった構造的補助変数にしたがって確実性を持って選出されるときには、選出されなかった企業の影響を推定するために計量モデルを使って、もっと高度な指標をつくり出すことができよう。

報告すべきこと	報告頻度
年次のカバー率	年ごと
正確性指標	月ごと
それらの信頼性指標をうるために使った方法	数年ごと

2. 2 非標本誤差

2.1.3 フレーム誤差

フレーム誤差は、過大カバレッジ、過小カバレッジおよび分類の誤りである。

標本抽出枠組みへの新しい単位の速やかな算入と死んだ単位の除外はフレーム誤差を限定するために必要である。

過大カバレッジは、実際には範囲外である単位を誤って分類するとか、実際には存在しない単位に関わる。

報告すべきこと	提出頻度
母集団の主な特性ごとに過大カバレッジを分割しての年次の率	年ごと
それらの数字を獲得するために使われた方法	数年ごと

過小カバレッジは、本当に生じたか合同したかのいずれかによる（新しい）単位をフレームに算入しないことと、誤って分類された単位に関わる。

それはまた、加盟国がカットオフした標本を故意に利用することからくるかもしれない。この場合には、偏りを評価するべきである。

報告されるべきもの	提出頻度
母集団の主な特性ごとに分割した過小カバレッジの年次の率	年ごと
カットオフした標本に由来する年次の偏り	年ごと
これらの数字を獲得するために使われた方法	数年ごと

分類の誤りは、目標とする母集団に属するするものを誤って分類することである。これらの分類の誤りは、経済現象あるいは特殊なケースを反映しているかもしれず、修正されるべきものか、あるいは修正されるべきでないものである。分類の誤りを考慮するためのルールは何だろうか？ これらの修正は統計にどのような影響を与えるのか？ それは統計が修正されるときにおこなわれるのか？

報告されるべきもの	提出頻度
分類の誤りの処理に関する方法論的ノート	数年おき／変更があったとき
分類の誤りの年次の率	年ごと
これらの数字を得るために使われた方法	数年おき

分類の誤りのあった単位は、幾多の層にわたる領域についての推定をもたらす。対応する変異は幾多の当初の層にわたる考察の対象になった領域のすべての単位の変異を考慮に

いれるべきである。

2.1.4 測定誤差

測定誤差はデータ収集時に起こる誤差である。測定誤差の源泉は多くあり、サーベいの装置（書式あるいは調査票）、回答者、情報システム、データ収集の様式（直接的面接、電話インタビュー、自己管理の郵送サーベイ、日記によるサーベイ、行政記録、直接観察、電子的観察）、および面接者等がある。

報告誤差は、特に EDI (electronic data interchange:電子的データ交換) が使われないときに起こりうる。

報告単位に由来する誤差

報告単位に由来する誤差は、事業サーベイよりも世帯サーベイにおいて一層重要である。これらの誤差について加盟国で行われた特別な研究は報告されるべきである。これは、例えば、過去の事象を思い起こすことの影響、他人が評価すると思われる特性を過大報告する傾向、回答者の注意の欠如、年齢や教育の影響、であろう。

統計家は、努めて調査票のデザインの改善や調査員の訓練を改善することによってこれらの誤差を最小にする。回答者に由来する誤差は、調査票や調査員の影響と区分することが難しい。そのためには、回答の変異性を測定するために、(副標本の) 回答者に対して、再度質問することが必要であろう。

調査票に由来する誤差

調査票の質を客観的に評価するのは難しい。

調査票を評価するためには、試験調査の説明と認識による(cognitive)実験ごとに調査票を分析した結論が必要である。

認識による実験的研究方法の第一の目標は、より正確な回答をもたらす質問戦略を注意深い分析を通じて認定することである。

例えば、個人の年齢を質問するサーベイでは、よく知られるタイプの測定誤差が生じる。若い年齢データが報告される傾向がある。生年月日を質問するのが望ましい。

調査員に由来する誤差

調査員からの影響は、調査員の選択がランダムでありうるので、他の測定誤差よりも測定するのが容易である。調査員に由来する分散を測定するために、標本設計をすることもありうる。

概念の測定に由来する誤差

例えば、

—統計が必要な対象期間は、データが基礎にしている期間と異なるかもしれない（暦年と会計年度）。

—純粋な価格変化の測定には、質的变化を考慮に入れる必要がある。

測定誤差を評価するために特別の技法（再面接、「調査員の影響」の推定、その他）を使うこともありうる。しかし、この種の実験の費用はしばしば、データ提供者とデータ収集者の双方にとって、考えられないほど高い。

国家統計機関は非常にひんばんに、検討を、調査票の特性間（あるいは調査票の特性と良質な何らかの外部情報との間）の一貫性のエディットに制限する。

エディットの叙述では、マクロエディット／マイクロエディットへ分割すること、そして後者を（欠損値／誤差／無記名）に分割すること、を考慮すべきである。

多年ベースで報告されるべきこと

- 1)この種の誤差を減らすために使った方法
- 2)調査員に由来する変異と偏り
- 3)誤差率

2.1.5 作業誤差

作業誤差は、コード化、キーづけ、編集、ウエイトづけ、製表といったデータ収集以後の過程における誤差である。

コード化に関しては、誤差率の何らかの尺度を品質管理の標準の技術を通じて獲得できる。

例えば、（キーづけの局面、あるいは国家統計機関のスタッフによるエディット処理の過程のいずれかにおける誤差のレベルを検討するために）処理された調査票の副標本の品質を検査することによって、である。

編集、ウエイトづけ、推定と製表については、誤差を測定するのが難しい。

報告すべきこと	提出頻度
異なるステップに対応して分割される処理誤差の率	数年ごと
推定についての方法論的ノート	数年ごと

2.1.6 無回答誤差

項目無回答と単位無回答は、サーベイによくある主要な誤差の源泉である。それらは管理されない仕方でも推定値に偏りをもたらしかもしれない。無回答についての伝統的で単純な指標は、単位か項目のいずれかの回答率である。それは、加重のものか（標本抽出のウエイトあるいは従業者、離職者他のような他の変数によって）、非加重のものでありえよう。実際の無回答者と範囲外の単位との区分は非常に重要である。測定できなかったのは無回

答によるのかフレーム誤差によるのかを国家統計機関が知らないなら、その単位を別個のグループとして区分にすることを推奨したい。

重要な問題は、回答単位が非回答単位とどう違うかである。無回答の一つの種類は、データの遅れからくる。そういった無回答の影響は、事後的な比較によって測定できる。無回答の第二の種類は、一般的にはさらに効率的なデータ収集技法（例えば、郵送質問ではなく面接）を使用して、無回答者の副標本の補助的サーベイを通じて、調査できる。

無回答（項目あるいは単位）は、補定（部分的あるいは全体的）あるいは回答単位の再ウエイトづけを通じて、生産過程のうちに処理されるべきである。この過程はランダムネスを取り入れるが、これは偏りや分散の推定に統合されるべきものである。

報告されるべきこと	統計の頻度		
	月次あるいは 四半期ごと	年ごと	数年ごと
	Eurostat への提出の頻度		
・ 単位と項目の回答率 ・	月/四半期ごと	年ごと	数年ごと
・ 無回答に対する補定及び/あるいは再ウエイトづけのために使用される方法の叙述	多年		
・ 補定による分散 ・	月/四半期ごと	年ごと	数年ごと
・ 無回答の偏り	数年ごと		
・ 上の推定についての方法論的ノート			

2.1.7 モデルの仮定誤差

モデルの仮定誤差は、統計あるいは指数を計算するための方法、すなわち、尺度調整 (calibration)、一般化された回帰推定量、完全範囲あるいは固定範囲に基づく計算、ベンチマーキング、季節調整、これまでの正確性要因にふくまれない他のモデル、の選択とともに生じる。

そういったモデルの品質は、分散や偏りによって評価されるべきである。この評価は統計に関連する信頼区間に統合されるべきであるが、この統合のための標準的方法はない。

季節調整モデル。季節調整の品質に関して非常に重要な点は、季節調整に先立つはずれ値(outliers)の探索および是正である。総合されたデータのはずれ値は、(ストライキのような) 不規則的現象に由来するもので、季節要因を推定するために是正されるべきである。

ベンチマーキング：

ベンチマーキング手続きは、一般的には年次統計と一貫させて改訂された年内系列を生

産するために使われる。この方法はそれが適用されたときに叙述されるべきである。最初の値と最終の値との比較が、この方法の影響を評価するために報告されるべきであり、ベンチマーク化された系列の正確性に対する影響が評価されるべきである（偏りの縮小、分散の増大）。

部分母集団についてのデータ収集ができないとき、総合された行政データが使用されるとき、必要とされるデータがなお入手できないとき（例えば、指数計算のためのウエイト）には、統計の一部を推定するために他の方法を使用することもありうる。

これらのモデルの選択手続き（すなわち、あるモデルが他のモデルに優先して採用された理由）は、おそらく、対応する推定値と結びついた推定誤差とともに、Eurostat に報告されるべきである。

系列を調整するために用いられた各モデルについて報告されるべきこと：

数年おきに、次の点に関するコメント

- ・ モデルの基礎にある仮定の検証
- ・ 既知の数量を「予測」するために歴史的データを使用することによる、モデルの予測力のテスト
- ・ そのモデルによって作成された結果と他の関連データ源泉との比較
- ・ 検査や横断的妥当性研究の利用
- ・ パラメータの推定に対するそのモデルの敏感性のテスト
- ・ そのモデルへのデータ投入の妥当性

3. 速報性と時間厳守性 (Timeliness and Punctuality)

生産過程での各段階の影響を分析するために、また遅れについて加盟国との討議を促進するとかヨーロッパのレベルでの事前推定値を発展させるという見地から、次の情報が有効であろう：

- －加盟国の統計的データ収集についての電子的なデータ転送の利用
- －加盟国で回答者に課せられている法的な締め切り期限
- －調査票と回収と追跡調査のために派遣される時と実地作業が行われる時を、具体的に指定したデータ収集の主な日程
- －週あるいは作業日ごとに受け取った調査票の数から見た回答のグラフ
- －編集局面の開始日と終了日
- －補訂局面のための日
- －データ収集局面が終了したと考えることができることを示す事象の詳細と日付
- －事前結果を計算し配布する日

- 品質検討（結果の一致<congruency>）の日と非開示手段
- 写真印刷用意のできた出版物が使用可能な日
- 出版日

Eurostat にデータを送る際の時間厳守性は、データ送付の定期性や遅れを指定する付録に収めた規則の構成にしたがって評価されるだろう。

4. アクセス可能性と明確性(Accessibility and Clarity)

Eurostat は配布過程に関する次の情報に関心を持っている。

- 出版物 1 冊
- 標本になった報告単位に送られる結果（あるのであれば）についての情報
- 結果の配布スキームについての情報（例えば、誰に結果が送られるか、利用者は誰か）
- 提供される統計に関するあらゆる方法論的な記録の写し

この情報は Eurostat レベルでの一貫した配布スキームを企画し、加盟国から電話をかける利用者によりよく伝えるためだけでなく、これが Eurostat 内での統計の解釈を助けになりうる各国の出版物にふくまれるあらゆるコメントを獲得する方法としても有用である。

5. 比較可能性 (Comparability)

5.1 空間的比較可能性

国における概念とヨーロッパの概念（統計単位、対象母集団、分類、観察された特性の定義）とに違いがある場合には、概念と、結果である統計に対する影響の推定の点からみた違いが叙述されるべきである¹。

同じように、加盟国は、測定についての既存のヨーロッパの勧告にその国が従っている

¹ 各国の統計が標本抽出サーベイの推定値からだけでなく、サーベイによって生産された統計と加盟国がめざす目的との間の何らかの乖離を調整するためのモデルから引き出されるときには、モデルの選択に由来する調整は、2.2.5 節「モデル仮定誤差」に含まれるべきである。ヨーロッパのモデルと、各国が行う可能性のあるモデルの調整の後に狙う基準との間の違いだけが、この節において報告されるべきである。より詳細な情報を得るためには、記録 Eurostat/A4/Quality/98/General Comparability をみよ。

かどうか、生じうるそれらの基準からの乖離、そして推定値への影響の対応する評価を、報告するべきである。

両方の情報を結びつけることで、統計の比較可能性におけるレベルの評価が提供されるだろう。

小さな副標本についての幾つかの補足的（自発的）サーベイによって、ヨーロッパの標準からの乖離を評価するべきである。

加盟国での実践と Eurostat の勧告との比較は、数年ベースで、重要な変化が生じたときに送り返られるべきである。可能な場合には、これらの違いが評価されるべきである。

5.2 時間にわたる比較可能性

ある利用者、特にマクロ経済学の領域で作業している利用者は、ある時点の推定値よりも時系列を使う。概念と測定方法の安定性は彼らにとって非常に重要である。定義、カバレッジあるいは方法の変化の詳細が与えられるべきである。統計単位の統合と分離がどう扱われているかの説明が与えられるべきである。統計単位（企業、世帯）の統合と分割は、しばしば、特に何らかの統計に大きな影響を与えるが、特にある分類によって分割されるときに著しい。二つの連続する時期からの統計の比較が、社会・経済的な傾向によってよりも、法律的な事象によって一層影響されるかもしれないときには、この点を報告することが重要である。連続性に影響を与える国のレベルでの法律におけるどのような変化も報告されるべきである。無視できない変化の影響もまた報告されるべきである。

季節調整された系列に関しては、比較は調整手続きにおける変化を考慮すべきであるが、季節調整モデルは 2.2.5 節「モデル仮定誤差」の下で報告されるべきである。

非連続性をもたらす大きな変化が起こるときには、加盟国は統計とともに理由を報告するべきである。それらの推定は後に Eurostat へ送付するべきである。

6. 一貫性 (Coherence)

6.1 予備統計と最終統計の間の一貫性

予備的指数と最終指数の一貫性は、予備的統計/指数が十分正確であるという意味で最大化されるべきである。

季節調整なしの最終的統計/指数と予備的統計/指数の間の年平均の違いのパーセンテージは、この平均/中位数のまわりの散らばりとして、要求されている統計毎に提供されるべきである。

6.2 年次と短期の統計の一貫性

幾つかの特性については、統計は年以下および年次の頻度の両方で生産されるべきである。利用者に対するイメージとしては、配布された情報についてある一貫性を狙うことが重要である。単純な確証の方法は、同じ概念の推定値ではあるが異なる頻度の統計からひきだされた推定値を比較することである。両方の頻度が水準についての推定量を提供するときには、平均的年次レベルあるいは総計、そしてそれら統計の少なくともひとつが指数のときには、成長率を比較するのが最も自然に思える。それらの比較は、国家統計機関において行われる可能性のあるベンチマーキングの前に行われるべきであり、加盟国は、対象とされた年次統計についてその年以下の統計をベンチマーク化するなら、別個に示すべきである。

各セクターに関して、加盟国は次の成長率を比較すべきである：

－短期のサーベイから計算した値	GRst
－年次サーベイから計算した値	GRyv
－発生、死滅と単位分類の変化の影響を除去して、年次サーベイから計算した値 ²	GRyc

生産された統計の動きに特定の影響を与える Eurostat の基準や勧告から乖離している習慣－基準の時期、指数のタイプ－ラスパイレス、パーシェ、フィッシャー他－もまた考慮されるべきである。

6.3 同じ社会・経済的領域での統計の一貫性

統計のあるグループ、おそらくは（貨幣価値で、大きさで、あるいは不変価格で、価格指

² より正確には、これは次のように定義される対象母集団に対応する…

－単位が n 年における主な活動に対応して区分される母集団

－母集団は $n+1$ 年における母集団と n における母集団との交わり部分(intersection)である。

標本は、この制限を考慮に入れるために階層化以後のものであるべきである。部分的に交替する標本の場合、信頼できる推定値は、標本が対応する階層で完全に更新されるのでなければ、 n 年の標本と $n+1$ 年の標本との交わり部分を考えることで成り立つ。

数で)異なるタイプの統計のグループは、よく同じ現象についての異なる角度からの表現をもたらす。利用者の質問を予想し、是正した行動を提供するために、それらの表現が一貫していることを検討することが非常に重要である。

例えば、離職、付加価値、株式の変動のような企業の短期統計は、生産価格によってデフレートした後に Prodocom 統計を伴った年次ベースで比較されなければならない。人口、失業、就業についての統計は比較されるべきである。同じく、意見サーベイと数量指標が比較されるべきである。社会統計(例えば、センサス)と企業統計によって測定された就業には、理想的には限られた乖離しかあるべきでない。個人の休日についてのサーベイは、旅行統計の数量的指標あるいはサービス部門(Horeca)の短期指標その他と何らかの一貫性を持つべきものである。

6.4 国民勘定と統計の比較

利用者の必要に最適な情報源泉について利用者に助言する位置にあるためには、サーベイからの統計を国民勘定と比較することも興味深い。したがって、他の情報源とそれら諸サーベイの全体的一貫性を示したものを利用者に提供することがありうる。国民勘定統計で使われる方法は、第一次データ源と施された調整をふくめて、対象とされた統計について叙述される必要がある。概念におけるズレもまた考慮されるべきである。

それらすべての比較の要約が Eurostat に毎年送られるべきである。

7. 費用と予算(Cost and Budget)

7.1 回答者の費用

加盟国は Eurostat に対して毎年、要求される情報を回答者が提供する平均費用を提供するべきである。企業あたり平均費用に標本の大きさを乗じて得られる母集団にとっての総費用もまた提供されるべきである。

加盟国はまた、個々の単位の間での統計の総負担のちらばりを測定するために、異なるサーベイの標本を調整しているか、あるいはサーベイに対する標本を交替させているかを示すべきである。

7.2 統計局の費用と予算

加盟国は Eurostat に対してサーベイに関わった総費用を提供するべきである。数年おき

に、(統計家、方法論者、プログラマーとアナリスト、調査員、事務職員とに区分した) 従業者s数が提供されるべきである。使用されているコンピュータ資源(ハードウェアとソフトウェア)や他の作業装置(光学的読みとり装置ほか)もまた詳細に示されるべきである。

用語集で定義されているタームの索引³

- 偏り、 3
- 一貫性、 3
- エディテイング、エディット、 5
- 電子的データ交換、 5
- 電子的データ移送、 7
- 推定値、 3
- 推定量、 3
- フレーム、 4
- フレーム誤差、 3,4
- 補定、 7
- 項無回答、 6
- マクロ・エディテイング、 5
- 測定誤差、 3,5
- マイクロ・エディテイング、 5
- 分類の誤り、 4
- 無回答、 6
- 範囲外、 6
- 過大カバレッジ、 4
- 作業誤差、 6
- 登録、 3
- 標準誤差、 3
- 過小カバレッジ、 4
- 単位無回答、 6
- 分散、 3

³ 資料 Eurostat/A4/Quality/98・General/Glossary を参照のこと。

2 統計における品質の定義

Eurostat

序

本稿は、品質に関する報告にふくめられるべき主要な論題をリストする。報告にとり入れられるべき詳細は、記録「品質報告への提案」でとりあげられている。

1. 品質の定義

品質は ISO8402-1986 において、「生産物あるいはサービスが、明言されているか暗示されているニーズを満足させる能力において持っている特性と属性の総体」と定義されている。

これは、総合的品質管理 (Total Quality Management) の枠組みの中で分析されうるだろう。我々は、過程全体ではなく、まずは「生産物」に集中するという、より限られたアプローチをとる。例外が2つある。すなわち、適合性 (relevance) は評価であり、速報性 (timeliness) は、過程の情報である日付に関わる。

統計の品質は、幾つかの基準との関連で定義することができる。

- 統計的概念の適合性
- 推定値の正確性
- 結果の配布における速報性と時間厳守性
- 情報へのアクセス可能性と明確性
- 統計の比較可能性
- 一貫性 (coherence)

品質の尺度ではないが、統計生産に利用可能な資源は、品質を制約するものである。品質のガイドラインに加盟国が従う能力を評価する際には、利用可能な資源を考慮することが必要である。

適合性

サーベイは利用者のニーズに対応しているなら適合的である。したがって利用者とその期待の確認が必要である。

推定値の正確性

正確性は、推定値と（未知の）真の母集団の値との間の近接度と定義される。ある推定値の正確性の評価は、その推定値と結びついた全体的誤差の分析をふくむ。

結果の配布における速報性と時間厳守性

大部分の利用者は、ひんばんにそして事前の予定通りに公表される最新の数字を求める。

情報へのアクセス可能性と明確性

統計データは、利用者にとってアクセスが容易であり、利用者の望む書式で入手可能であり、記録が十分であるときに、最大の価値をもつ。統計の利用と解釈に際しての援助もまた提供者から手近に与えられるべきである。

統計の比較可能性

ある与えられた特性に関する統計は、その特性についての推定値が場所と時間を通じて信頼できる比較を可能にするときに、最大の有用性を持つ。比較可能性の構成要素は、ヨーロッパレベルで合計された統計の意味を評価するために、同じ統計を各国間での比較に重きをおく。

一貫性

統計がただひとつの出所からとられたものである場合には、基本的概念をより複雑な仕方でも結合しても信頼をおくことができる点で一貫している。統計が異なる出所からとられ、特に頻度の異なる統計的サーベイからの場合には、統計は、共通の定義、分類、方法論的基準に基づいている限りにおいて、一貫している。そうであれば、統計が利用者に伝える声明は相互に明確な関係を持つか、少なくとも互いに矛盾はしないだろう。統計間の一貫性は、通常は異なる仕方でも、また異なる用途のために作成される異なる統計の比較についてのものである。

費用について

統計の品質と統計を作成するために利用可能な資源との間には強いつながりがある。品質を評価する過程では、費用の評価が念頭におかれるべきである。

2. 正確性

品質のこの側面は、多くの統計機関や大学人によって広く研究されてきた。正確性では標本誤差と非標本誤差に区分される誤差の分析が焦点となる。

誤差は以下のように区分できる。

2.1 標本誤差

2.2 非標本誤差

2.2.1 フレーム誤差

2.2.2 測定誤差

2.2.3 処理誤差

2.2.4 無回答誤差

2.2.5 モデルの仮定誤差

Groves(1989)、Lessler and Kalsbeek(1992)、Särndal, Swensson and Wretman(1992)、Biemer and Fresco(1995)、Quality Measurement Model of the BLS(1995)は同じ区分を使っている。

これまで、統計家は主として標本誤差に焦点をおいてきた。統計機関が発表する正確性についての推定は、通常は標本誤差だけについてのものであった¹。しかし、非標本誤差についての研究が急速に発展しており、そういった誤差の第一次的評価に使うことの出来る幾つかの方法が今では利用可能である。例として、Biemer et al.(1991)を参照してもらいたい。

3. 速報性と時間厳守性

利用者は一般に、統計的情報について、最小限の時間で生産され、利用可能になると同時に公表され、適切な場合には、定期的に利用可能になることを求める。生産時間を常に最短に保つということは、効率的な生産技法を伴い、しばしば大きな費用を要する。データの収集、エディティング、補定、推定、配布の全過程は、作業期間を最小にするための管理下に常におかれるべきである。この問題の部分的解決は、回答者の一部の標本に基づく、速報推定値を提供することである。それら速報推定値の発表とそれに続く改訂された情報の配布にも注意が必要である。

非常に速やかではあるが信頼性の劣る形で生産され配布される（しばしば民間の）統計情報と、政府機関からのより正確な統計とは、ときとして競合する。速報性と正確性との

¹ 実際には、推定された分散値は、実践的には回答率と何らかの非標本誤差一例としてはランダム測定誤差に依存しているが、それはすべての非標本誤差を考慮しているわけではない。特に、大部分の系統誤差は、この正確性の測定にはふくまれていない。

間のトレードオフは、ある利用者には明らかな影響を与える。このことは、統計機関に高い基準の品質を落とすことなしに統計の速報性を改善するという課題を与える。

4. 統計の明確性とアクセス可能性

配布は情報の連鎖において決定的なステップである。統計局内部のどこかに貯蔵されている「良い統計」があるというのでは不十分である。統計は、すべての潜在的利用者に対して、適切な形で入手可能にされなければならない。第一に、利用者は、どんな統計が入手可能であるかを簡単に知る位置におかれるべきである。第二に、統計への物理的アクセスが便利であるべきである。第三に、統計は、概念と方法についての必要な情報を伴ったものであるべきである。主題についての専門家とそうでない者との違いをおくために、レベルの異なる説明がときとして望ましい。最後に、統計の分析は、統計生産物についてのサービスの次元を強調してよい。

5. 比較可能性

各国の定義とヨーロッパの定義との間には違いがありうる。例えば、ある参加国で粗投資の概念は、金融リースを通じて獲得された資本財の金額は除外しているが、ヨーロッパの定義にはふくまれていることを考えればよい。個別企業のレベルで、金融リースを通して得られた金額を収集することができないなら、推定値は集計レベルで行われる必要がある。比較可能性はEU内部での比較可能性に限られない。すなわち、Eurostatは、他の諸国あるいは国のグループ（アメリカ合衆国、日本）とEU統計との比較可能性を評価する位置にいないなければならない。

比較可能性はまた時間にわたっても必要である。対象とする概念あるいは測定過程の修正からくる変化は記録されるべきであるし、それらの変化の影響が評価されるべきである。同じようにして、継続性に影響をもたらす社会の変化一般（例えば、新立法、統合や分割）が考慮に入れられるべきである。

6. 一貫性

多様な源泉からの類似の統計がある場合には²、それら統計を確認し、可能ならどのような差異も数量化するべきである。異なるサーベイによって生産された2つの統計セット間の食い違いは、データの収集過程の違い、あるいは報告単位の違いが異なる結果をもたら

² ここで「源泉」とは、政府統計制度が国レベルで行うサーベイあるいは行政データから計算された統計を意味する。

すこと、よることがありうる。この情況は、ベンチマーク化（例えば、年次結果上での月次あるいは四半期統計）によって、あるいは異なるサーベイ結果の結合によって改善されるかも知れない。どのような場合であれ、異なる概念に対して異なる用語を使用することによって、利用者の誤解を防ぐべきである。

さらに、複雑な概念を推定する統計（比率、弾性値ほか）は、（対象母集団、特性値、対象期間そして統計単位についての通訳可能な定義を伴った）一貫した基本的統計に基づくべきである。

7. 費用の制約

品質は費用と関連している。サーベイ管理者が使用可能な資源に関する情報は、彼らが受け取る予算をどう最大限に活用しているかの評価を可能にする。加盟国が利用可能な資源は、品質の諸尺度を満たす上での各国の本当らしい能力を評価するときに、考慮されるべきである。費用については2つの構成部分を考えることができる。すなわち、統計局にとっての費用と、報告単位（代表的には企業あるいは企業の一部）にとっての費用である。回答者にとっての費用は、主に企業規模や活動部門ごとで違う³。新しいサーベイに対する費用は、既存のサーベイの費用よりもはるかに大きいだろう。行政データの利用によって、企業の負担は劇的に減少するだろう。個人あるいは世帯に対する費用は考慮されない。個人への負担の影響は、非回答誤差で考慮に入れられるべきである。

費用は、品質それ自体の構成要素というよりも品質の改善のための制約なのである。

注1：品質の異なる構成要素にはトレードオフがある。特に：

速報性 / 正確性

正確性 / 場所間の比較可能性⁴

適合性 / 時点間の比較可能性

適合性 / 正確性

大きな領域での一貫性 / 部分的領域での適合性

注2：品質の構成要素への上の分解は独自のものではない。他の組織が統計の共同利用について他の分解を提案してきている。

1) *United Nations*(1983). 全体で9つの点は、以下をふくむ：

— 「時点間の比較可能性」

³ 異なる部門では異なる質問がおこなわれるからである。

⁴ 場所間の比較可能性は、全国的な—しかし比較できない—概念よりも測定の困難な共通の概念を課すかもしれないからである。

- 「他の統計との比較可能性」
- 2) *Statistics Canada*(1992). 全体で9つの点は、以下をふくむ
- 「時点間の比較可能性」.
 - 「他の源泉からの統計との比較可能性」
- 3) *Beekman and Struijs*(1993). 3つの主な分類は以下である：
- 「統計生産物の比較可能性」：1つの統計からの統計と、他の統計からの統計とを、時間（時系列の継続性）における比較可能性と同時的統計の比較可能性の見地から、比較するか結合する。
- 4) *Statistics Netherlands*(1995). 「一貫性の必要」についての Section II.2.は7つの見地をあげている。
- 経済活動の異なる領域にふれた統計の加法性、例。商業や製造業での資本形成
 - 異なる主題についての統計の比較可能性、例。労働サーベイからの雇用データおよび生産統計からの離職統計
 - 短期統計の結果の年次統計との一貫性、
 - 地域統計の結果の全国統計結果との一貫性、
 - 期間 t についての結果と期間 $t+1$ の結果との継続性、
 - 国の統計の国際統計との比較可能性、
 - 企業統計の概念と結果の国民勘定体系(SNA)やヨーロッパ勘定体系(ESA)との調和性、
- 5) *Statistics Sweden*(1994)：4つの構成要因、すなわち、内容、時間、信頼性、アクセス可能性がとりあげられている。内容は他の統計との比較可能性をふくみ、時間は時点間の比較可能性をふくむ。非常に重要な、まさに利用者本位のアプローチが、品質の宣言に付されている。

文献

Biemer, P., Groves, R.M., Lyberg, L.E., Mathiowetz, N.A. and Sudman, S. (eds) (1991) *Measurement Errors in Surveys*, New York: Wiley

Beekman, M. M. and Struijs, P. (1993). *The quality of economic concepts and definitions*. Statistical Journal of the United Nations ECE 10, 1-15.

Biemer, P.P. and Fesco R. S. (1995) *Evaluating and controlling measurement error in business surveys*, in Cox and al. (eds) Business Survey Methods, New York: John Wiley, pp.257-281

Bureau of Labor Statistics (1995), Quality Measurement Model, presented at the International Conference on Survey Measurement and Process Quality, Bristol (UK)

Groves, R.M. (1989) *Survey Errors and Survey Costs*, New York: John Wiley

Lessler, J.T. and Kalsbeek, W. D. (1992) *Nonsampling Errors in Surveys*, New York: John Wiley

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

Statistics Canada (1992). *Policy on Informing Users of Data Quality and Methodology*. Policy manual, Statistics Canada, April 7, 1992.

Statistics Netherlands (1995) *Reference Manual on Design and Implementation of Business Surveys*. Koeijers, E. and Willeboordse, A. (editors). First draft, on request of Eurostat.

Statistics Sweden (1994) Quality definition and recommendations for quality declarations of official statistics (authors: Chris Denell, Eva Elvers and Bengt Rosén).

United Nations (1983) *Guidelines for quality presentations that are prepared for users of statistics*. Statistical Commission and Economic Commission of Europe, Conference of European Statisticians, Meeting on Statistical Methodology, 21-24 November 1983.

付録

使用されている基本概念

統計の基本概念は統計的特性である。これは、次の作業によって定義される。与えられる統計的尺度が、特定グループ内の単位の変数の値を要約するために適用される。単位の全体的グループは通常母集団と呼ばれる。しばしば、異なる部分集団、研究領域が関心の対象になる。母集団、単位と変数はすべて、対象期間、時点あるいは期間を持つ。

主題に関連した問題に関する質問に答えるためには、適切な統計的特性の集団、関心の対象である特性、がある。ときとして、それらを統計的サーベイにおいて操作可能にすることが難しい。サーベイは、サーベイの見地からは実行可能であり、関心に十分近い目的の特性に向けて行われる。

サーベイは、統計的な、目標とする特性の値の推定値をもたらす。ひとつの推定値は、調査からのデータ、標本での観察、モデル、あるいはそれらの何らかの結合に基づくかもしれない。補助的情報はときどき標本データとともに使われる。

3 統計の質に関する用語解説

Eurostat

この用語解説の範囲は、この文書の姉妹編である文書「品質の定義」、や「統計の品質報告書に関する提案」、そして特定領域での品質に関する報告書についての他の提案、を読むときに、あいまいさを避けるために必要なものに限られている。これは、統計の品質についての共通の分析を企画する上で必要な条件である。

これらの定義は、US Department of Commerce (1978) Lesser and Kalsbeek(1992)による類似の作業を改作したものである。それらの文書自体は、サーベイにおける誤差についての入手可能な文献の論評と分析である。

サーベイ・サンプリングに関する定義は、Särndal et al(1992) から改作した。

文献

Lesser, J.T and Kalsbeek, W.D (1992) ,*Non Sampling Error in Survey*,
New York : John Wiley

US Department of Commerce(1978),*Glossary of Non Sampling Error
Terms : An Illustration of a Semantic Problem in Statistics*,
Statistical Policy Working Paper 4 , Office of Federal Statistical
Policy Standards.

Särndal,C-E. Swensson,B. and Wretman,J.(1992) *Model assisted survey
sampling*, Springer-Verlag.

基準年 (Base Year)

商品のウエイトが消費者の選択にしたがって固定される特定の年をさす。それは(連鎖指数とは反対に) 固定基準指数を計算するために使われる。

偏り (Bias)

偏りは、推定量の期待される値と推定されている真の値との差である。

連鎖指数 (Chain Index)

連鎖指数は、現在時を最近時と比較して計算される。さらに、固定基準指数とは反対に、連鎖指数の商品ウエイトは最近時に基準が変えられる。

実際には、最近時とウエイトが計算される時期には時間のズレがしばしば生じる。ウエイトはしばしば、最近の構造的サーベイの結果の利用可能性に依存しており、年ベースで計算される。この時間のズレは、予測モデルを使うことで縮小されるかもしれない。

コード化 (Coding)

コード化は、言語的情報をより簡単に数えたり、表にすることができる数字あるいは記号に転換する手続きである。

コード化の誤り (Coding Error)

サーベイへの回答に誤ったコードを指定すること。

一致 (Consistent)

推定量は、誤差がある与えられた大きさ以上になる確率が、標本が大きくなるにつれてゼロになる傾向を持つとき、一致性をもつ。

変異係数 (Coefficient of Variation)

推定量の変異係数は、平均に対する標準偏差の率に等しいちらばりの相対的尺度である。実際には、これは点推定値の推定標準偏差を点推定値で除した値であり、%で表現される。

固定範囲 (Constant Scope)

固定範囲は枠になる母集団に関わる。それは発生、死滅、統合や分割をふくむべきでなく、完全範囲 (full scope) とは対照的に、その母集団は、指数が推定される両時点について同一であるべきものである。

標本の調整 (Coordination of Sample)

標本を独立してそれぞれ抽出するのではなく、幾つかのサーベイでの標本の重複分を増やすことが、積極的調整として知られている。積極的調整は、二つのサーベイからの相関する変数に依拠して統計のより良い正確性を獲得するため、時の経過にわたって繰り返されるサーベイ (パネル) においてしばしば追求される。異なるサーベイについての標本の重複を減らすのが消極的調整として知られている。消極的調整は、回答単位の回答負担をより等しくするために追求される。

現在時期 (Current Period)

統計が生産される時期をさす。

死滅単位 (Dead Unit)

死滅単位とは現在時期中に存在しなくなった単位である。企業では事業の閉鎖により、個人には死亡によって生じる。

輪郭描写 (Delineation)

統計単位の輪郭描写は主として、法的単位のグルーピングあるいは分割、そしてEUの調和規則を適用することと/あるいは分類を使用することによって行われる。まずい輪郭描写は二つの単位が一つとみなされたり、反対の場合（例えば、スーパーマーケットに付されたカフェテリア）に起こりうる。輪郭描写での難点が起こりやすい他のケースは、総就業を、1年間のフルタイムに相当するよりも少なく示す非常に小さな企業で生じる。誤った輪郭描写は世帯についてもまた起こる。

領域 (Domain)

点推定は、関心領域に対してひきだされる。領域は、全体母集団、あるいは別個の推定が企画される何らかの特定部分母集団である。この領域は標本抽出の階層と完全に重複するか、それらを越える。

エディティング、エディット (Editing, Edits)

エディティングは、データの収集や獲得の結果であるデータ記録における個別の誤差を探索して調整する手続きである。コンピュータに支援されたエディティングにおける、欠損値、間違った値あるいは疑問のある値を識別するためのチェックは、エディット規則あるいはエディットと呼ばれる。エディットによる変更は、誤差が認定されたときの処置の結果として、項目（問題）値が調整されるときに起こる。

電子的データ交換 (Electronic Data Interchange: EDI)

EDIとは、事業の共同者間の標準的な書式による事業取引情報の電子的な移動である。EDIは事業の内部的システムから情報を電子的に検索し、コミュニケーション・ネットワークを通じて取引の共同者/供給者/顧客/政府に対してその情報を提出する機会を事業に対して提供する。

誤差 (Error)

サーベイ値と対応する真値の差。

推定値 (Estimate)

推定値は、推定量を用いて計算された値である。

推定量 (Estimator)

推定量は、特定のパラメーターについての推定値を計算するための数理的関数である。

期待値 (Expected Value)

同じ基本的条件の下で繰り返し行われると想定される全てのサーベイからの仮説的平均値。

フォローアップ (Follow-Up)

選出された標本のメンバーのうち、ひとつのデータ収集戦略（例えば、電話、面接あるいは郵送）によっては回答が得られなかったメンバーに、回答率を増加させるために、同じである異なるデータ収集戦略によって接触する手続きである。この用語は、回答者のパネルの間での繰り返しサーベイをさすときにも使われうる。

フレーム (Frame)

サーベイのためのフレームとは、目標である母集団の要素について、境界を定め、認定し、アクセスを可能にするリストあるいは単位のリストである。

フレーム誤差 (Frame Error)

フレーム誤差は、インプットされたデータの本来的限界によってか、データの獲得や処理の遅くれ及び誤差によって生じうる。フレーム誤差は、

- ・ カバレッジ誤差－誤った算入や除外および重複
- ・ 分類誤差－産業別、地理別、あるいは大きさ別に分類されなかったり、誤って分類される単位。
- ・ 接触誤差－不完全なあるいは誤った接しよくによるデータを伴っている単位。

完全範囲 (Full Scope)

完全範囲とはフレーム母集団に関わり、固定範囲とは対照的に、指数が推定される期間中の発生、死滅、合併、分離をもふくむべきものである。

ヘドニック法 (Hedonic Method)

ヘドニック・アプローチは、商品は物理的およびパフォーマンス上の特性をひとまとめたものと見ることができ、品質とはそれらの特性の変形であること、

ひとつの商品の市場価格の変化は品質価格構成部分と純粋な価格構成部分とに要因区分できること、ひとつの製品の価格はその特性と相互関連していること、を想定する。ヘドニック価格を推定するために、回帰法が、品質変化の影響を除去（し純粋な価格変化を反映）するために使われる。

補定 (Imputation)

補定とは論理的エディットあるいは統計的手続きを使って欠損データを代置することである。これは無回答項目あるいは無回答単位を問題にする。

現在価格の指数、あるいは価格指数 (Index at Current Price or Index in Value)

固定ウェイト要因を考えることによって、現在価格で表された価格の変化を測定するのに用いられる指数。

固定価格の指数あるいは数量指数 (Index at Fixed Price or Volume Index)

固定価格を考えることによって、数量の変化を測定するために使われる指数。

無回答項目 (Item Non-response)

無回答を参照。

項目回答率 (Item Response Rate)

項目回答率は、ある項目について、利用可能なデータ数を、利用可能データと欠損データの数で除すことによって計算されるべきである。

ラスパイレス指数 (Laspeyres Index)

価格（.数量）ラスパイレス指数は、（パーシェ指数と対照的に）対象の状況において観察された数量（.価格）ウェイトを用いる。

マクロ・エディテイング (Macro-editing)

合計量を検討するか、すべての記録あるいはそれらの部分集団に対して統計的方法を適用するかによって、疑問のあるデータを指摘する手続き。

平方平均誤差 (Mean Square Error)

推定量の分散プラス偏りの平方。

測定誤差 (Measurement Error)

測定誤差は、サーベイ中に収集された値と真（個別）値との差と考えることができる。それらの誤差は次のものから発生する：回答者が真の値を報告する点での失敗（回答者誤差）、調査員が値を正しく記録することでの失敗（調査員誤差）、あるいは値を正しく測定するための装置（調査票）の失敗。

マイクロ・エディティング (Micro-Editing)

個別の観察を検討することによって誤差を見いだすための悉皆的チェック

分類の誤り (Misclassification)

分類の誤りとは、目標母集団に属していて（例えば、産業別、地理的領域別あるいは規模別に）、誤って分類された単位である。

無回答 (Non-response)

無回答は、資格の無い単位（死滅した単位、範囲外の単位、無接触）を除いて、選択された標本において、全ての単位について完全な情報を収集する点での失敗の結果生じる。無回答には2種類ある。第一は、接触した抽出単位が回答するのに失敗すること。これは単位無回答である。第二は、単位が調査票に不完全に対応すること。これは項目無回答といわれる。

無回答による偏り (Non-response Bias)

これは、サーベイ分析を、観察された単位に限ることから生じる偏りである。

無回答誤差 (Non-response Error)

無回答誤差は、無回答による欠損データの補定の結果である。

非標本誤差 (Non-sampling Error)

非標本誤差は、標本抽出誤差以外の全ての源泉に帰することのできる誤差である。非標本誤差はすべてのタイプのサーベイにおいて、計画、実施、データ処理、および最終的な推計段階で生じる誤差である。

範囲外の単位 (Out-of-scope Unit)

現期間に、目標母集団に属していないので、標本抽出のフレームにふくまれるべきでない単位である。現期間中に死滅、合併あるいは分割した単位は、範囲外の単位とは考えられない。

過大カバレッジ (Over-coverage)

目標母集団に属さないが標本抽出フレームに含まれる単位である。これらのケースは、通常は、接触した単位について観察されるが、非接触単位あるいは標本から除外される単位には必ずしも生じない。過大カバレッジの理由は、範囲外単位と更新されないフレームである。レジスター（記録簿）における過大カバレッジは、一般にその抽出リストから引き出された推定量に偏りをもたらす。

パーシェ指数 (Paasche Index)

価格（数量）パーシェ指数は、（ラスパイレス指数と対照的に）現在の状況において観察される数量（価格）ウェイトを用いる。

処理誤差 (Processing Error)

処理誤差は、正しく計画されたサーベイ方法を誤って実施することによって生じる、最終サーベイ結果における誤差である。処理誤差は、調査票の印刷とともに、収集後の総ての作業をふくむ。大部分の処理誤差は個々の単位についてのデータにおいて起こるが、製表や推定にも入ってくる。

記録チェック (Record Check)

一つのデータ収集法によって獲得された個別単位についてのデータを、異なるデータ収集法で得られた利用可能な記録からえらえる同じ単位についてのデータと対比してチェックする研究（例えば、センサスで報告された年齢を、出生証明からの年齢についての情報と比較する）。

記録保持誤差 (Record-Keeping Error)

回答に使われる記録における不正確性から生じる誤差。

対象期間 (Reference period)

指数との関わり：

一つの指数系列に 100.0 の値を与えられる期間である。対象期間と基準期間は固定基準指数では同じであるが、連鎖指数では異なる。

他の関わり：

対象期間は統計が生産される期間である。（また、指数のフレームワークの中では、現在期間と呼ばれる）。

拒否率 (Refusal Rate)

面接には成功したが、必要な情報を与えることを拒絶する企業の割合である。これは、もともとの標本全体の大きさを拒絶数を除することで計算される。

記録簿 (レジスター : Register)

これは、総ての統計単位の更新されたファイルであるべきであり、そこからフレームが取り出される。レジスターは一般には行政データからひきだされる。

相対標準誤差 (Relative Standard Error)

変異係数を参照。

再ウエイトづけ (Reweighting)

再ウエイトづけは、単位無回答を処理する一つの調整手続きである。それはまた、補助的情報の使用を通じて正確性を増大するために使われる。標準的方法は、事後階層化、尺度調整および回答傾向のモデル化(response propensity modelling)である。

標本設計 (Sample Design)

標本設計は、目標標本および最終標本の大きさ、階層定義、および標本選定方法についての情報を提供すべきである。その情報はまた、方法の変化の詳細とそれらの変化が時間にわたる比較可能性に与えるすべての影響をもふくむべきである。

標本誤差 (Sampling Error)

標本誤差は、目標母集団の全企業(businesses)の完全なセンサスを行うのではなく、母集団の部分集団だけをサーベイすることによる誤差である。

標準誤差 (Standard Error)

これは、推定量の分散の平方根として計算される。

サーベイ設計 (Survey Design)

サーベイ・設計は次の問題を取りあげる。すなわち、生産される統計の種類、母集団は何か、何時のか、どれだけの正確性を持たせるか、関心のある母集団のどの単位についてどんなデータが収集される予定か、要求される統計を生産するために、データを収集し処理する方法は何か、である。作業可能な、組織的および行政的問題が通常取りあげられる。

目標母集団 (Target Population)

これは、サーベイの対象となり、統計が生産される対象になる要素（分析の単位）についての概念的母集団である。

真値 (True Value)

現実の母集団の値であり、これは、一次データの収集と数理的操作の遂行との両方において、完全な測定装置を使い、どのようなタイプの誤差もないときに獲得される。

過小カバレッジ (Under-coverage)

目標母集団に属する総ての単位をフレームにふくめる点での失敗。これには、主として、実際の発生あるいは分割、および誤って分類された単位、のいずれかによる、フレームにふくまれない（新しい）企業がある。

無回答単位 (Unit Non-response)

無回答を参照せよ。

単位回答率 (Unit Response Rate)

情報を獲得するのに適格な標本の%である。面接サーベイについてなら、計算式の分子は面接した数である。分母は、標本の大きさマイナス不適格な回答者、すなわち、マイナスその特定研究に対して定義された潜在的回答者に関する基準に合わない者、である。加重回答率がより有効である。これに加えて、次のウエイト抽出フレームからの何らかの既存の重要な変数（離職あるいは大きさ）に合わせた標本ウエイトを使った、もう一つの無回答率も計算できる。後者の数字は無回答の影響を説明する。

変異 (Variance)

確率変数の分散は、平均の周りでの経験的値のちらばりを反映している。それは、平均値の周りでの変数の平方平均偏差である。

4 統計規範の品質

Raoul Depoutot¹(Eurostat)

1 序

統計規範統計規範の品質は、統計の品質に比べて、文献においては非常にわずしか分析されてこなかった。これは、一部には、統計規範が統計生産過程における中間生産物とみなすことができること、一部には、規範の発展は国家統計局においては希な出来事だからによる。これに対して国際統計機関は、多くのエネルギーをこの活動に投入した。これら機関は、ごく希にしか統計生産に関わらないので、規範を最終生産物とみなしている。Eurostat はその TQM プロジェクトの枠組みの中で、統計における品質の測定のための規則を採択し、現在では規範の質に関して類似の作業を仕上げようとしている。本稿はこの作業への寄与の1つである。

2 統計規範

統計規範とは何か？ それは、

－現実を、標本抽出理論の数理的形式に対応するように、体系的な仕方で叙述する道具である。

－統計家が、統計利用者とコミュニケーションするときの補助的装置である。

－基本的な個別的情報を提供する者（世帯、個人、企業）との間でのコミュニケーションのための参照フレームである。

統計規範のあらゆる側面を、一般的なコミュニケーション問題（透明性、客観性、メッセージの効率性、情報の費用、文化的背景の影響、その他）の特殊な場合とみることもできる。統計規範には異なったタイプがある。すなわち、母集団、統計単位、分類、特性、対象期間と「統計的要約」である²。統計生産、配布手続き、メタ情報についても規範がありうることに注意してもらいたい。それらは、一般に低い地位にあるか、未展開なので、ここでは検討しない。

¹ この論文に示されている見解は必ずしも Eurostat の公式的立場を反映するものでない。

² 母集団 P の特性の統計的要約は母集団 P 上のこの特性によってとられる値の単一変量であれ多変量であれ、数値的関数の値である。例は、平均 (the mean)、平均 (the Average)、合計、分散、分位点、3 つの四分位数、他である。

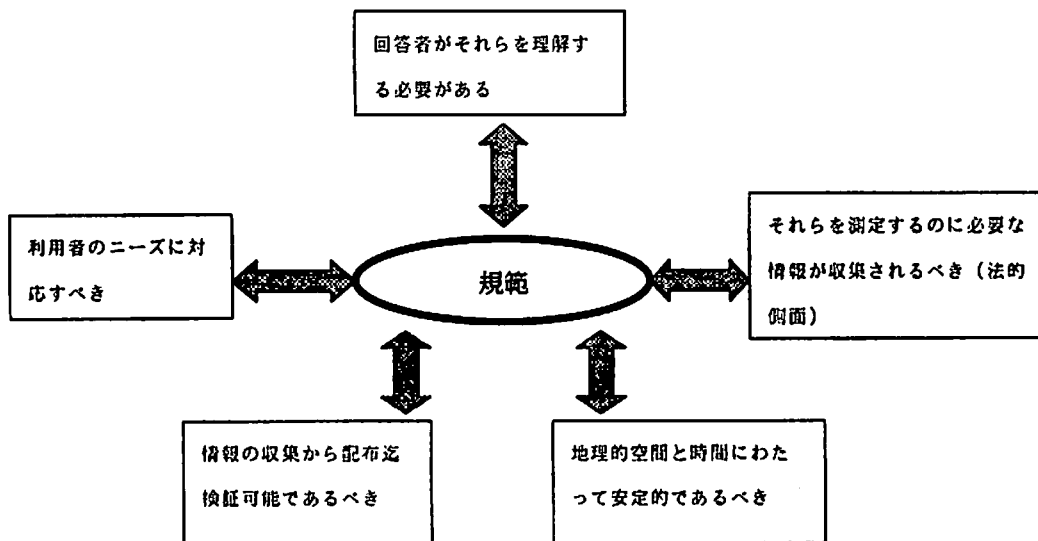
3 統計と統計規範と統計生産

統計は生産過程において統計規範を適用した結果である。Eurostat(1998a)を改作すると、統計は以下のように定義される。「ある対象期間に、統計単位のある母集団についての選ばれた特性に関する測定と統計的要約の推定値の計算の結果である数であり付属するメタ情報であり、それはおそらくある与えられた分類にそって区分され、ひとつあるいは多くの方法にそって配布される」。この定義において、われわれは統計的要約、特性、母集団、統計単位、対象期間、分類は統計規範であることを認めている。統計生産は他の面に存在する。同じ統計規範が幾多の統計によって同時に使われることを思い起こすことが重要である。

4 統計規範の品質：

規範の品質を一般的により正確に定義するためには、われわれは、構成された ISO アプローチによる経験によって示される「同時的規範」を対照するだろう。経験的アプローチは、Boltanski(1970)、Thevenot (1982)、Volle et all (1971) による幾つかの分析を使う。

経験的アプローチ：



構成的 ISO アプローチ：

われわれが品質についての一般的アプローチ、すなわち ISO8402 規範に固執するならば、品質は「生産物あるいはサービスの、明言されたか暗示された必要に対する能力における特性と特徴の総体」である。統計規範は主として何に使われるのか？ それは、統計の生産、統計の利用、行政や民間機関のレジスター上における単位の分類のような非統計的目的である。

われわれの見解では、統計的サーベイの回答者は規範の使用者ではない。彼らは、統計規範を正しく理解しようとする個人的な動機を持たない。というのは、彼らが質問に対して答えることは、彼らに直接的影響を与えないからである。したがって、統計規範は、回答者（社会保障書式、所得税書式ほか）に対して何らかの利益あるいは負担を条件づけうる行政的書式のために使われる概念とは違うのである。

一般に、これらの規範を使う統計の生産者は、主として公的組織に、そして非常にひんぱんに国家統計局自体にいること、民間機関はこの点で何らかの特定の必要を示すことは希であること、単一の単位を分類する目的はわずかの重要性しか持たないこと、を考えると、統計の利用者に主として焦点をおくのが合理的である。彼らは、ある意味で統計規範の最終的利用者なのである。

統計規範の品質は、したがって、「この特定の規範に基づいたすべての統計の利用者たちの、明言されたか暗示されたニーズを満たすよう得られる統計の生産に貢献する能力における特性と特徴の総体」と言い換えることができる。われわれは、ここで証拠として Beekman and Struijjs(1993)が示した結論に到達する。

5 統計規範と統計の品質

Eurostat(1998)に示された統計における品質の7つの構成要素とその部分的構成要素の幾つかを使って、それらの各々と統計の品質とをどう関連づけることができるかを見ることにしよう。われわれは、規範の幾多の見地に依拠して幾つかの概念を実際に測定するのは、はるかに難しいものでありうるという意味で、規範と測定との間の相互作用を明確に考慮に入れる。この点は、例えば Grais(1998)が述べている。

われわれは、ここで、統計における品質の構成要因について、何を規範に帰属させることができ、そして何が生産の側面に依存するかについて検討できる。

統計の適合性は、統計内容が利用者の必要に対応することを求める。このことは、考察される統計が使う統計規範は、利用者の必要を満たすべきことを意味する。

正確性は、生産過程にすべて依存する標本誤差と非標本誤差の大きさの結果である。非標本誤差のうち、フレーム誤差は、使われた規範（対象期間、統計単位、階層分類、外因性の特徴・・・）に密接に依存するサンプリング・フレームの品質の問題を生む。測定誤差と処理誤差（特にコード化）は、回答者と国家統計局のスタッフによる規範の理解という問題を生む。もし提案された規範が直接的には収集不可能なら、モデリングが要求されうる。この規範が複雑で長期のデータ収集を指導するものなら、非回答率を増加させ、非標本誤差の増加をもたらさう。明らかに、これらの誤差の一部はまた、過程それ自体にぶら下がっている。

統計の速報性（と定期性）は、一部には同じく使われる規範にも依存することができる。いくつかの規範は、とりわけ極端に速やかに観察されうるし、他の規範は統計処理における遅れをもたらす。しかし、規則性との明確なつながりはない。

明確性（とアクセス可能性）は、統計が基礎を置く規範とつながっているのは明確である。規範が複雑で法外なものであればあるほど、統計に明確なメッセージを付するのはより難しくなる。

比較可能性は、規範の適用の均一性、および地理的空間と時間を越えた幾つかの生産の側面の均一性という問題を生む。比較可能性と、規範と測定の間につながりについての詳細な提示に関しては、Eurostat(1998b)を参照のこと。

一貫性は、異なる統計を一貫した仕方です結びつけることを要求する。この点で、生産の側面が無視されることはありえないとしても、規範は非常に重要である（Eurostat,1998c）。

統計体系の完全性は、必要な規範が開発され、それに加えて必要なサーベイが行われてきていることを要求する。

最後であっても重要性に変わりがないのだが、規範は、サーベイへの回答者と統計局との双方に対して、統計生産の費用において大きな影響を与える。

6 統計規範の品質を特徴づける

上に示した統計規範の品質を、7つの構成要素、すなわち「普遍性と安定性」、「適合性」、「明確性」、「検証可能性」、「収集可能性と極度の儉約」、「一貫性」、および「完全性」との関連で特徴づけることが可能であり、便利である。各構成要素は以下のように定義づけることができる。

普遍性/安定性は、目標とする母集団をつうじて、そして十分に長期の期間を通して、規範が均一な仕方です、どれだけ深く理解されうるのかについて記述する。

適合性は、ひきだされた統計が利用者のニーズに対応するための可能な仕方はどれかを特徴づける。

明確性は、この規範からひきだされた統計を利用者がどれだけ十分に理解できるかを評価する。

検証可能性は、検証しようと要求する者すべてが、規範を十分な一致性をもって適用することを保証する必要を扱う。これは、提供される規範の精密性、規範にもとづくすべての記録へのアクセス可能性、実施における猶予の限界の明確な定義、個別の情報の提供者が与える基本的個別情報の正確性をチェックできる可能性、におよぶ。

収集可能性は、そのような規範を実施することを直接的あるいは間接的に要請される要素を、実際に収集できる法的あるいは文化的可能性の問題におよぶ。極度の儉約は、ひきだされる統計の統計的生産を行う間に、規範の選択によって課せられる予算への影響をふくめることに関わる。

一貫性は、同じタイプであれ³、異なるタイプであれ、異なる規範の一貫性に焦点をあてる。

完全性は、規範の全体集合に関わっており、利用できる規範が、利用者の主なニーズを満足させる統計を定義するのを可能にするかどうかである。

これらの構成要因は、統計規範の次のタイプ、すなわち、統計単位、分類、母集団、特性、および対象期間に関して、意味があることを容易に見ることができよう。最後のタイプについては、統計的要約、普遍性および検証可能性要因がいかなる場合でも保証され、収集可能性は無関係である。

注：幾つかの構成要因の間にはトレードオフがあり、ひとつの要因は、他の要因を減じる危険を持ちながらでしか、増大することはできない。例えば、極度の儉約は、一般には普遍性を低め、検証可能性を低めるだろう。

7 統計規範と統計の品質構成要因

さて、われわれは上の分析をさらに一層発展させて、統計規範のどの品質構成要因が、統計の品質構成要因に影響を与えるか、統計規範のどのタイプについて、このこ

³ 統計単位、母集団、分類、特性、対象期間と要約

とが意味を持つか、を示すことができる。統計の品質への総ての影響を分析するために、統計生産が規範に加えられる。

われわれは、統計規範のどのタイプに対して、この依存関係が妥当するかも同じく検討した。多くの場合に、品質の構成要因間のある所与の関係を述べることを選んだ理由を見いだすことは非常に容易ではあるが、誤解を避けるためには、われわれの選択を正当化する値はあろう。

表：規範と統計の品質構成要因間の連携

統計の品質		適合性	正確性	速報性 定期性	明確性、 アクセス 可能性	比較可 能性	一貫性	完全性
規範の 品質	普遍性	P,C,S, Ch,T	P,C,S, Ch,T		P,C,S, Ch,T	P,C,S, Ch,T	P,C,S, Ch,T	
	適合性	P,C,S, Ch,T,S						
	明確性				P,C,S, Ch,T,S			
	検証可 能性		P,C,S, Ch,T			P,C,S, Ch,T,S	P,C,S, Ch,T,S	
	収集可 能性、 極度の 儉約		P,C,S, Ch,T	P,C,S, Ch,T				
	一貫性						P,C,S, Ch,T,S	
	完全性							P,C,S, Ch,T,S
統計的 生産の 品質		X	X	X	X	X	X	

P=母集団,C=分類、S=統計単位,Ch=特性,T=対象期間,S=統計的要約

8 規範の品質についてどう報告するか？

統計規範の品質のこの分析は、品質の評価を容易にしより客観的にするなら有用である。標準報告に向けての主な方向を、以下に提案する。

普遍性と安定性：は何らか種類の知覚的な種類の手続きによって評価されうるし、試験サーベイあるいは回答者との面接からの報告を聞くことによって補足されうる。法的あるいは行政的理由によって課せられている標準化された語句あるいは規範を使

用することによって、それらの規範を知っている公衆に対して適用されるとき、高い標準化のレベルを保証するであろう。

適合性：は、とりあげられている規範とつながりを持つ統計の利用者が表明した必要を—それが満たされようがいまいが—を総合する。このことは、与えられた統計のすべてのタイプの利用者が表明した意見を合計する手続きだけでなく、異なる統計の利用者にとっての同じ手続きをも定義することを意味する。

明確性：利用者は、通常の言葉、言語、および文化的な参照を以て表された統計をよりよく理解することは非常に良く知られている。研究者あるいは学生に提供された規範は、それら利用者が規範の十分性に対する見方を表明し、それが唯一利用できるものだからというだけで利用しているのでなければ、アカデミックな研究論文、モノグラフあるいは統計で使われている概念に関して評価されうる。同じように、法的概念、行政的概念、テーマに関する概念（健康、疾病）、あるいは一般的コミュニケーションについてすぐに言及する他のあらゆる部類の利用者に対して、規範がどう容易に橋渡しをするか評価されるべきである。概念の名前の選択は非常に重要である。メタ情報の用意は、非常に有用ではあるが、明確性を訴えるには不十分なことは確かである。

検証可能性：この概念は、統計単位の定義について提起されるすべての問題—輪郭描写(delineation)、分類、特性の内容、母集団への算入あるいは除外、対象期間にふくめるか除外するか—が、安価に、理論的にユニークな仕方でも、しかるべき時間内に、に答えることができることを意味する。この情報が利用できず登録されている場合は、規範に付与される許容限界が満たされるなら、評価のために使用される。同じように、規範に付された標準にしたがってこの規範を適用するために使われる手続きは記録され、この手続きの実施は標準に従ってチェックされうる。これに加えて、回答者によって提供される個別の情報は、「真値」という微妙な概念を持つために、検査可能であるべきである。

収集可能性と極度の儉約：例えば試験調査や後には面接者の報告を聞くことを通じて、フォローアップ等に責任を持つスタッフ、規範によって請求される必要な情報を収集可能なのはどんな条件下なのか、特に遅れについてはどうか、回答者の拒絶、法的請求の数、統計法の下でそういったデータ収集を強制する可能性、回答者と国家統計局にとっての費用（収集されるべきデータの量）他が評価されるべきである。このことは、規範自体は直接的に収集可能であるべきことも、期待される概念の収集を可能にする間接的情報を収集することが可能なことも、意味しない。

一貫性：は同じタイプの規範の内部と、異なるタイプの規範の間との両方で考察されるべきである。同じタイプ内部では、共通の特徴とそれら規範の間で異なる特徴を考察することが不可欠である。特に、用語の（非）類似性は、意味の（非）類似性に

照らしてチェックされるべきであり、それらを説明し測定する可能性が評価されるべきである。異なるタイプの規範の間では、どの組み合わせが意味を持ち、どの組み合わせが避けられるべきかが考察されるべきである。

完全性：は、統計システムの全体的レベルで評価されうる。それは、一方では、利用者からの統計に対する登録された要求と、それら要求に与えられる適切なウエイトと優先度を考慮に入れて、他方では異なる領域で異なるテーマについての統計調査を行うことを可能にする使用可能な規範とを、比較するべきである

9 結論

われわれはここでは、統計の幾つかの側面を十分に正規化することが何故必要なのかを論じなかった。例えば、統計生産の幾つかの側面は、完全に標準化できる。実際に、部分的調和および全体的調和の間には、幾つかの継続性がある（Grais,1998を参照）。つつめていえば、統計の品質に対する要求が増大するときには、正規化への圧力が増加する、ということにしよう。これはヨーロッパ連合の統計家の経験である。統計規範にける品質の評価の方法は、統計体系（system）の品質を高める過程で、規範発展の過程と統計生産が十分に結びつけられていることを保証することによって、より一層の効率性とより大きな保証をもたらすだろう。

文献

- Eurostat (1998a), '*Definition of quality in statistics*', doc. Eurostat/A4/98/Quality/General/Definition rev.1
- Eurostat (1998b), '*Assessment of Comparability*', doc. Eurostat/A4/98/Quality/General/Comparability rev.1
- Eurostat (1998c), '*Assessment of Coherence*', doc. Eurostat/A4/98/Quality/General/Comparability rev.1
- Beekman, M.M. and Struijs, P. (1993). '*The Quality of Economic Concepts and Definitions*': Statistical Journal of the United Nations, ECE 10, 1-15. IOS Press.
- Boltanski L. (1970), '*Taxonomies populaires, taxonomies savantes: les objets de consommation et leur classement*', Revue Francaise de Sociologie, XI, 1970, 34-44
- Grais B., (1998), '*Statistical Harmonisation and Quality: The case of social Statistics*', paper presented at the 4th Mondorf Seminar on 'Harmonisation on social statistics and quality'

Thevenot L.,(1982) '*Measurement errors and coding procedure analysis*', International Meeting on Analysis of Sample survey data and on sequential analysis, Jerusalem,14-18 June 1982

Guibert B. Laganier J. and Volle M. (1971) '*Essais sur les nomenclatures industrielles*', *Economie et Statistique*, n°20, fev.1971, INSEE

5 サブ・グループ：規範の品質

最終報告

Eurostat

1998年7月

目次

節	表題
0	実務的要約
1	背景とサブ・グループの課題
2	規範（基準）；統計の生産と利用におけるその役割；規範の異なる種類
2.1	「規範」という語の定義
2.2	統計の生産と利用における規範の役割
2.3	規範の種類
3	サブ・グループがとりあげる範囲：検討されるべき規範のタイプ
3.1	とりあげる範囲内の種類
3.2	関連する規範の例
4	品質の概念とその統計規範への適用；考慮される利用者グループ、品質の評価のための基準
4.1	品質の定義
4.2	規範への品質の適用
4.3	規範の利用者
4.4	ニーズに応えること
5	実践における品質の測定
5.1	規範の品質の評価方法
5.2	規範の品質を評価するための可能な方法
5.3	方法 A—説明
5.4	方法 B—説明
5.5	方法とアプローチの種々の組み合わせの相対的な長所と短所
6	サブグループの結論と勧告
7	文献
8	サブグループのメンバー
付録 1	NACE 修正版 1 を用いての方法 A の作業例
付録 2	方法 A の品質基準の詳細チェックリスト
付録 3	NACE 修正 1 と方法 B を用いた標本の品質の報告

0 実務的要約

Eurostat が採用した全体的品質管理は、組織の全産物の品質が評価されることを要請している。規範（基準）は重要な産物なので、それらの品質を測定するための枠組みが要求される。1997年5月に、品質保証グループは適切な枠組みを描くためにサブ・グループをつくることを決めた。サブ・グループの任務は、規範の品質の定義とその品質を測定する手段を提案することであった。それは、品質の保障、品質の評価を何時、誰が行うべきかについて提案することまではふくまれていない。

この報告は、規範というタームを定義し、規範の分類体系を提案し、適切な規範の例を示す。そして、主な三つの利用者グループを認定し、彼らのニーズを考えながら、品質について ISO 定義をとりあげ、これを規範に適用する。

第5節は、規範の品質を評価するための関連する二つの方法を提案する。すなわち一つは、簡単で速やかに行われるが、かなり粗く主観的なものであり、もうひとつはより詳細で客観的であるが、またより時間がかかるものである。各方法の相対的な長所と短所が、両方法の作業例とともに示される（付録1と3）。

サブグループの結論と勧告は第6節に示される。それらは次のように要約できる。

- ・ 規範の品質を客観的に評価するのは難しい。
 - ・ 個々の規範の品質は、孤立的に評価されるべきではなく、規範のより大きな枠内でのその役割を考えることが重要である。
 - ・ 利用者のニーズを満たす規範の能力は、その品質を評価する上での最善の基準である。このことは品質に関する ISO8402 に従うものである。
 - ・ 規範の品質を評価するのに投入されるリソースはその規範を開発する費用に比例する。
 - ・ 将来、品質保証は規範を開発する手続きの総合された一部となるべきである。
 - ・ この報告で述べられた方法は、少数のキイになる規範の品質を評価するのに使用されるべきであり、得られた経験は次に、必要なら方法を改善するのに用いられるべきである。
 - ・ 統計の品質と規範の品質についての作業間の緊密な協力が必要である。
- Eurostat の品質管理者の役割は、この点で重要である。

1. サブ・グループの背景と任務

Eurostat の主要な任務の一つは、「偏らない、信頼の置ける、適切で、費用効率的な統計が、連合全体で生産されるように、一団の規範と方法」を開発すること、そして「連合の統計を生産する際の費用効率性ととともに、比較可能性を改善するために、

加盟国が連合の統計基準を採択するように準備し、一層発展させ、促進すること」¹である。

Eurostat が採用した総合的品質管理アプローチは、Eurostat の総ての生産物の品質が評価されるべきことを要求している。もし、品質の現在のレベルがわかるなら、品質の将来の改善を評価することがより容易になる。規範の質を測定する枠組みあるいは方法は存在しないので、この問題に取り組むための品質保証グループのサブグループを創設することが決められた。

このサブグループの任務は、規範の品質についての定義と、その品質を測定する手段を提案することであった。そこには、品質の保証、誰が Eurostat 内で規範の品質を評価する課題を遂行するか、あるいはそのような作業の時間的予定に関して、提案することはふくまれていなかった。狙いが規範を生産物として評価することであって、それらに責任を負う人々を評価することではないことが、ここで強調されるべきである。

多くの規範は、諸委員会内の長い討議の生産物であり、したがって理想的な解決ではなく妥協の産物である。品質評価の結果としてのそういった規範を変更するには時間がかかる。規範の品質が乏しいと評価されとしても、特に新しい立法が必要ななら、これは改善に向けての圧力を強めることができるだけである。

2. 規範（統計基準）；統計の生産と利用におけるその役割；規範の異なる種類

2.1 「規範」というタームの定義

辞典の定義；

規範-（大工の広場を意味するラテン語の *norma* から）、標準、パターンあるいはタイプ。

基準(Standard)；基準や、何らかのタイプ、事実、もの、基準(criterion)として役立つその他として設定された尺度。

サブグループの作業目的に関して、われわれは「規範：norm」と「基準 standard」という語を同義語として使うことに賛成する。そして、規範を「一貫した統計的データの収集、処理と配布のための枠組みに向けて貢献する、同意されるか承認された協定」と定義する。

2.2 統計の生産と利用における規範の役割

¹ 連合の統計生産に関する Eurostat の役割についての 1997 年 4 月 21 日の委員会決定(97/281/EC)

規範は統計の生産と利用にとって決定的道具である。すべての統計はその基礎にある概念や定義に依拠しており、多くの統計はデータを分類したり同質の集団を定義するために分類体系を利用する。通常は、設定され、承認された方法が、生のデータを統計にまとめるために使われる。調和のとれた規範は、統計を理解しやすくし、意味のある比較を、特に国際的レベルで、可能にすることで統計の有効性を大きく高める²。

2.3 規範の種類

規範を分類する仕方には、次のものをふくめて種々ありうる。

- (a) ・概念
 - ・分類体系—大きさ、タイプ、活動、位置、その他
 - ・単位（収集、処理、測定、生産）および単位の分類
 - ・特性—観察した変数
 - ・引きだされる概念、例えば、人口、GDP
 - ・要約的統計、例えば、平均、分散、四分位数
- ・生産
 - ・サンプリングの方法
 - ・データ収集と処理
 - ・計算と結果の表示
- ・背景
 - ・インフラストラクチュアと法的枠組み
 - ・秘匿性
 - ・習慣、伝統他
- (b) ・強制（法律かEUの規制によって強いられる）
 - ・望ましい（ガイドライン、勧告、最善の実践のマニュアル、紳士協定などに示される）
 - ・習慣的（書かれておらず、伝統的に使われている）
- (c) ・統計的生産物
 - ・データの収集
 - ・データの移送
- (d) ・ヨーロッパ統計制度(ESS:European Statistical System)内で使用されている
 - ・ESS内では使われていない

規範はまた、上の分類のすべての基礎にある用語に示され、この用語が多くの言語に翻訳されるべきときに、特に重要になる。

² "L'harmonisation des concepts et des definitions comme etape indispensable de l'integration des statistiques"
Michel Poulain, GÉDAP, Université Catholique de Louvain

ヨーロッパ勘定体系(ESA:European System of Accounts)³のようないくつかの法的なあるいは同意された枠組みを、単独の集められた規範、あるいは規範の総合化された集まりのいずれかで見なすことができる。もし、後者の場合なら、個別的規範が全体的枠組みの中で相互に影響しあう仕方を念頭におくことが重要である。規範の相互関係は種々のレベルにある。ある場合には、規範は実際レベルで非常に密接につながっており、他の場合には、そのつながりは、性質としてはより理論的であるかもしれない。

3. サブ・グループの対象範囲：検討される規範のタイプ

3.1 対象内の分類

サブ・グループの狙いは、ESS 内で使われる統計規範に集中しながら、すべての統計規範の品質を評価することに使うことのできる普遍的な枠組みを提供することである。したがって、検討されるべき規範の主な分類は、概念に関するそれである。すなわち、分類体系、単位、特性、および引きだされた概念である。（しかし、要約統計に関係するものは違う）。統計の生産に関係する規範、例えば第一次登録簿の使用、も検討される。強制的および望ましい規範の両方がふくまれる。それらはしばしば補足になるからである。それら規範は、統計的産物、データの収集あるいは移送に関連することもある。

ESS 内で使われる規範のうち、ごく少数のものだけを、厳密な意味で Eurostat が実際に作成しており、あるものは他の機関が使ったものから開発され、あるものは修正なしで採用されている。しばしば、Eurostat は他の国際機関とともに、世界規範を開発する際の一部を担っている。できるだけ最も早い機会にこの過程に関わることが望ましいのは明らかである。Eurostat が特定の規範に対してもつ管理の程度は、その規範の品質を評価するときに考慮されるべきである。

3.2 適合的な規範の例

3.2.1 分類規範

- ・ 規模-SME 定義（委員会の勧告）
- ・ タイプ-ESA 制度部門（金融機関、非金融会社、世帯他）
- ・ 活動あるいは職業-NACE 修正版,ISCO
- ・ 生産-CPA,CN,PRODCOM リスト
- ・ 立地-NUTS

³ 連合のヨーロッパ国民および地域勘定体系についての 1996 年 6 月 25 日の理事会規制(EC 2223/96)

- ・ 社会的および人口的-ISCED,ICD、社会指標

3.2.2 単位の規範

- ・ 統計単位、例えば、統計単位についての理事会規制(696/93)に定義されている企業統計についての単位、そして「世帯」のような社会統計に使われている単位
- ・ 測定単位、例えば購買力平価

3.2.3 特性の規範

- ・ SNA/ESA における離職、就業、他の定義
- ・ 人口・住宅センサス、および世帯サーベイの調整されたコアの変数
- ・ 労働力サーベイについての ILO の定義

3.2.4 生産の規範

- ・ 統計目的のための企業登録簿を作成する際の共同体の調整についての理事会規制

4. 品質の概念と統計規範へのその応用；考慮されるべき利用者グループ；品質評価のための基準

4.1 品質の定義

品質についての ISO8402 の定義は、「明言されているか暗示されているニーズを満足させる能力を持つ生産物あるいはサービスの特徴および特性の総体」である。統計の品質を決定する Eurostat 内で一般的に使われるのは7つの基準である⁴。

- ・ 適合性
- ・ 正確性
- ・ 速報性と時間厳守性
- ・ アクセス可能性と情報の明確性
- ・ 比較可能性（空間と時間に関して）
- ・ 一貫性
- ・ 完全性

回答負担をふくむ費用は、外的制約と考えられている。統計の品質についてのこれらの基準のすべてが、規範の品質の影響を受ける。そして、それらの修正版を、規範の品質を決定するときに使うことができ、したがってここには2様の道がある。

⁴ Eurostat Quality Assurance Group によって定義された

4.2 規範への品質の適用

規範の品質を評価するために提案された基準を以下にリストする。これらは、さらに付録2で拡張されている。幾つかのタームは、これが論理的なところでは結びつけられる。例えば、明確性、具体性とアクセス可能性は、すべて利用者に規範をどう伝えるかに関わるので、結びつけられる。

- ・ 適合性 — 使われる規範が提供者にとって適合的であり、意義があるなら、データを与えることはより容易である。規範はまた、統計の生産者と生産された統計の利用者にとって、適切であり適合的でなければならない。

- ・ 明確性、具体性およびアクセス可能性 — 規範はすべてのグループが容易に理解できるものであるべきであり、特に用語に関しては、他の言葉に翻訳するのが容易であるべきである。ときとして、概念を定義する幾多の文章があるべきである。というのは、ある単語の意味は翻訳によって、歪められるからである。規範はまた、利用者に対して、適切なフォーマットで容易にアクセス可能であるべきである。

- ・ 一貫性 — 一貫性には二つの側面がある。利用者が規範を正確に一貫した形で適用しているかと、規範が外的要請、例えば、データ保護や ISO の要請と一致しているか、（すなわち、規範を実施することが可能であるべきことと、またそれらが適切に応用されることをチェックできるべきこと、である）。

- ・ 収集可能性と柔軟性 — 特定の規範に関連するデータ収集の容易さ（と費用）が考慮されるべきである。規範の適用のされ方における柔軟性、例えば、国際分類体系をさらに国に分割する可能性、あるいは共通の目的に向けて積極的に動かす柔軟性、はそれに適用する規範という見地から、ある規範の品質を改善することができる。時間にわたっての適合する柔軟性はまた、規範がその適切性を維持する助けとなることができる。しかしながらまた、柔軟性は普遍性と安定性にマイナスの影響を与え、規範がそういった比較を行ううえで望まれる必要性に応える能力を減じる危険がある。

- ・ 普遍性と安定性 — 普遍性は国の間の比較に関わる。例えば、ヨーロッパの目的のためには、規範はすべての加盟国に適用されるべきである。もし、それらがまた国際比較を可能にするグローバルな脈絡で使用できるなら有効である。安定性は時間にわたっての一貫性に関わる。例えば、分類体系が時代遅れになるなど、規範の品質は時間とともに変化しうるものであるが、規範の変化は時系列の不連続性の理由になりうる。

- ・ 一貫性 — 異なる統計分野にわたって多くの目的に使うことができ、なお基本的ニーズに応える調整された規範が望ましい。そういった規範には広い範囲の利用者があるので、それらをすべてのニーズに対応させることはより難しいことを意識することが重要である。非常に特定した目的のために指定された規範は、

当初は、規範がごく限定された利用者のすべてのニーズに対応するために、高い品質のものであるように見えるかもしれない。関連する分野での規範の高い評価もまた重要であり、特に、国民勘定といった大きな体系においてそうである。非統計的（例えば、会計的）規範との両立がしばしば望まれる。

- ・ 完全性 - 完全性には二つの側面がある。規範の内部での完全性、すなわち、説明されているか測定されていることのすべての必要な特徴をカバーしているかどうか、と規範体系の完全性、すなわち、さらに何らかの規範がある統計を生産するために必要であるかどうか、である。すべての可能性がカバーされる場所では、分類体系にとって完全性が特別に重要である。そしてユニークネス、すなわち、あるものをひとつ、そしてただひとつの部門にだけ分類することだけが可能であるとする原理もまた重要である。

4.3 規範の利用者

規範の品質は声明されるか示されているニーズを満たすその能力によって決定される。統計的規範の生産者は以下に示す3つの主なグループ（これらグループの間にはいくつかの重複がありうる）を考えるべきである。

- ・ データ提供者
- ・ 統計の生産者 - 地域、各国、国際的レベルでの
- ・ 統計の利用者（中間および最終の） - Eurostat に関わる限り、利用者はさらに6つの部分グループ⁵ に分割できる。
 - ・ ヨーロッパ連合の機関、例えば、委員会、議会、協議会
 - ・ 加盟国内の政府機関、政府統計機関をふくむ地方、地域、全国機関
 - ・ 国際機関、例えば、国連、OECD
 - ・ 他を代表する機関、例えば、商業団体、労働組合、メディア
 - ・ 経営での利用者、例えば、企業
 - ・ 他の利用者、例えば、アカデミーや一般公衆

品質の全体的尺度を得るには、各グループの利用者を認定し、かれらのニーズを判定し、それにどれだけうまく満たされているかを見る必要がある。多くの利用者を持つ規範にとっては、どの利用者が、統計的過程に最も適合的かあるいは重要であるかを考察すること、そして彼らのニーズに集中するのが価値あることであろう。統計規範は、最初に意図した利用者のニーズに十分には対応するが、それに次ぐ、おそらく未知の利用者のニーズには対応しえないかもしれないことを忘れないことが重要である。したがって、規範がそのライフサイクル内に到達した段階が、その品質を

⁵ Eurostat Quality Assurance Group によって定義された

決定するときに適切であり、新しい利用者がわかったときに、この規範の品質を再評価することが必要であろう。

さらに、第四のグループ、統計規範を非統計的目的に使う者、例えば行政目的に統計分類体系を使う者を考えることも、ときとして必要である。それら規範の利用者の多くが未知である可能性は大きいので、彼らのニーズを評価することは難しいか不可能である。この試みの目的は統計規範の評価の手段を提供することであるから、主な焦点は、統計分野での規範の利用者のニーズにおかれる。しかしまた、適切なところでは、他の利用者、特に統計目的に使われるデータを提供する行政機関のニーズについての何らかの考慮をふくめるのがよい。

4.4 ニーズへの対応

統計の利用者と生産者はしばしば、データの供給者が与えようと意図している以上の詳しさを求める。常に、供給者に対する統計的（そして行政的）負担、したがって財政的費用を最小にしようとする圧力がある。優れた規範はここで助けとなる。したがってデータの供給者にとって重要な要因は、適合性、収集可能性および明確性であろう。

統計の生産者はまた、適合性、収集可能性および明確性を重要なものとする。特に、生産者がかれらのデータに規範を適用しなければならないときにそうである。彼らのデータをチェックするためには、彼らはまた普遍性、一貫性および時間にわたっての安定性を重要とする。もし彼らのデータを公表しなければならないなら、彼らは法的制約をもたまた考えなければならない。

統計の利用者にとって、規範の品質において重要な要因はもっと多様であろう。特殊な利用者は、統計を扱う経験を持っていて、普遍性、安定性、一貫性を求めがちであるが、統計の全般的利用者、偶然的利用者はおそらく、統計が彼らに語っていることを容易に理解できるように、明確性と適合性をより強調するであろう。

「ニーズ」というタームは、規範の利用者の三つすべてにとっては、費用と時間という要因をふくんでいる。ほとんどの要請に見合う規範は、すべての要請に見合うがその生産がかなり高価で時間がかかる統計よりも、より高い品質のものであると考えるかもしれない。

5. 実際における品質の測定

5.1 規範の品質を評価する方法

規範の品質を評価するためには、各種の方法が可能である。

- ・ 自己評価－これは可能なところでは実証的方法、例えば、分類体系のグループ化の内部あるいは相互間での分散の比較
- ・ 利用者のサーベイ－規範の利用者に、規範がどれだけニーズに応えているかを質問すること。規範がどのように改善するかに関する考えもまた、規範の将来的改善を考えるときには有効である。
- ・ 同輩による検討－例えば、競合する外部機関を招いて品質を評価する。これはまた、ベンチマーキング－すなわち、相互の利益のために他の機関との比較をする－の要素をふくみうる。

選択されるアプローチが可能な限り客観的であることが重要である。実証的方法は、最も客観的であるようだが、多くの規範に適用するのは難しいかもしれない。

任意の規範の品質を、その取り上げ率を観察することによって評価することも可能であろう。規範が関係者のニーズを満足させることが大きいほど、それが使われる可能性が大きくなる。

5.2 規範の品質を評価することが可能な方法論

ここでは二つの別個の方法を検討する。第1（方法 A）は、相対的に速やかで、やさしく、したがって低価格のアプローチである。第2（方法 B）は、規範の品質に関する総ての側面を、事前に決定した基準のチェックリスト（付録2参照）を使用して、よりはるかに徹底的に評価しようとするものである。両方法とも、自己評価、利用者のサーベイ、そして同輩による検討アプローチに適合させることができる。

これらの方法は、統計の品質の測定に使われる方法とは、数値得点を規範の発展の叙述と結合した報告をふくむ点で、いくらか異なる。規範の発展についての叙述は、それが有用な背景情報を提供し、規範が何らかの仕方でも開発した理由を説明する助けになるので、必要である。規範を開発するのに使われた過程は中立的な調子で書かれるべきであり、この過程の判定を企てるべきではない。これは、生産物の品質を測定する狙いを減じるからである。報告のこの部分はまた、規範の改訂が考察されるべきかどうかの評価を含むことがありうる。

数値得点を品質の絶対的尺度としてでなく、ひとちのガイドとして受け取るべきことが重要である。それらは、目的と利用の類似性が高い規範を検討するときに最も妥当する。これは、非常に特殊な目的やごく限定された数の利用者のためにだけ開発され、使用された規範は、利用者のニーズをより満たし、したがって多くの利用者と広い適用に供され、必ずしも「すぐれた」とはされない規範よりも、高い得点をうる傾向があるからである。

5.3 方法 A—説明

- ・ 規範の起源、変化、Eurostat による統制の程度をふくめた、規範の分類の決定
- ・ 規範の利用者を判断して、それらを、データ提供者、統計の生産者、統計の利用者にグループわけする。
- ・ 関連する規範、その規範が一部をなしているより大きな枠組みを考慮に入れて、それら利用者の各グループのニーズに規範がどれだけよく対応しているかを評価する。
- ・ 適切な場合には、その規範の全体的な品質を、再度、規範のより広い枠組みとのつながりで判断する。これは、規範の利用者の主な三つのグループの各々について得点を与えることで行うべきだろう。すなわち、次のようになる。

各グループにとって、	規範がすべてのニーズに応える	点数 = 5
	ニーズの 80% から 99% に応える	点数 = 4
	ニーズの 60% から 79% に応える	点数 = 3
	ニーズの 40% から 59% に応える	点数 = 2
	ニーズの 20% から 39% に応える	点数 = 1
	ニーズの 20% 未満に応える	点数 = 0

そして、0 から 15 点までの範囲で、規範に全体的点数を与えるために、3 つの点数が個別にあるいは加算して一緒に供することができる。しかし、極端な場合（例えば、5,5,0 の点数）には、単独の合計は規範の品質について誤った印象をもたらす危険がある。

- ・ 上述のステップを詳細化し、規範の発展についての簡潔な説明をふくみ、関連する規範との一貫性の必要といった制約を考慮に入れ、適切ならば、改訂のための勧告をふくむ報告を作成する。しかし、制約についての検討が、低い品質の点数を合理化する手段として単独に使われないように注意してもらいたい。

この方法を使った規範についての品質評価の作業例は付録 1 にある。

5.4 方法 B—説明

この方法は、方法 A の最初の二つのステップ（すなわち、規範の分類と利用者の判断）の次に、その規範のどの利用者を考え、彼らのにニーズがどのようにして考慮されたかの説明を加える。次のステップで、特定の規範の品質に関する異なった構成要素を詳細に検討する。

付録 2 に検討されるべき基準のチェックリストがある。

方法 A では、規範利用者の三つのグループが同じウェイトを持つものと仮定されているが、この仮定は必ずしも常に妥当するものではない。方法 B では、三つのグループに 0 から 15 点までの範囲内での相対的ウェイト (W_g) がそれぞれ与えられ、それ

らウエイトの合計 (ΣWg) は 15 点である。例えば、この規範が統計の提供者あるいは生産者にとって重要性は少ないが、統計の利用者にとって決定的なものであるときには、ウエイトはそれぞれ、2;2;11 になるのである。

付録 2 の基準リストは、次に各利用者グループごとに検討される。しかし、七つの基準がすべて、規範のすべての利用者にとって同じく適格的ではないことを認めなければならない。

ほとんど総てが、統計利用者にとって適格的であり、統計生産者にとっては大部分が適格的であるべきである。データ提供者のニーズはかなり異なるので、基準の幾つかは、このグループにとっては、わずかしき適合性がないだろう。

これを補うために、七つの基準の各々にウエイト (Wc) が、それらウエイトの合計が (ΣWc) の合計が 100 になるものとして付与される。特定グループにとって適合性の小さな基準は、したがってそのグループに対しては低い (あるいはゼロすらの) ウエイトを与えることが可能である。各基準に対する点数 (Sc) は、その基準の下にリストされた点数を参照して、 $0 \leq Sc \leq Wc$ として決定される。次に総合点数 (Sg) が計算され (すなわち、 $Sg = \Sigma Sc$)、加重された点数 ($SgWg$) を与えるその利用者グループのウエイトを乗じられる。三つすべての主要な利用者グループに対するこの過程がすべて終了すると、三つの加重された得点が総計されて、その規範の全体的点数 ($\Sigma SgWg$) が出される。この得点は 0 と 1500 の間にあるべきであり、100 で除すなら、方法 A から得られた点数と比較できる。

統計規範が非常に多様であるということは、各基準内の詳細な点の幾つかが、あるタイプの基準について、他よりもより適格的であることを意味する。それらの点数は、直接的に規範の品質と常には結びついているわけではなく (例えば、3.1.2 全ての利用者は規範に一貫した形で対応するか?)、品質にありうる欠点を示すかも知れない。ある場合には、他の点数を、それらが検討されている基準の範囲内にある限りにおいて、付録 2 にリストされている基準とともに、あるいはその代りに、考えてみるのが適切かもしれない。

数的点数とともに、規範の品質についての報告は次の点をふくむべきである。

- ・ 政策的背景、交渉やロビー活動による妥協、方法論的枠組み、そして関連する規範との一貫性というニーズのような制約をふくめた規範を創り出すために行われる手続きの詳細な叙述
- ・ 規範の実施とその後の発展についての適切な情報
- ・ その規範が他の規範とどう相互に作用しあい、それが規範のより大きな枠組みにどのように適合するか
- ・ より高い品質をもつかも知れない何らかの類似の規範の参照

しかし、制約についての考察は、乏しい品質得点を正当化する手段としてだけ使われるべきではないという点に注意してもらいたい。

この方法を使った規範の品質評価の作業例は付録3に与えられている。

5.5 方法とアプローチの種々の組み合わせの相対的長所と短所

各方法の主な長所と短所は次の通りである。

方法 A 長所

- ・ 理解するのに簡単であり、易しい。
- ・ 非専門家によって、わずかの訓練で使うことができる。
- ・ かなり速やかに行うことができ、したがって安い。

	方法 A	方法 B
自己評価	自己評価は自己の正当化になってはならない。	
	評価があまりに主観的であり、ほとんどの規範が 10 点周囲の点を与えられる危険がある。	再度、主観的である危険があるが、この方法のより構成的な性格が、これを少なくする。
利用者のサーベイ	このアプローチは利用者が認定されうことを仮定している。これはしばしば、規範の主な利用者にあてはまるが、ある者、特に未知のそして二次的利用者の見方は、考慮に入れられないだろう。	
	規範の利用者に接触する時間と費用が与えられているとき、方法 B のアプローチに基づく基準は、おそらくより適切であろう。	すべての基準を評価することは、非常に詳細な調査票を要求するので、選択する必要が恐らくあろうが、あまり多くを導入しないよう注意を払ってのことである。
同輩による検討	この規範をつくりだすか、あるいは「ライバル」の規範をつくりだすことに関与しなかった十分資格のある同輩を見つけだすことは難しいかもしれず、したがって全体的客観性は、常には保証されない。	
	この方法は、おそらく意味のある同輩にとっては、単純である。	

短所

- ・ 非常に主観的である。
- ・ 粗っぽい点数法である。

方法 B 長所

- ・ 品質の全側面の完全な評価である。
- ・ 比較目的のためのより正確な点づけ方法である。
- ・ より構成的なアプローチである。
- ・ 品質報告書のための適切な構造を提供する。
- ・ 方法 A より客観的であるが、なお、部分的に主観的である。

短所

- ・ より時間がかかり、したがって高価である。

方法とアプローチの各組み合わせはまた、潜在的な長所と短所を持つ。表は、各組み合わせがぶつかる主な問題を示す。

6. サブ・グループの結論と勧告

規範の品質を判断する際に検討すべき多くの異なる側面がある。この領域ではこれまで作業はごくわずかしち行われてこなかったが、幾つかの類似物が統計の品質についての作業の面からひきだされる。この報告書は、規範の定義、役割、分類を論じ、次に品質を定義し、その定義を規範にどう適用できるかを論じた。これの基本的特徴は、規範の利用者のニーズをよりよく判定できるように規範の利用者の確認とグループわけである。

実際に、規範の品質を明確に評価することは難しいようである。行われるべき十分に客観的な実証的評価を可能にするデータは、しばしばわずかあるかないかである。したがって、規範が利用者のニーズをどれだけ十分に満たすかについての意見の評価を基礎におくことが必要だろう。このことは、そういったどの評価も基本的には主観的であるが、評価を形成するために明白なフレームワークあるいはテンプレットを利用すれば、少なくとも客観性の要素を持ち込むことができる。規範の背景や発展についての情報が、評価をより意味あるものとするのに役立つことができる。

二つの方法が規範の品質を評価のために提案された。これらの方法は補足的なものと思われるべきことを意図している。方法 A は理想からほど遠いが、非専門家にとって、速やかで、安く、易しいという長所を持つ。方法 B はもっと全体的であり、したがって時間を要する。品質基準のチェックリストに従うことによって、個々の規範についての詳細報告を作成することが可能である。

幾つかの規範は明らかに他よりもより重要であり、利用者に最大の影響を与える評価に資源を集中すべきである。一般に、規範を発展させる費用は、その影響におおよそ比例することは期待できる（避けることのできない例外は特殊な場合として扱われる）。したがって、規範の開発費用の $x\%$ は、品質評価に宛てられるという一般的

規則を、資源の配分に使うことができる。(xはおそらく範囲 2-5 にあることになろう)。これは、使われる方法とアプローチの結合を判定するのに役立つだろう。規範の開発に関しては Eurostat のスタッフの間で、かなりの知識と経験があり、規範の品質を評価するときには十分な活用されるべきである。

統計の品質と規範の品質の間には二方向のつながりがあることに注意を払う必要がある。規範の品質についての報告を用意する際に、その規範を使って作成される統計をカバーする既存の品質報告が考慮されるべきであるし、逆もそうである。

要約すると、主な結論と勧告は、

- ・ 規範の品質は、客観的に評価することは難しいように見える。
- ・ 個々の規範の品質は、単独では評価されるべきではなく、その役割をより大きな規範の枠組みの中で考慮することが重要である。
- ・ 規範の利用者のニーズを満足させる規範の能力は、その品質を評価する最善の基準である。これは品質についての ISO8402 と調和する。
- ・ 規範の品質を評価するために投入される資源は、その規範を開発する費用に比例するべきである。
- ・ 将来において、品質評価は、規範を開発する手続きの全体の一部になるべきである。
- ・ この報告書で説明された方法は、少数の基本的な規範を評価するために使用されるべきであり、獲得された経験は次に、必要であればその方法を改善するために使うことができる。
- ・ 統計の品質と規範の品質についての研究の密接な調整が必要である。Eurostat の品質管理者 (Quality Manager) の役割はこの点で重要である。

7. 文献

規範の品質に関する直接的な文献は少ないが、脚注でふれた文献に加えて、次のものが少なくとも、参照に値する幾つかの材料をふくんでいる。

Abrahamse, A. (1996) Harmony by regulation: on the impact of EU Regulations on international harmonization of business statistics. Proceedings of the 82nd DGINS Conference – Statistical Business Registers: Problems and Opportunities, Vienna, Austria, 1996. Eurostat.

Beekman, M. and Struijs, P. (1993) The Quality of Economic Concepts and Definitions. Statistical Journal of the United Nations ECE 10, 1-15. IOS Press.

Daney de Marcillac, L. (1994) Lessons learned in implementing NACE Rev.1 in the European Union. Proceedings of the Eighth International Roundtable on Business Survey Frames, Heerlen, The Netherlands, May 1994.

Defays, D. (1995) Is harmonisation possible? Paper prepared for the conference on methodological issues in official statistics in Stockholm, Sweden, June 1995.

Depoutot, R. Quality definition and evaluation. Paper presented at the Fourth Seminar on the Future of European Social Statistics / Harmonisation on Social Statistics and Quality, 26-27 March 1998. This paper contains many further references that may be of interest.

Rainer, N. (1995) The Revised System of International Classifications. Eurostat.

8. サブ・グループのメンバー

Mr Tapio LEPPÖ

(座長)

Mr Alfonso ARPAIA

(Mr Paolo CARIDIが交替した)

Mr Fausto CARDOSO

Mr Louis DANÉY DE MARCILLAC

Mr Raoul DEPOUTOT

Mr Marcel ERNENS

(Mr Paolo CARIDIが交替した)

Mr Svein GAASEMYR

Mr Peter STRUIJS

Mr Steven VALE

(報告書作成者)

付録1－NACE Rev.1(改訂1版)を用いた方法Aの作業例⁶

このセクションは、方法の例解としてだけ見られるべきであり、この規範の品質の明確な評価を意味するものではない、ことを注意していただきたい。

背景の詳細－脈絡とコントロールの程度

NACE Rev.1は経済活動の分類体系である。その利用はヨーロッパ連合内では企業統計については強制されている。それは、データの収集と統計的生産物の両方に使用される。国連で作り上げられ、世界中での利用を勧告されたISIC Rev.3から採用されてきたものである。NACE Rev.1は、特に加盟国に関連する経済の領域、についてより詳細をふくむが、容易にISIC.Rev.3に総合することができる。したがって、Eurostatは、NACE Rev.1に対する部分的コントロールを持つとすることができる。

この規範の利用者

- ・ データの供給者－国との関わりでは、それはしばしば企業である傾向にあり、その役割では通常NACE Rev.1と直接的接触はない。
- ・ 統計の生産者－このグループはNACE Rev.1を、企業を区分し、同じ活動をしている企業をくくるために使う。これは集計的データの生産にとって決定的である。このグループには国とEurostatと世界のレベルの統計家が入る。
- ・ 統計の利用者－このグループは集計データを求める傾向があるが、しばしば経済のある部分の詳細を、要求しがちである。彼らのニーズは4.2節で述べた区分すべてをふくむ。

ニーズはどれだけ満たされるか？

- ・ データ供給者－このグループのニーズは、論理的、实际的で、負担を最小限にする簡単なシステムに向けられている。NACE Rev.1は、このグループのニーズのおそらく85%に対応するといえることができる。
- ・ 統計の生産者－このグループのニーズは、領域間の一貫性、普遍性、時間にわたる安定性、そして国のレベルでは、明確性、収集可能性、柔軟性、そして、コードがしばしば実業界が提供す活動の叙述からひきだされる必要があるように、適合性である。NACE Rev.1は他の経済データとの幾つかの一貫性を提供しており、そこでは、他の分類体系、例えば生産物についてのCPC、CPA、国際貿易のためのHSとCNと連結している。それはEU全体で使われており、世界で使わ

⁶ この規範についての一層の情報については、「NACE Rev.1：ヨーロッパ連合での経済活動の統計分類」

れている ISIC Rev.1 と一貫性を持つ。それは、企業統計のより広い体系の重要でうまく組み込まれた一部であるが、また他の領域でも使われている。それは短中期の期間にわたってかなり安定しているが、経済発展が、長期的にはさらなる改訂を必要にすると思われる。それは、経済のすべての分野をカバーする総合的体系であり、国のレベルでは、それらが NACE Rev.1 のクラスに集計される限りで、一層の分割というその他の柔軟性を可能とする。それはかなり明確であり、説明に基づいてコードをあてはめる必要のある利用者にとっては容易にアクセス可能である。高いレベルの正確性を持ってこれを行うことのできる多様なコンピュータ・パッケージがいまでは利用可能である。したがって、NACE Rev.1 は、このグループのニーズの 90%に見合うと思われる。

- ・ 統計の利用者－利用者となる可能性のある者は広くにおよぶので、一般化することは難しいが、

適合性、一貫性、普遍性、安定性は重要と思われる。NACE Rev.1 はこれらのニーズにかなり応えるが、ときどき、特に民間部門から、この体系が提供するよりもより詳細あるいは異なる構成への分割に向けての要求がある⁷。したがって、おそらくニーズの 85%は満たされている。

全体得点

節 5.3 の点数制度を使用すると、NACE Rev.1 は 4,4,3 の点数で、合計点は 11 である。

過去そして将来の発展

NACE Rev.1 は経済分類の総合体系の一部として発展した。それはまた、ISIC Rev.3 のような他の国際的な活動分類との一貫性をもつ。それは、EU のすべての加盟国からの専門家の間での長期にわたる詳細な協議の後に作成された。したがってそれはある程度は妥協の産物である。それは統計の領域の内外で多くの広い適用をもたらしている。それは、そういった応用に対して常には最適の分類体系では必ずしもないかもしれないが、しばしば利用されるのは、単純に既に関連されていて、広く利用可能であり、認められているからである。

経済的、技術的発展は、どんな活動分類も折々改訂される必要があることを意味する。それは、意味のある時系列統計をつくり出すために、時間にわたっての一貫性と

(NACE Rev.1 : Statistical classification of economic activities in the European Community", Eurostat 1996

⁷ "Alternative Aggregation and the Standard Industrial Classification" Shaila Nijhowne, Statistics Canada for the 12th meeting of the Voorburg Group on Service Statistics, Copenhagen, 15-19 September 1997. 経済分析をする目的にとっての共通の集計についての 1996 年 2 月 8 日の委員会の勧告 Q1 No.138,16.2.1996 も参照

いうニーズとバランスしていなければならない。統計の利用者の範囲が広いことによって、活動の異なる定義とグループ分けに向けての圧力が常にあるが、誰でもを満足させることは不可能である。現在では NACE Rev.1 を改訂することへの強いニーズはないように見えるが、おそらく 5 から 10 年での改訂のための計画をスタートさせるのが適切であろう。

付録2－方法Bに対する品質基準の詳細リスト

1. 適合性

- 1.1 この規範の利用者は誰か？
- 1.2 この規範の利用者のニーズは何か？
- 1.3 この規範はこれらのニーズに適合的か？
- 1.4 この規範は一般的に受け入れられるか？
- 1.5 この規範は非統計目的のためのものか？
- 1.6 もしそうなら、この規範はそれらの目的にとって適切か？

2. 明確性、具体性およびアクセス可能性

2.1 明確性

- 2.1.1 その規範は、すべての利用者にとって明確かつ容易に理解できる仕方で具体化されているか？
- 2.1.2 技術的専門用語は必要最小限にとどめられているか？
- 2.1.3 その規範の根拠が明確に説明されているか？
- 2.1.4 その規範の背後にある論理は明確に説明されているか？
- 2.1.5 その規範の利用者に対して、ガイダンスあるいは技術的評注が提供されているか？
- 2.1.6 それらの評注は明確かつ総括的か？

2.2 具体性

- 2.2.1 その規範の具体化において、正確で承認された術語が使われているか？
- 2.2.2 その規範の具体化は既存の慣例に従っているか？
- 2.2.3 他の規範あるいは立法の参照は実際に正しいか？
- 2.2.4 その規範はすべての公式の言語に翻訳されているか？
- 2.2.5 規範の翻訳は技術的に正確であるか？

2.3 アクセス可能性

- 2.3.1 その規範はどう配布されているか？
- 2.3.2 その規範は利用者にとって適切なフォーマットで容易に利用可能になっているか？
- 2.3.3 潜在的利用者がその規範の存在を知るようになっているか？

3. 一致性(Conformity)

3.1 規範との一致性

- 3.1.1 すべての利用者がその規範を正しく適用しているか？
- 3.1.2 すべての利用者がその規範を一致性がある形で(consistently)適用しているか？
- 3.1.3 すべての利用者がその規範を正しく／一致して適用していないとすれば、その理由は何か？
- 3.1.4 すべての利用者がその規範を正しく／一致して適用していないとすれば、その結果は何か？

3.2 その規範と外的要請との一致性

- 3.2.1 その規範は、各国および国際的な法的要請を考慮しているか？
- 3.2.2 その規範は秘匿性要請を考慮しているか？
- 3.2.3 その規範は ISO の要請に対応しているか？ すなわち、
その規範を実際に実行することは可能か？
その規範が正しく適用されていることをチェックすることは可能か？

4. 収集可能性と柔軟性

4.1 収集可能性

- 4.1.1 その規範を適用するのに必要なデータを収集するのがどれだけ容易か？

- 4.1.2 その規範はデータ提供者にどのレベルの回答者負担を課すか？
- 4.1.3 その規範を使用することに伴う財政費用は少ないか？
- 4.1.4 その規範は、統計の生産を合理的な期間内で可能にしているか？

4.2 柔軟性

- 4.2.1 その規範を実施することのできる仕方に何らかの柔軟性はあるか？
- 4.2.2 もしそうなら、この柔軟性に対して明確に規定された限界があるか？
- 4.2.3 その規範のすべてあるいはある部分の実施を延期するような不面目を認めてきたか？
- 4.2.4 もしそうなら、それらは明確に規定された期間についてか？

5. 普遍性と安定性

5.1 普遍性

- 5.1.1 その規範は以下すべてに使われているか？
 - a) ヨーロッパ連合
 - b) ヨーロッパ経済地域
 - c) 他の国のグループ
 - d) 世界
- 5.1.2 その規範は、上述の地域で同時に適用されているか？

5.2 時間にわたっての安定性

- 5.2.1 その規範は何時作られたか？
- 5.2.2 その規範の有効性は時とともに変化しているか？
- 5.2.3 その規範の適用範囲は時とともに変化したか？
- 5.2.4 その規範は、その導入以来改訂されたか？
- 5.2.5 もしそうなら、以前のバージョンとのリンクを作ることは可能か？
- 5.2.6 時系列統計の生産に対して改訂はどんな影響を与えたか？

6. 一貫性

- 6.1 その規範は統計のどの分野で使われるか？
- 6.2 その規範はより大きな規範の体系（例えば、ESA）の一部か？
- 6.3 その規範は他の統計的分野に使われている規範と互換性を持つか？
- 6.4 その規範は他の諸国あるいは国際機関で統計目的のために使われている規範と互換性を持つか？
- 6.5 その規範は、非統計的規範（例えば、会計基準）と互換性を持つか？

7. 全部性

7.1 規範の全部性

- 7.1.1 すべての規範-その規範は説明されているか測定されているものの総ての属性をカバーしているか？
- 7.1.2 分類体系-その分類体系は、分類される各項目に対して単一独自の分類になっており、すべての可能性をカバーしているか？

7.2 諸規範の体系の全体性

- 7.2.1 この規範がその一部になる規範の体系に何らかの空白はあるか？

付録3－方法Bを使った品質報告例-NACE Rev.1

このセクションは、方法を使って作成されうる報告のタイプの例解としてだけ見られるべきであることを注意していただきたい。これは、その規範の品質の明確な評価を意味するのではない。それは品質評価の数学的部分に主として焦点をおいているので、この方法の下で作成される完全な報告書のように包括的なものではない。

背景の詳細－脈絡とコントロールの程度

NACE Rev.1は経済活動の分類体系である。この使用はヨーロッパ連合内では企業統計にかんして強制されている。それは、データの収集と統計的生産物の両方に使用される。国連で作り上げられて世界中での利用を勧告されてきたISIC Rev.3から採用されてきたものである。NACE Rev.1は、特に加盟国に関連する経済の領域、についてより詳細な内容をふくむが、容易にISIC Rev.3に総合することができる。したがって、Eurostatは、NACE Rev.1に対する部分的コントロールを持つということが出来る。

NACE Rev.1の利用者

- データの供給者－各国との関わりでは、それはしばしば、役割的には通常NACE Rev.1と直接的接触はない企業である。企業はしばしばその活動の記述を提供することを単に求められるだけである。しかし、それらはなお規範の利用者である。というのは、企業が提供する記述は、人手あるいは機械によってコード化されなければならないからである。企業が使う記述や用語を対処する分類体系の容易さが、したがってコード化の正確性に比例する。記述が単一の分類にうまく入らないなら、企業の回答負担を大きくすることになる企業との再接触をするか、誤ってコード化する危険があるかであり、したがって、将来不適當なサーベイを受けるかもしれないことになる。だから、明確性、具体性および収集可能性は企業にとっての重要な要因であると思える。

多国籍の脈絡では（例えば、ESS内部で）、データの提供者はしばしば国家統計機関あるいは他の国家的あるいは国際的組織である可能性がある。データ供給者としての役割においては、NACE Rev.1の利用者は、適切性と普遍性に最大の強調をおく。というのは、彼らのデータの供給する先は標準的枠組みとして、すべての人によって利用され理解される必要があるからである。

- 統計の生産者－このグループには国、Eurostatおよび世界レベルでの統計家がいる。このグループは、NACE Rev.1を企業を区分し、同じ活動をしている企業をくくるために使う。これは集計的データの生産にとって決定的であり、この分類体系やガイドの注記が、明確、正確そして容易にアクセス可能であることを必要とする。彼らの仕事は、企業にコード化するのに必要なデータを収集し検査するのが容易なら、容易になる。

諸領域の間で一貫性と時間にわたる安定性の両方が、彼らが学習しなければならない異なる体系の数を減らすことで、助けになることができる。

統計の利用者—このグループは集計データを要求する傾向がある。しばしば経済のある部分に対してはある詳細を要求するのであるが。利用者グループについての Eurostat サーベイが識別した六つの利用者グループはすべて、ある程度まで規範 NACE Rev.1 の利用者である。

- (i) ヨーロッパ連合の機関、例えば委員会、議会、理事会。
- (ii) 加盟国の地方と地域、および国家統計機関をふくむ国家の、政府機関。
- (iii) 国際機関。例えば国連、OECD。これら利用者の基本的ニーズは普遍性、安定性、一貫性である。
- (iv) その他を代表する団体、例えば商業団体、組合、メディア。
- (v) 産業の利用者、例えば企業。
- (vi) 他の利用者、例えば、アカデミーや一般大衆。これらグループにとっては、適切性、明確性、アクセス可能性がより重要であろう。

利用者グループのウエイト

NACE Rev.1 は、統計の生産者と利用者にとってかなりの重要性を持つが、データの提供者にとっては重要ではない。したがって利用者グループのウエイトは、以下のようになる。

データ供給者	2.5
統計生産者	6.5
統計利用者	6.0
$\Sigma W g$	15.0

各利用者グループにとっての品質の評価

これは、可能なところではどこでも、その規範の利用者が提供する情報—例えば利用者サーベイを通じて—に基づくべきである。

1 データ供給者

適切性—これは経済活動をコード化することを求められるデータ提供者にとっては非常に重要である。使われる分類表は論理的であり、データ提供者が認知している経済活動の自然な下位区分を反映しているべきである。そうでない場合には、データ供給者は分類表を理解し活動を正しく区分するために長時間を要して負担が増大する。NACE Rev.1 は一般的にはこれらのニーズを満たすが、他の分類表と共通に、混合した活動をしている企業を区分するときには問題を持つことがある。したがってウエイト=25、得点=21である。

明確性、具体性およびアクセス可能性-

利用者各グループのウエイトと点数は表の形で要約されるべきである。すなわち、

	ウエイト	点数
適切性	25	21
明確性、具体性、アクセス可能性	5	3
一致性	10	8
収集可能性と柔軟性	30	25
普遍性と安定性	10	7
一貫性	5	4
全部性	15	12
	100	80

総合点

統計の生産者と利用者にとっての総点がそれぞれ 85 と 70 と仮定すると、合計点は、

$$(2.5*80)+(6.5*85)+(6*70)=1172.5$$

背景情報

1. NACE Rev.1 を創るために従った手続きの、政策的背景、方法論的枠組み、関連する規範との一致性の必要といった制約をふくめた叙述。

NACE Rev.1 にとって、ISIC Rev.3 や調整されたヨーロッパ分類体系の発展のような他の国際的活動分類とのある程度迄の一致性という要望は注記すべき重要な要因である。一致性と比較可能性の強化という点での長所は、分類がそれらの制約のために最適ではないという短所と比較することができよう。

NACE Rev.1 を作成する作業は 1986 年にスタートした。Eurostat とヨーロッパ連合加盟国からの専門家委員会が設置された。この委員会は ISIC Rev.3 (この開発に Eurostat もまた大きく関与した) を出発点とし、それをヨーロッパのニーズに当てはめた。結果として、NACE Rev.1 はより高いレベルでは ISIC Rev.1 と一致性を持ち、低いレベルではこの一致性を保持するように集計することができ、他方で、必要などころでは、ヨーロッパの目的のためにより細かく分割することが可能である。理事会規制 1990 年 10 月 9 日 3037/90 は、経済活動に関するデータをまとめ、移送するときには、1993 年 1 月 1 日から NACE Rev.1 を使うように EU 加盟国に強制した。

2. NACE Rev.1 の実施とその後の発展に関する関連情報

NACE Rev.1 は、今ではすべての加盟国と幾つかの他のヨーロッパ諸国で実施されている。これは、国の目的にとって必要なところでは、標準の4桁コードに最大5桁をつけるという柔軟な対応を加盟国に許すことによって一層容易にされた。NACE Rev.1 についての情報とその利用のためのガイドをふくむ出版物が加盟国の代表の協定に従って出版された。それら代表者は SPC(NACE)と呼ばれる統計計画委員会 (SPC: Statistical Programming Committee) の小委員会の形で定期的に会合した。NACE Rev.1 の発展に関する問題が、判例法に基づく規則をふくめて、これらの会合で同意された。

3. NACE Rev.1 が他の規範とどう相互関係を持ち、それが規範のより大きな枠組みにどう組み込まれるか。

NACE Rev.1 は包括的な経済分類の一部である。それは、CPA や Prodom List のような生産物分類、そしてまた商業目的のために商品进行分类するために使われている結合分類表 (the Combined Nomenclature) と密接にリンクしている。NACE Rev.1 はまた、ESA のような規範のより大きな枠組みの重要な構成部分であり、多くの統計分野、例えば企業で、社会的に、そして地域的に、使われている。

4. より高い品質を持ちうる他の類似の規範の参照。

NACE Rev.1 と北米で使われている経済活動分類体系である NAICS との間で比較はできる。二つの体系間の主な違いは、NAICS は特にその目的のために創られた方法を使って、既存の国際分類を参照せずに開発され、従って技術的にはより適切な体系であるかもしれないが、NACE Rev.1 で可能な国際比較可能性のレベルを欠いていることである。

6 国際統計機関における統計生産物の品質¹

Raul Depoutot²

要約

過去4年間にEurostatで行われた品質に関する作業に基づいて、本論文は、国際統計機関における品質の分析の特殊性を示す。本論文は規範、記録簿、統計、統計分析や勘定体系のような主なタイプの統計的生産物の品質に対する一貫性を持った分析の枠組みを提案する。参考用の用語集に基づいて、論文は、品質報告の正規化した手続きが、完全な調整へのひとつの選択肢としてどう使われるかを説明し、品質についての信頼できる尺度のために、どんな条件が必要かを強調する。統計的生産物の品質についての測定体系も、可能なところではEurostat内部でかヨーロッパ連合加盟国代表との間で同意された提案に基づいて、示される。

キーワード： 品質、国際統計、比較可能性、統計生産物

1. 序

アメリカ合衆国連邦統計政策・基準局(1978)、カナダ統計局(1992)、スウェーデン統計局(1994)といった他の統計機関における類似の作業の蓄積をふまえて、過去4年間にEurostatで行われた品質に関する作業に基づいて、本論文は国際統計機関における品質の分析を提示する。われわれは規範、記録簿、統計、統計分析や勘定体系のような主なタイプの統計的生産物の品質に対する一貫性をもつ分析の枠組みを提案する。参考用の用語集に基づいて、論文は、多様な統計的生産物の品質報告の正規化した手続きが、どう実施されるかを説明する。品質の測定が、どう完全な調整に対しての選択肢になるかを示し、国際機関における品質管理の特殊な問題を示す。

2. 用語

2.1 規範

統計規範には異なるタイプのものがある。それは、統計単位、母集団、分類、属性、対象期間、統計分布の要約(例：諸平均、総計、分散、分位値)をふくむ。より一般的には、統計規範は、第1に、サンプリング理論の数学的形式主義の内部で現実を描く道具であり、

¹ Proceedings of the Joint IASS/IAOS Conference: Statistics for Economic and Social Development, september 1998

² Raul Depoutot, Eurostat room C5/96, Batiment Jean Monet L2920 Luxembourg, raul.depoutot@eurostat.cec.be

第2に、統計家が統計利用者と伝達しあうときに強制的な計器であり、第3に、基本的な個別情報を提供するもの（個人、企業）と伝達しあうための参考になる枠組みである。

統計規範に関する総ての問題は、伝達問題（透明性、客観性、メッセージの効率、情報の費用、文化的背景の影響、他）の特殊な場合と見ることができる。

2.2 統計

どんな統計も、対象とする母集団—それ自体は、統計単位の選択によって、そして母集団を包括的に定義する分類とによって定義される—、対象期間、特性、関心の対象になっている属性（平均、モデル、合計、指数、他）の分布の要約、測定過程、与えられた特性の名称、そして何らかのメタ情報によって、定義される。

2.3 生産過程

生産過程は、概念の代理(Proxies)の選択、推論戦略（すなわち、標本設計とモデルの組み合わせの選択）、何らかのメタ情報を伴った（データ収集手続き、収集されたデータの処理、配布作業からなる）測定過程によってなりたっている。

2.4 統計的勘定

統計勘定の体系は、統計規範の体系、ある概念の数量化についての何らかの仮定、およびそれら規範によって定義される統計、からなる。

2.5 分析

分析は、諸特性をリンクし、その相互影響を数量化するモデルからなる。このモデルは、対象期間にわたって妥当する統計単位の選択と母集団を包括的に定義する分類によって、定義される対象母集団に適用される。モデルでの特性のリンクは、個別の統計単位についてか、統計に集計されるかした属性について行われる。ここで考察されるモデルの特性とおそらくモデルのパラメータにはある名称³が与えられ、測定過程からひきだされたデータの基づいて数量化され、利用者のために、おそらくは観察された特性についての何らかのメタ情報で補足される。

3. 統計における品質の定義

3.1 各国レベルでの品質の定義

品質を評価するための通常の4つの構成要因として、適合性、正確性、速報性およびアクセス可能性がある。追加として、異なる統計についての一致性(consistency)を強調するこ

³ 例えば、「俸給水準に関する労働需要の弾力性」

ともできる。このタイプの定義は、国連（1983）、カナダ統計局（1992）、オランダ統計局（Beekman and Struijs、1993）、スウェーデン統計局（1994）、Deren and Elvers（1997）が与えている。

3.2 国際レベルでの品質の定義

3.2.1 国際的統計活動の特殊性

集計可能性と比較可能性：国際統計機関は、国家間で比較可能な情報を生産し、また多くの国にわたる集計値を生産することを狙って、各国統計を収集しなければならない。このことは、統計が比較可能であるか、比較可能性のために改善できることを必要とする。これは、事実、実践においての非常に強い要請であり、利用者のために不可欠である。比較できない統計は、国際レベルでは非常にわずかしか利用されない。

部門的情報と一貫性(coherence)：統計の国際的標準化の過程は、典型的には分野別に展開されてきている。ワーキンググループや委員会が、農業、産業、あるいは就業や環境その他といった特定の論題などの特定分野の統計と取り組むために設置されている。この過程は極端に長く、ときとして同じ国の異なる政府統計部門間の複雑な調整構造を必要とする（典型的には、経営に関する統計の責任は、諸省の異なる統計部門と中央統計局とがともに負っている）。このことは、関連する統計を扱う他のワーキンググループとの調整を同時に行うことを極端に難しくしている。したがって、統計の一貫性は保証されない。

3.2.2 品質の定義

われわれの作業は、品質の1986年の定義、ISO8402に基づいている。それは、品質とは「明言されているか暗示されているニーズを満たす生産物とサービスの特性と属性の総体」であると述べている。利用者とその期待をリストし優先させた後で、われわれは上述の属性を少数の同質の構成要因に区分する。すると、統計の品質は、7つの構成要因：適合性、正確性、速報性と規則性、アクセス可能性と明確性、比較可能性、統計体系の一貫性と完全性、からなるものとして定義される。

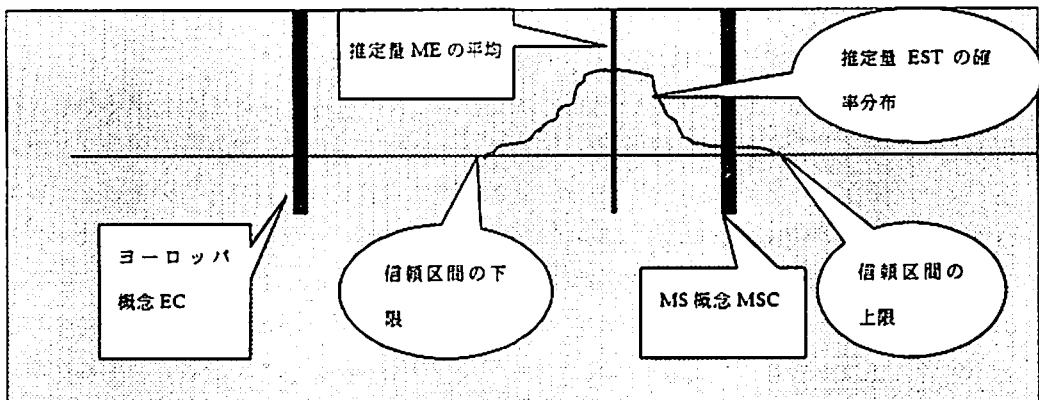
適合性は、利用者のニーズの充足に対応する。これを評価するには、利用者とその期待の確認が必要である。推定値の正確性は、推定値と母集団の（未知の）真値との接近度として定義される。結果の配布における速報性と時間厳守性とは、ひんばんにそして事前に公表した日に時間どおりに公表される最新の数字を獲得するという大部分の利用者のニーズに対応する。情報のアクセス可能性と明確性とは、利用者にとって、望む形態で、容易にアクセス可能で、十分な記録になっている、という統計の数値を特徴づけている。統計を利用し解釈する際の援助もまた、提供者から手近に与えられるものであるべきである。統計は、空間や時間にわたる基礎にある特性に関する数値について信頼できる比較を可能にするなら、比較可能とよばれる。統計が異なる出所から、特に異なる頻度の統計的サー

ベイから生産されているときには、互いに明確に関連づけられているか、あるいは少なくとも相互に矛盾しないメッセージを利用者に伝達する限りにおいて、統計は一貫性を持つ。統計が入手可能な分野が、その統計体系の利用者が表明するニーズや優先度を反映しているなら、統計体系は完全性を持つということが出来る。費用、すなわち回答者の負担と統計機関が利用できる必要な予算とは品質の構成要因とはみなされない。

3.2.3 比較可能性への一層の注目

そこで、2つの統計の比較可能性を、2つの側面から見る。すなわち、1-これらの統計によって測定される未知の真値の比較可能性と、2-比較に使われる指標を十分に数量化する位置にいるための、真値をめぐる推定値の正確性。

第2点は、各統計について、および各加盟国ごとに独立して品質報告で取り扱われる（以下を見よ）。第1点は、品質報告においてどこでもとり組まれているわけではない。これは、比較可能性の次元の中心部分である。EU に関しては、各加盟国の統計が基礎におくことのできるヨーロッパ参考文献があるべきである。



上の図において、われわれは、EU 規範を使っていて、どんな種類のどんな誤差をも伴わない統計の理想的な目標の値を、EC で表している。MSC は、理想的であるが、国家の規範による目標に対応し、EST は測定過程（推定量の定義、標本設計、非標本的測定過程）からひきだされる推定値に対応する。われわれは、推定量 ME の平均と MSC の距離は、サンプリングの偏りと非サンプリングの偏りからなっていること、確率分布と信頼区間は、サンプリングと非サンプリング（U で表現）のランダムな見地を考慮に入れること、を認めている。

加盟国 A と B とを、検討中の特性について比較するときには、われわれは、比率 $R(A,B)=EST(A)/EST(B)$ を考える。略号を使うと $[MSC(I)-EC(I)] = DC(I)$ であり、 $[ME(I)-MSC(I)] = MB(I)$ である。

容易に以下のように示される。

$$R(A,B) = \frac{EC(A)}{EC(B)} \left(1 + \underbrace{[DC(A)/EC(A) - DC(B)/EC(B)]}_{\text{概念の調整の欠如に由来する比較可能性のバイアス}} + \underbrace{[MB(A)/EC(A) - MB(B)/EC(B)]}_{\text{測定の調整の欠如に由来する比較可能性のバイアス}} + \underbrace{U(A)/EC(A) - U(B)/EC(B)}_{\text{測定に由来する変異}} \right)$$

概念の比較可能性の欠如についての調整をモデル化する枠組みを提案することは可能である (Arondel and Depoutot, 1997)。「比較可能性の調整」は、母集団の構造における違いの調整と概念の違いの調整とを統合する。測定における比較可能性の欠如の調整とは、単純にバイアスの修正のことである。

4. 他の統計生産物の品質定義

4.1 国際機関の統計生産物

統計生産物は、最終生産物（説明なしの統計、説明付きの統計、経済勘定と分析）と、中間生産物（統計規範、要求へのサービス、支援、とノウハウの交換）からなる。

4.2 それら生産物の各々についての一貫性という品質の定義の仕方

統計分析はモデル化の追加的積み重ねが加えられた統計とみなすことができるから、統計分析の品質は、次の要因、すなわち、適合性、正確性、速報性、アクセス可能性と明確性、比較可能性、一貫性、透明性、公平性との関連で定義できるものと考えることが提案されている。公平性(Impartiality)とは、説明モデルの選択が、非常に頑健な方法によっており、十分正当化されるべきことを意味する。透明性(Transparency)とは、著者の各判断について、仮説と著者が行った科学的論証との間の区別について、読者が気づくようにすることを著者に要求するものである。これは、モデル自体の評価、あるいは確認のために使われるテストの妥当性、あるいはモデル選択の手続きに適用される。

中間生産物に関しては、高い品質の統計を獲得するのに必要な属性の集合として分析できる。統計規範の場合には、統計の品質が3つの項目、規範の品質、推論の品質と生産過程の品質、の合計と等しいことを同じく注意することが有効である。このことは、統計規範の品質として次の構成要因をもたらす。すなわち、適合性、普遍性と安定性、明確性、検証可能性、収集可能性と極度の儉約、一貫性、完全性、である。普遍性/安定性(Universality/Stability)とは、規範が、狙いとしている母集団を通じて、そして十分に長い期間を通じて、同じ仕方ですべて理解されうるかを示す。検証可能性(Verifiability)は、検証を要求するすべての者が、規範を十分な一致性(conformity)をもって適用することを保証する必要を扱う。これは、提供された規範の正確性、規範についてのすべての記録へのアクセス可能性、実施における許容限界についての明確な定義、個別情報の提供者が与え

る基本的個別情報の正確性をチェックできる可能性、を扱う。収集可能性(Collectability)とは、そういった規範を実施するときに、直接的あるいは間接的に要求される要素を実際に収集できる、法的小よび文化的可能性の問題を扱う。極度の儉約(Parsimony)は、ひきだされた統計の統計的生産過程で、規範の選択によって課される予算への影響を組み入れることである。適合度、一貫性および完全性は、統計に対してと同じ仕方で定義される。

同じアプローチを使って、記録簿の品質を次の構成要因からなるものとして定義するのが論理的である。すなわち、カバレッジの全面性、ふくまれる個別情報の信頼性、統計単位の一致性、速報性、である。

同じように、勘定体系の品質は、規範の体系の品質、何らかのモデル化の基礎になっている分析の品質、および結果たる数量化の品質、の合成である。そこで、これは次の構成要因を導く。すなわち、適合性、正確性、速報性、アクセス可能性と明確性、比較可能性、他の統計との一貫性、透明性、公平性、である。次に、適合性は、選択された規範が利用者のニーズに対して十分性を持つかに焦点をあてており、統計規範についてと同じ構成要因(普遍性と安定性、概念の明確性、検証可能性、収集可能性と極度の儉約、勘定体系内部での一致性、その体系の利用者の期待からみでの完全性)に分割することができる。行政データからひきだされる統計の品質は、同じ構成要因を使って叙述できることを明瞭に声明すべきである。すなわち、もしデータが個別的であるなら、統計の品質に関する構成要因で十分であり、そしてもし統計が集計された行政データからひきだされるなら、その情況は、統計分析の品質を測定することと極めて類似している。

5. 統計的生産物の品質の測定

5.1 利用者と品質

Eurostat の生産物の利用者は、以下のように区分することができる。

－機関：ヨーロッパのレベルでは、委員会、理事会、ヨーロッパ議会、ヨーロッパ中央銀行、その他のヨーロッパ機関をふくむ。加盟国では、国家あるいは地域レベルで、経済あるいは財務省、他の諸省(部門の比較に関して)、国家統計機関(NSI: National Statistical Institutes)や他の統計機関(規範や訓練他のような生産物に関して)がある。最後に、OECD、国連、IMF 他のような多国籍機関もまた我々の生産物の幾つかの利用者である。

－社会的関係者：ヨーロッパ、各国、および地域レベルでの経営者団体、組合、ロビイスト。

－メディア：数字と分析/コメントの両方に関心を持つ国際的、各国のあるいは地域的な、特定しているかあるいは一般大衆向けのメディア。

－研究者、学生：統計、分析、特別なサービス、個別データへのアクセスを求める。

－企業：(非常な大企業では)自らのマーケット・リサーチ活動のためか、情報部門で何

らかの相談サービスを提供するために。

－Eurostat 内部：多くの高度な統計は他の単位が生産した結果に依存する（比率あるいは集計された勘定のようなより複雑な指数の計算のために）。

利用者と生産物の上述の分類は、当然ながら市場のマトリックス表現にいたる。他の次元が、分析にとって非常に速やかに重要にみえる。すなわち、統計で描かれる分野（経済統計、社会統計・・・）、その頻度、それらが利用者に配布される仕方、・・・である。セルの可能な数は、満足調査で包括的に調査できることを越える。さらにまた、利用者の意見は利用者の特徴に応じて非常に多様であり、そして標本調査の結果は標本の選択に対して非常に敏感であるという一経験に基づいた一感覚がある。

同じく他の困難も生じる。多くの利用者は個人でなくて組織である。生産物の品質に関する意見は、調査された者に応じて非常に大きく違う。特に、それらの組織で異なる責任を負う人々は、品質の異なる側面により大きな重要度を与えるだろう。情報を獲得する責任を持つ者は、特に「アクセス可能性と明確性」および「規則性と速報性」に関心を持つだろうし、社会経済分析家はおそらく「適合性」と「一貫性」により大きな重要度を付与し、意思決定者は「完全性」、「一貫性」および「速報性と規則性」により関心を持つだろう。

最後であっても重要度は小さくないのだが、統計サービスの品質の判断のために利用者の意見を活用することは、4つの制約に直面する。

－第1。統計生産物は無料な－あるいは限界価格よりもより安く売られる－ので、利用者が表明するニーズがどれだけ切実なのかは大部分無視される。ここで考慮される費用は、NSIs による統計の生産のために費用だけでなく、調査票に記入するか面接で回答する回答者の費用でもある！

－第2。利用者の利害のために利用者がその本当の意見を隠すことによって、回答に偏りがある潜在的危険がしばしばある。例えば、行政の部署間、あるいは個別情報の提供者の地位と統計の利用者との間の何らかの利害対立によってである。

－第3。利用者は必ずしも、将来のニーズを考慮しないし、既存のデータ源の可能性を知ってはいない。利用者の判断にこだわると、近視眼的判断をもたらすことがある。

－第4。統計生産物の幾つかの品質を利用者は直接的には評価できない。これは特に、正確性、比較可能性、一貫性に関する情報が統計生産者から与えられざるを得ない統計についていえることである。このことは、真の品質評価よりも品質イメージが支配的になることにつながる。

5.2 品質測定の原理

どんな組織においても、品質を測定する補助的方法は3つある。利用者の満足を通じて、

自己評価を通じて、ベンチマーキングを通じてである。ベンチマーキングは、他からの教訓を既存の法的・社会的背景に導入できる限りで、有用なやりかたである。国際機関は、法的・社会的・政治的根拠に対しては非常に具体的なので、ベンチマーキングが適切な道具とは思えない。自己評価は、非常に詳細な規則をもって行われる限りにおいて、またそれが利用者満足度調査との比較によって、自己満足あるいは自分勝手を定期的に防止するならば、価値があるだろう。

5.3 各統計生産物の品質測定

Eurostat は、NSIs と協力して、統計の品質についての標準報告 (Eurostat,1998a) を企画しつつある。内部的には、統計規範の品質 (Eurostat,1998d) についての報告制度を現在完成しつつある。Arondel and Depoutot(1998)は、統計分析の品質について報告する方法を提案した。統計勘定の品質評価を構成するための幾つかの試みがスタートしようとしている。他の統計生産物に関しては、TQM イニシャティブ (“Qualistat” と呼ばれる) の枠組み内で作業が計画されている。

統計の品質に関する標準報告は、われわれの最初の 6 つの品質構成要素を分割した次の項目をふくむ。

適合性

NSIs と Eurostat はその主な利用者の意見について定期的に調査を行うべきである。

正確性

標本誤差

確率的標本抽出：偏りと分散についての情報が報告されるべきである。

非確率的標本抽出：標本のカバー率の測定、標本の構成に対する感度分析、あるいは選定手続きの事前と事後のモデル化が行われるべきである。

非標本誤差

枠組みの誤差：過大カバレッジ、過小カバレッジ、分類ミス、および統計単位の輪郭描写(delineation)における問題が報告されるべきである。

処理誤差：データ入力誤差、コード化誤差、エディティング誤差の大きさが定期的に評価されるべきである。

無回答誤差：無回答率 (項目無回答と単位無回答) を示す統計、補定に使われた方法の記述および/あるいは、再ウエイトづけが提供されるべきである。無回答の偏りについての報告が Eurostat に対して定期的に送付されるべきである。

測定誤差：このタイプの誤差を探し出すために使われた方法 (例えば、属性間の一貫

性チェック)に関する報告と報告単位に由来する誤差率の対応するレベルが、定期的に Eurostat に報告されるべきである。

モデルの仮定誤差：モデルの使用とそれらモデルの選定過程に由来する誤差が報告されるべきである。これらのモデルは、部分母集団についてデータ収集ができないときあるいは集計された行政データが使われるときに、統計の一部分の推論に使われるかもしれない。

速報性と時間厳守性

統計生産過程における重要な日付が報告されるべきである。

アクセス可能性と明確性

ヨーロッパ連合全体にわたる情報へのアクセスと明確性を評価するために、Eurostat は加盟国の出版物の範囲、標本にふくまれた報告単位に送付された結果、結果の配布計画、そして提供される統計に関連する方法の記録に関する情報に関心を持っている。

比較可能性

場所的比較可能性：NSIs の概念とヨーロッパの定義とに違いがあるところでは、違いの影響の評価が定期的に行われるべきである。同じように、EU 以外のアメリカ合衆国や日本のような主要国との違いも非常に有用である。

時点間の比較可能性：定義の変化あるいは例外的人口事象（非常に影響のある個別統計単位の合併あるいは分割）に由来する非連続の発生の際には、それらの推定値は Eurostat に伝達されるべきである。

一貫性

暫定値と改訂値との：改訂の体系的評価が行われるべきである。

年次結果と短期結果との：類似の属性が異なる頻度で調査されるときには、構造的そして短期のサーベイから推定されたそういった特性の成長率あるいは年次レベルが比較されるべきである。

センサスと年次統計との：センサス結果と前年（あるいは次年）に関する適切な年次サーベイとを、入手できる短期サーベイからの成長率を考慮し調整して比較するべきである。

国民勘定との：概念と測定方法の違いに由来する調整は、第一次のデータ源をふくめて、叙述されるべきである。

他の源泉との：類似の母集団についての調査の異なる分野を目標としたサーベイは、ときどき同じ特性あるいは2つの非常に類似の特性についての情報をもたらす。2つの源泉からの統計を、概念あるいは測定方法におけるいかなる違いをも考慮して、比較するべきである。

5.4 管理の道具としての品質

5.4.1 ある特定のサーベイのために：総合的なサーベイ・デザイン

もし品質と費用についての情報が、統計生産過程のすべて、あるいは殆どすべての局面について入手可能なら、全体的最適モデル (a global optimisation model) を企画することが可能である。これは、フォローアップ、データ・エディティング、補定その他…のより合理的な管理につながる。オランダ統計局は Eurostat の指導の下に、そういったパイロットモデルを開発した (Depoutot and Huighen,1998)。

5.4.2 統計局のために：より合理的な管理

優先度の選択：そういった情報は、十分な品質レベルを達成しないプロジェクトに対して、利用者にとっての重要性を考慮に入れた上で、焦点をおくために、予算配分のための適切な基礎になることは明らかだろう。

修正活動：異なる統計における弱点は、頻度と重要性とによって同じようにランクづけされうるだろうし、統計の品質へのその影響の合理的な順位ごとに取り込まれるだろう。

5.5 ヨーロッパ統計体系における品質の測定

5.5.1 品質、国際的な調整と助成

品質は、ヨーロッパ統計体系(ESS)の建設において決定的問題である。政治的統合はヨーロッパ統計が、EU 合計および比較統計の両方において高い品質を持つ必要度を高めた。統計のよりよい比較可能性のためには、一般に既存のサーベイの調和が要求されると考えられる。そういった再設計の費用は、政府統計部局と回答者の両方にとっては一般的には高すぎて、調整に制約を与える。これは、強い政治的意思によってだけ正当化されうるし、ヨーロッパ連合のような政治的に統合された構造においてだけ想定することができる。しかしそこにおいてすら、問題を解決するために最も適切なレベルが責任を負うべきとする EU のいわゆる「補助的ルール」が、そういった政治的構造の内部においてすら、調整に対しては高い制約があることを強調している (Champsaur,1996)。最近まで、調整はまさに 2 分法、すなわちサーベイが調整されたかそうでないか、において叙述される。完全に調整されたサーベイと調整されないサーベイとの 2 分法の代わりに、品質報告の利用可能性になると、すべての範囲の可能性を開く。ヨーロッパ基準が正確に包括的な仕方 で定義されるなら、各加盟国は、その推定値とヨーロッパ基準から引きだされる推定値との間の (任意の)ズレについて報告できる。このことは、EU 機関の要求と加盟国による統計の調整を高い優先度のある経済的目標に限ろうとする圧力の増大との間のズレに橋をかけることを可能にする。

ESS のもうひとつの独創性は、真の連邦統計制度と比較して、Eurostat と NSIs との間の

役割の再分割を具体化する厳密な補助的規則にある。国家統計機関は一般に組織の全体的自由とサーベイ装置の選択を維持している。このことは、共通の組織的規則と ESS 全体に古典的な TQM アプローチを採用することを妨げる。それにもかかわらず、共通の品質プロジェクトに対するイニシャティブは、最近、NSIs の局長のインフォーマル委員会によって論じられてきた (Franchet,1998 を参照せよ)。

5.5.2 品質要求

品質測定を長期に遂行することは、品質への要求を採用することであろう。これは多くの問題を生む困難な問題である。それは多目的である。というのは、それは EU と加盟国レベルでの品質と国家統計の十分な比較可能性を目標としているからである。これは、多くの制約を持つ問題である。すなわち、異なる加盟国の予算は非常に異なり、移転は不可能である。各国の GNP に統計の費用を比例させることを仮説として考えることは合理的であろうか？ 標本抽出率に対する正確性の周知の非比例性は、明らかにこの原則と大きく矛盾する。非回答者への負担は、あれこれの加盟国について比較可能であるべきか？ われわれは、特にある国の地域が他の国より大きいことを考えながら、地方の要求を統合するべきか？ 既存の連邦の例が非常に違う (アメリカ合衆国、ドイツ) ことを知っているときに、小さな加盟国により大きな制約を与えながら、すべての加盟国に国家的情報の公表を強制的とするのか？ われわれは EU の集計結果からの力をかりて、小さな加盟国の国家統計を生産する小地域推定テクニックを使うことを目論むことができるか？

6. 結論

ヨーロッパ統計体系は、その完全性を改善する最初の局面の後、今や機関の利用者の要請を満たす品質問題と明白に取り組んでいる。政治的統合の進展は、ヨーロッパ中央銀行の最近の発足とともに、品質の測定と一層の改善に向けての大きな刺激である。補助規則はそれにもかかわらず、全システムの品質を特定の角度からのみわれわれが考えることを余儀なくする。特に、ヨーロッパ統計体系が、品質の測定を越えて TQM イニシャティブを発展させる場合には、そうである。というのは、既存の基準 (ISO 9000) と品質のマニュアルのどれも、異なる国籍からの公的行政のネットワークと情報サービスを作り出すことに適合していないからである。

文献

Aronde P. and Depoutot R.(1997), "Quality and Comparability of international statistics", paper presented at the Comparative Analysis of Enterprise Data Seminar, Bergamo, Italy, Dec 1997, Eurostat Working Paper 9/1998/A/4

- Arondel P. and Depoutot, R.(1998), ' Overview of Quality Issues when dealing with socio-economic products in an international environment', XXXth ASU Meeting, Renns,France, 25-29 May 1998
- Beekman,M.M. and Strujis,P.(1993) ' The quality of economic concepts and definitions', Statistical Journal of the United Nations ECE 10, 1-15
- Champsaur P.(1996), ' Harmonisation internationale et comparabilite statistique', Conference on ' Consequences of Globalisation on Statistics Canada's programs'
- Depoutot,R. and Huighen,R.(1998),' A quality model for business surveys management',paper presented at the 1998 IASS Conference, Aguascalientes,Mexico.
- Elvers E. and Rosen B.(1997),' Quality Concept for Official Statistics', entry in the forthcoming update of the Encyclopedia of Statistical Sciences, Wiley & Sons, Statistics Sweden R&D Report, ISSN 0283-8680
- Eurostat(1998a),' Glossary on Quality of Statistics', Document A4/98/Quality/General/Glossary
- Eurostat(1998b),' Definition of Quality in statistics',Document A4/98/General/Definition
- Eurostat(1998c),' Proposal for a quality report in statistics', Document A4/98/Quality/General/Standard report.
- Eurostat(1998d),'Report on the assessment of the quality of statistical norms', Qualistat Internal report
- Franchet Y.,(1998),' Improving the quality of the European Statistical System', discussion paper for the DGINS Seminar, Stockholm, May 1998.
- Statistics Canada(1992),'Policy on Informing Users of Data Quality and Methodology Policy Manual', Statistics Canada, April 7.
- Statistics Sweden(1994),' Quality definition and recommendation for quality declaration of official statistics' (authors: Chris Denell,Eva Elvers and Bengt Rosen)
- United Nations(1983), ' Guideline for quality presentations that are prepared for users of statistics', Conference of European Statisticians, Meeting on Statistical Methodology,21-24 November 1983.
- U.S.Office of Federal Statistical Policy and Standards(1978), Statistical Policy Working Paper 3: ' An error profile: Employment as Measured by the Current Population Survey'.

論文

統計の品質（統計の真実性と関連諸要因）

—最近の国際的論議を参考に—

伊藤陽一¹

1 はじめに

今日、先進国の政府統計機関は、国民の政府統計活動への協力度が低下するなかで統計活動の様々な革新を模索している。この中で、統計の真実性を高め、統計の真実性の程度についての説明を丁寧に公開しようとする動きと論議が、「統計の品質」というタイトルの下に活発になってきている。統計の品質の問題は、これから新たな統計生産に着手しようとしている諸国、既存の統計生産を今日の国際的な標準に見合うよう努力している諸国においても最重要な問題の1つである。

国際統計界では、特に1990年代に統計の品質をめぐる研究・論議が活発になった。

日本の社会統計学では、過去70年以上にわたって「統計の真実性の吟味」の研究が中心問題のひとつであった。ここでの理論的蓄積には、目下の「統計の品質」論に貢献する部分がある。他方で国際的論議は、より多面的に展開しており、日本においても参考にすべき点が多くある。

本報告は、これまでの日本の社会統計学での蓄積をふまえ、統計の品質をめぐる論議を大略的に整理することを課題にする。

2 日本における先行研究と報告者の検討視角

筆者は、これまでの日本の社会統計学での蛭川以降の論議と最近の国際的研究・論議の各々について次のように考える。

(1) 統計生産過程を、客観（の集団的数量的側面）を認識する過程として位置づけ整理した蛭川（蛭川、1932）の見地と幾つかの指摘は、統計生産過程（統計調査）論を客観的に展開する上での、基本的枠組みを与えるものとして継承すべきである。同じ

¹ 法政大学日本統計研究所/経済学部

この論文は、第3回日中経済統計学会議(1999.10.10-11、於：嘉興市)において、配布・発表した。その後、若干の修正を加えた。

く、統計利用者の立場の重視は、今日、国際的に重要視されている点から見て、極めて先見性があったといえる。(2) 蛭川の統計生産論が認識論、方法論に重点がおかれていたのに対して、統計制度を重視すべきとした大屋の指摘も重要である(大屋、1995)。筆者は、政府統計の生産過程は、認識過程と統計行政という2側面を持つ、「行政的」「認識過程」として統一的に把握すべきものとする(伊藤、1980、1988)。この行政過程として把握は、行政改革、統計予算・資源の制約、国民の協力度の低下、統計行政サービスの強化、等の今日的課題を論じるときに不可欠の視角である。(3) 国際的統計論議で1990年代に活発化した「統計の品質」論には、統計利用者の要求を重視する点と、統計生産物を一般の生産物における品質管理の角度からみるという新しい見地がある。これによって、① (i) 統計利用者について、顧客(customer)とみ、顧客の満足を配慮する、(ii) 統計資料を統計生産物とみなして、顧客のニーズにそった生産物を提供する、このために統計生産過程の工程管理を強め、生産物については十分な説明を添えて顧客に提供する、(iii) このような見地から、政府統計活動と統計資料を検討する、等の新しい見地がある。さらに、② 論議の内容は、統計生産過程を中心に、かなり包括的かつ詳細になってきている。但し、③ これらの議論は、蛭川が提起し、筆者も継承する統計生産論の基本的枠組み、すなわち、統計生産過程のステップ論や認識過程と行政過程の統一として把握する見地が不足しているように思われる。

統計の品質をめぐる国際的研究・論議が提起しており、我々もまた重視する論点としては、以下がある。(1) 統計データの品質の構成要素(第一次データ、分析加工結果、国内生産物と国際的生産物)、(2) 統計データの品質を規定する要因、(3) 統計の品質の測定、(4) 統計の品質分析による品質管理。そこで、本稿では、この(1)、(2)を中心に、品質論議を概略的に整理することにする。

あらかじめ図1によって、本稿の課題を位置づけておく。まず統計的生産物がある。統計的生産物には、統計生産(調査)の成果としての一次統計数と、(これを総合加工したGDP等)、統計分析結果がある。これらを一括して最終生産物とよぶことにする。この最終生産物の各々の品質の良し悪しを、品質を構成する要素にそって検討することが、統計利用者にとっては最も重要である。この最終生産物の品質に影響を与えるものとして、統計方法や統計基準等がある。これらを統計の中間生産物ということにする。この中間生産物の品質も問題とされる。最終、中間の統計的生産物に影響を与えるものとして、さらに統計制度がある。この統計制度についても品質が問われる。第1図は、一国の統計についてのものであるが、国際統計についても同じように、国際統計機関をふくめて質が問われる。

図1 統計の品質をめぐる概念図（統計的生産物の品質の構成要素、および統計的生産物に影響を与える要因）

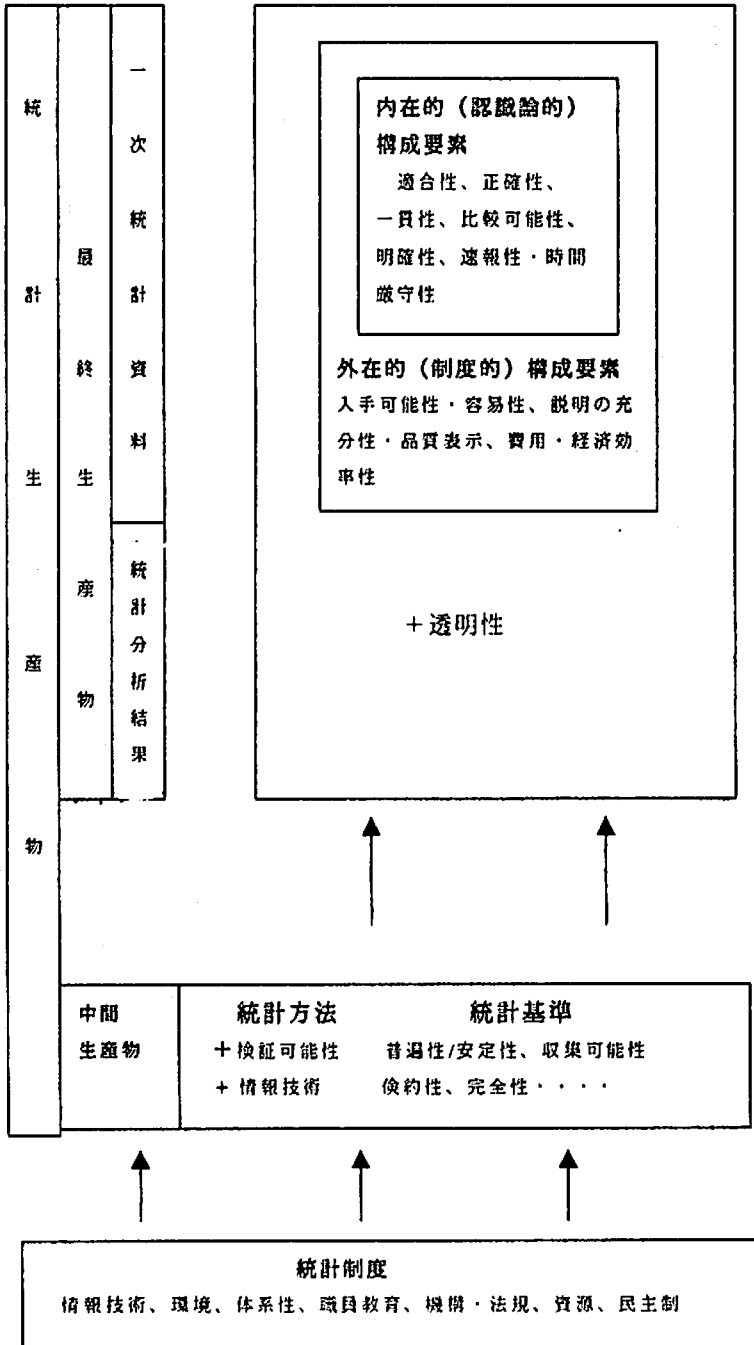


表1 統計生産物の品質の構成要素、及び関連要因とその品質

	品質構成要素	統計生産物（最終）				統計生産物（中間）		統計制度
		国内統計		国際統計		統計方法	統計基準	
		一次統計データ	統計分析結果データ	一次統計データ	統計分析結果データ			
1	内在的（認識論的）要素 適合性(Relevance)：階級・階層分析視角、ジェンダー分析視角の重要性	○	○	○	○	○	○	
2	正確性(Accuracy)	○	○	○	○	○	○	
3	一貫性(Coherence)	○	○	○	○	○	○	
4	比較可能性 (Comparability)	○	○	○	○	○	○	
5	明確性(Clarity)	○	○	○	○	○	○	
6	公平性(Impartiality)					○	○	○
7	速報性と時間厳守性 (Timeliness and Punctuality)	○	○	○	○			○
8	外在的（制度的）要素 入手可能性・容易性 (Availability and Accessibility)	○	○	○	○	○	○	○
9	説明の十分性と品質表示 (Sufficient explanation and Quality Presentation)	○	○	○	○	○	○	○
10	透明性 (Transparency)		○		○	○	○	○
11	費用と経済効率(Cost and Economic Efficiency)	○	○	○	○	○	○	○
12	普遍性/安定性 (University/Stability)						○	
13	検証可能性(Verifiability)					○	○	
14	収集可能性 (Collectability)						○	○
15	儉約性 (Parsimony)						○	○
16	完全性 (Completeness)						○	○
17	情報技術の導入・活用					○	○	○
18	(地球) 環境問題への配慮							○
19	統計生産・貯蔵・配布分析の体系性（重複、空白の排除）						○	○
20	職員教育							○
21	以上をふくむ統計革新を遂行する機構や法規、必要な資源の確保など							○
22	統計生産の企画・実施・蓄積・配布における民主制 (Democracy in the Statistical Services)							○

3 統計データの品質の構成要素(1)：内在的（認識論的）要素

統計データの品質を構成する要素に関しては、これまで多くがかかげられてきた（Depoutot, R.

1998a, b, Eurostat, 1998a-f, Nanopoulos, P. 1997）。すなわち、適合性、正確性、速報性・時間厳守性、比較可能性、一貫性、費用・予算、透明性、入手可能性、入手容易性、明確性、説明の十分性、高潔性などである。筆者は、これまで統計生産・配布過程の民主性という要素を重視してきた。この要素をも付け加えたい。さらに筆者は、これら構成要素は、統計生産物の品質の内在的

（認識論的）構成要素と間接的（制度的）構成要素に分けることができると考える。以下、この区分によって、それぞれを簡単にみていく(表1)。

まず、統計生産物の内在的（認識論的）構成要素とは、統計的生産物が、調査対象（統計対象）を真実性をもって反映しているか、という属性をさしている。これは、適合性、正確性、一貫性、比較可能性、速報性と時間厳守性、からなる。

3.1 適合性(Relevance)

その統計が当該問題の検討や政策形成にとって適切であるか（適合性：Relevance、日本の社会統計学のタームでは信頼性）という要件である。これは目指される統計内容が、利用目的に十分対応している、言い換えれば、適切である、という要請である。この適合性は、統計生産者が、その統計を生産する目的に即して、理論をしっかり持ち、調査対象、内容・項目等を適切に選択・定義・設計すること、すなわち、調査過程において調査票の設計までの段階で確保される。一方、統計利用者側は、自らの利用目的に、入手した統計が対応しているかを吟味して、この適合性を判断する。これは「調査すべき問題を適切に調査しようとしているか」の問題である、という言い方をしてもよい。この適合性の評価は、統計利用者の種類（政府関係、研究者、一般国民など）・関心・統計利用目的によって異なる。統計の価値・有用性は、その経済・社会にとって、基本のもの、第二次のもの、第三次のものとしてある程度のコンセンサスは得られようが、ここでも階級・階層、性別によって評価は違ってくる。例えば、企業にとっては、雇用労働者に関わっての負担額全体が利潤との関連で基本的関心事であって、雇用労働者の可処分所得とこれによって営まれる生活内容は、三次、四次の問題でしかない。国民諸階層の生活の立場から、また、女性の活動と状況を男性と対比してみるジェンダー視角にたって（伊藤、1994, 1997）統計の必要度・有用性を評価するという見地は、これまで、しばしば企業や政府の側には希薄であった。今日では、利用者あるいは顧客重視としてこれら一般国民の要求に注目する動きが強まっている。

3.2 正確性(Accuracy)

調査の目的や調査に利用可能な資源の制約等の下で、調査対象と調査内容について正確な数値を提供しているという要件であり、一般には、誤差の大小の問題である。この要件については、これまで、統計学や統計生産論が詳しく論じてきた。したがって、ここでは主として Eurostat の

説明に基づいて(Eurostat,1998c)簡単に項目を掲げるととどめる。大きくは、統計生産の形態として標本調査によったことから生じる標本調査に固有の標本誤差と、生産(調査)形態の如何にかかわらず総ての調査において問題となる非標本誤差がある。

3.2.1 標本誤差 (sampling error) には、確率的標本抽出法による誤差と、非確率的標本抽出による誤差がある。

3.2.2 非標本誤差(non-sampling error)

①フレーム誤差(Frame error)：対象集団リストの誤りから生じる。過大カバレッジは、実際には範囲外の単位を誤って分類するとか、実際には存在しない単位にをふくめる誤り。過小カバレッジは、対象範囲阿内の単位を、分類の誤り等によって、除外してしまう誤りである。

②測定誤差(Measurement error)

測定誤差はデータ収集時に起こる誤差である。測定誤差の源泉は多くあり、サーベイの装置(書式あるいは調査票)、回答者、情報システム、データ収集の様式(直接的面接、電話インタビュー、自己管理の郵送サーベイ、毎日のサーベイ、行政記録、直接観察、電子的観察)、および面接者等である。

報告誤差は、調査者と回答者の間での伝聞・通信の際のミスである。

報告単位に由来する誤差。報告単位に由来する誤差は事業サーベイよりも世帯サーベイにおいてより大きい。例えば、これは過去の事象を思い起こすことの影響、他人が評価すると思われる特性を過大報告する傾向、回答者の注意の欠如、年齢や教育の影響による。

調査票に由来する誤差。調査対象や項目は正しく認識されながら、調査票の質問、定義や言い回しなど設計が不十分であることからくる。例えば、個人がその年齢を質問されるサーベイでは、切り下げられた年齢データが報告される傾向がある。生年月日を質問するのが望ましいのである。

調査員に由来する誤差。訓練されていない調査員による調査内容の理解不足や回答者への対応の不適切、回答内容の検討不足によって生じる。

③作業誤差(Processing error)。作業誤差は、コード化、キーづけ、エディティング、ウェイトづけ、製表といったデータ収集以後の過程における誤差である。

④無回答誤差(Non-response error)。これには、項目無回答と単位の無回答があり、サーベイでの主要な誤差の源泉である。それらは管理されない仕方による場合には推

定値に偏りをもたらす可能性がある。重要な問題は、回答単位が非回答単位とどう違うかである。

⑤モデルの前提誤差(Model assumption error)。モデル前提誤差は、統計あるいは指数を計算するための、ウェイト付け、一般化された回帰推定子、完全範囲あるいは固定範囲に基づく計算、ベンチマーク化、季節調整、これまでの正確性要因にふくまれない他のモデル等の、方法の選択によって生じる。

3.3 一貫性(Coherence)

統計の一貫性とは、異なる仕方、異なる用途のために生産される異なる統計間で、それらを結合して利用しても、定義、分類、方法等での一貫性あるいは調整可能性を持っていて、食い違いがないという要件である。ここでも Eurostat の説明に依拠する。

3.3.1 暫定値と最終値の間の一貫性は、予備的統計/指数が十分正確で最終結果数と違いを持たないことである。

3.3.2 年次と短期の統計の一貫性

3.3.3 同じ社会・経済的領域での統計の一貫性 統計のあるグループ、おそらくは異なるタイプの（貨幣価値で、大きさで、あるいは不変価格で、価格指数で）は、ひんばんに同じ現象について異なる角度からの表現をもたらす。それらの表現は、一貫していなければならない。

3.3.4 国民勘定と統計の比較 サーベイからの統計を国民勘定との一貫性も問われる。したがって他の情報源とそれら諸サーベイの全体的一貫性の表示を利用者に提供することがありうる。国民勘定統計で使われる方法は、第一次データ源と行われた調整をふくめて、とりあげられる統計について叙述される必要がある。概念におけるズレもまた考慮されるべきである。

3.4 比較可能性 (Comparability)

同一属性についての異なる地域、あるいは異なる時期についての統計系列・統計表の比較可能性である。地域、時期の違いにおいて概念の違いや変更がある場合には、それが明確にされて比較可能な統計に調整されていること、あるいは調整できる情報が与えられているべきという要件である。比較統計表では、比較可能な程度を示す評価が与えられるべきである。

3.5 明確性 (Clarity) 説明や統計表示が、用語、表形式（タイトル、単位、時、クロス形式、実数、比率等の表現等）にわたって、意味が明確であり、簡潔かつ分かり易いことである。

3.6 速報性と時間厳守性(Timeliness and Punctuality)

利用者は一般に、統計的情報について、最小限の時間で生産され、利用可能になると同時に公表され、適切な場合には、定期的に利用可能になるべきという要件である。この要件は、与えられる統計データに内在的な要素なのか、次項(4)の外在的な要素

なのか、迷うところがあるが、一般に生産物でも鮮度が問われ、特に統計の場合には、最重要の要素なので、ここにふくめる。

4 統計の品質の構成要素(2)：外的（制度的）要素

統計生産物の内容が真実性を持っていたとしても、第一に、それが、統計利用者のもとに、利用しやすい形で提供されなければ、利用者・消費者が利用・消費すべき生産物である意味を持たない。したがって、統計生産物の入手可能性、入手容易性、説明の十分性（説明が沢山あるか、説明は明解か）、品質表示という構成要素が必要である。第二に、統計データの生産において、投入される資源（費用、人員）の大きさが問われる。一般に、十分な資源が与えられたときには、品質の高い統計データを生産することが可能であろう。しかし、今日の統計機関は、限られた資源の下で質の高い統計データをを生産することが要請されている。ここで、費用・経済効率（Cost and Economic Efficiency）が要件になる。

4.1 入手可能性・容易性(Availability and Accessibility)。

統計利用者、さらには広く国民が統計に関心を持ち、必要とする統計データを容易に入手できる体制と情報が用意されるべきという要件である。このためには、第一に、各種の統計生産物に何があるか、どこにあるかが明示されていること、第二に、統計資料が公立図書館や学術機関の図書館など公衆の身近に十分備えられていること、最近ではコンピュータ用の電子的媒体によって統計が配布される形、さらにはウェブサイトを（インターネット）を通じての統計の展示・配布が強化されるべきことがある。統計生産物の入手可能な者とそうでない者との不公平が拡がったり、一部の者だけの統計にならないことが必要である。

4.2 説明の十分性と品質表示 (Sufficient Explanation and Quality Presentation)

統計利用者が統計データを入手できても、その内容が理解しやすく、疑問・質問に対応する体制を整えるべきという要件である。ここでは、その統計データの質の構成要素についての説明と質の評価結果が、統計データの生産過程の説明とともに公表されるべきである。この説明は、専門的利用者と一般利用者など異なるレベルの利用者に対して異なる内容で用意されるべきだろう。

4.3 費用と経済効率(Cost and Economic Efficiency)

統計作成が、資金と人員を投下して行なわれ、さらに回答者側に経済的・時間的負担を強いる。その統計の有用性は投入資金・回答負担費用との関連で評価されるべきであり、統計の質の構成要素に入れてよいと考える。これに対して、Eurostat (1998b) は、費用については「質それ自体の構成というよりも質の改善のための制

約である」として構成要素から除外する。筆者は、同一の正確性をもつ統計が、一方ですぐの費用で生産され、他方で多くの費用で生産されている場合、総合判断において、前者の場合に質は高いという評価を与えて良いと考える。統計機関はこのような質の向上をめざしている。統計の質の評価において正確性と費用とは同時的に取り扱うべきとみて、構成要素に入れることにした。もとより、この費用・経済効率性を追求する過程で、統計職員や調査員数、その給与・労働条件、その他統計生産のための関係費用を、みだりに切りつめることは、統計全体の質の低下を招く。この事態は、しばしば起こりがちなので特に注意を要する。

5 統計分析結果、その他の統計生産物の質、そして構成要素間のトレードオフ関係

5.1 統計分析結果の質

統計データには、統計生産（調査）過程の結果である統計表（そこには簡単な構成比と指数が添えられることがある）を一次統計と位置づけ、多くの異種あるいは同種（時系列）の一次統計をもとに数理的手法やモデルを使用して分析した統計分析結果がある。この統計分析結果の質は、一次統計の質に全面的に依拠する。したがって一次統計の質の構成要素は統計分析結果の質の構成要素になる。そのうえで、なお統計分析結果についてさらに固有の構成要素を論じていくことは丁寧といえる。デポト (Depoutot, R. 1998c) は、ここで、公平性 (Impartiality) : 説明モデルの選択が、非常に頑健な (robust) な方法によっており、十分正当化されるべきこと、と透明性 (Transparency) : 分析者の各判断について仮説と分析者の科学的論証との区別が、読者にわかりやすいこと、という二つの要素をあげている。このうち、透明性は、先にふれた説明の十分性要素の一部分になるように思える。公平性は、後にふれる方法の質に関連する。

5.2 その他の統計生産物 Eurostat の論議では、その他に国際機関による統計生産物、記録簿 (レジスター) の品質、勘定体系の品質等、統計の中間生産物、最終生産物、そしてその中での個別生産物等についても、品質を検討できることが示唆されている。中間生産物とは、統計生産と統計分析に関与する方法や基準である。統計の品質論議は、統計生産物とこれに関わる全分野にまで広げられるべきものである。統計方法と統計基準については、6 でふれる。

5.3 質を構成する要素間のトレードオフ

以上に見てきた質を構成する要素の間には、ある要素を高めようとすると、他の要素が低められるというトレードオフの関係が生じうる。Eurostat (1998b) は、速報性/正確性、正確性/場所的比較可能性、適合性/時点間比較可能性、適合性/正確性、大きな領域での一貫性/部分的領域での一貫性の五つをあげている。費用を質の構成要素

とみる筆者の立場からは、費用要素は、全要素と、そして特に、正確性、速報性、入手・利用可能性、説明の十分性とトレードオフ関係にある。
実践的な場では、このバランスをとることが重要になり、またこれら要素を総合的にどう評価するかが焦点の問題になる。

6 統計データ（生産物）の品質を規定する関係分野の品質：統計方法、統計基準、統計制度

上に4で示した統計生産物の質の構成要素に影響する関連統計分野としての統計方法、統計基準、統計制度についても、Eurostatは、その品質を評価しようと提起している。

6.1 統計方法の品質

Eurostatは統計方法の品質については示していない。統計方法は、統計生産物を生み出す過程で採用されるものであり、これまでの統計学が主要課題の一つとして、基礎理論と応用によって絶えず開発と改善をめざしてきたものである。そこでは、方法の科学性を基礎に、それら方法を適用した結果の有効性との関連で、資源等の制約など現実諸条件の下で使用の是非が検討されてきたものであった。これには、統計生産物の品質要素のうち、正確性はその生産物の評価によって結果的に評価されると見ることにすると、適合性、明確性、一貫性、公平性、透明性、検証可能性、費用・経済性、が特に問われるといえるだろう。

6.2 統計基準の品質

Depoutot,R.(1998b)は、適合性、明確性、一貫性の他に、普遍性と安定性、検証可能性、収集可能性と儉約性、完全性という要素をあげている。ここで普遍性/安定性(University/Stability)とは、基準が、目標母集団を通じて、そして十分長い期間を通じて、同じ仕方で理解されること、とされる。検証可能性(Verifiability)とは、検証を要求する総ての者が、基準を十分な一致性(conformity)を以て適用することの保証である。これは、提供された基準の正確性、基準に関する総ての文書へのアクセス可能性、実施における許容限界についての明確な定義、個別情報の提供者が与える基本的個別情報の正確性をチェックする可能性をふくむ、という。さらに収集可能性(Collectability)とは、こういった基準を実施するときに、個別情報を実際に収集できる法的および文化的可能性、儉約性(Parsimony)とは、基準の選択で予算に与える影響、また完全性(Completeness)とは、統計基準の集合について、それらが利用者の主要な必要を満たすかどうかについての要素である、とされる。儉約は、既にふれた費用要素として論じることができる。

6.3 統計制度の品質

Eurostatは、統計制度の品質内容に関する論議を示してはいない。統計制度に関しては、何よりも中心的に、品質の高い統計生産物を提供するための体制が問われる。そして、特に入手可能性、入手容易性、説明の十分性、品質表示、経済効率が注目される。統計制度に関しては、「高潔性あるいは自立性 (Integrity)」という要素が語られることがある (Bonnen, J.W. and others, 1980)。これは、統計家が専門性と倫理性を持ち、客観的な統計を提供すること、制度全体の公開を通じて政治的圧力に対して自立的であること、これらによって統計制度が大衆からの信頼をうること、を意味している。筆者は、この高潔性の中身の大部分は、この報告でとりあげている品質の諸構成要素に分解できると考える。政治的圧力からの独立という要素も、統計の正確性や以下にふれる民主制にふくまれると考え、独自の要素とはみなさないことにする。筆者は、今日の国際的な政府統計機関における統計改革の動き、そして統計の品質問題への関心の高まりと利用者を重視する見地の強調、といった点を考えると、統計制度の品質の構成要素として、①情報技術の導入・活用、②(地球)環境問題への配慮、③統計生産・分析・貯蔵・配布の体系性(重複、空白の排除)、④職員教育、⑤以上をふくむ統計革新を推進する機構や法規、⑥必要な資源の確保などとともに、特に重要な点として、⑦統計生産の企画・実施・蓄積・配布における民主制 (Democracy in the statistical services) をあげたい。

ここで民主制だけにふれると、統計データの生産が企画・実施・貯蔵・配布される全過程において、国民本位であるという要件である。この要件を満たすことで、政府統計は、国民が必要とするものであり、国民の協力の下に生産され、質の高さが確保されることになる。この中味としては以下がある。

第一に、何を調査するべきかの企画の段階で、国民の意向を汲み上げる体制としては、特定分野、さらには特定調査について、労働組合など国民各階層の要求・意見を汲み上げ、検討する恒常的審議機関の存在、臨時の公聴会(地方に出張しての)等の開催、さらには、要求・意見を受け付ける窓口が存在すること、である。

第二に、統計生産の実際過程が国民の納得の行く形で進められべきである。まず、情報社会における民主制保障の二大要件、すなわち、情報公開とプライバシー保護の基本点が徹底されることが重要である。この場合、情報公開とは、国の活動が、国民に周知され、国民による監視が行き届いていること、プライバシー保護とは、個人情報秘匿性 (Confidentiality) が保障されるべきこと、個人情報の扱いがその個人の了解の下に行なわれることを、基本の精神とする。さらに言えば、これらは、国が持つ情報の扱いにおいての主権在民=民主主義をより徹底すること、直接民主主義に近付けること、といっても良い。この情報公開の実質化によって、統計資料とともに国の統計活動の内容は国民の知り得るところとなり、理解も深まるし、統計要求も出やすくなる。プライバシー保護が実質化されると、統計調査活動は、目的と必

要性について国民への説得が必要になり、調査への協力は強制か任意か、強制の場合、調査拒否をした場合どのような罰則があるかを事前に告知すること、調査個票の秘匿状況がどうなるかの説明も行なうことになり、実際状況への国民によるチェックをも一定程度許すものとなる。政府統計活動がこのような説明によりながら国民の協力をとりつけるべきことは、本来的なことなのであって、法的強制は、せいぜいが虚偽申告を引き出すことになる。

とはいえ、民主制が必要なのは、結局、非民主制が統計の真実性を損なうからなのではない。統計の真実性の如何にかかわらず、場合によっては統計の真実性を低下させても、統計の在り方—統計調査の在り方は民主的でなければならない。

第三に、同じく調査実施にあたっては、調査の意義と必要性についての説明をとめないながら、調査指導員—調査員を通じて、被調査者にあたり、調査の現場での疑問・質問に十分に対応することである。日本の場合、この調査の現場を支えるのは、地方統計職員であり、その下につく調査員であるから、彼ら自身が調査の意義、必要についてまず納得がいくよう、調査企画・設計の段階から、彼らの意見・要求を吸収するシステムは必須である。そしてそもそも住民に密着した形での地方統計活動の活発化が現場を担う地方統計職員の士気を高め、活動の前進への寄与を生みださう。統計調査員については、手当てと労働条件の保証が基本におかれる。

第四に、統計生産物の貯蔵・配布において、容易に統計を理解し、利用できる体制が、用意されていること。これは先に、入手・利用可能性と説明の十分性としてふれた。さらに、速報性と時間厳守性もかかわる。

7 統計の品質の測定、結果の表示、品質管理（統計の品質向上に向けての管理体制）

以上、統計生産物の品質を構成する要素と、統計生産物に影響を与える分野を中心に論じてきた。これらの品質論議は、統計生産物の品質を評価し、その向上に努めて、統計利用者（顧客）に高品質の生産物を提供することを目的としている。したがって、統計の品質論議は、さらに、個々の生産物にいたるまで、品質構成要素ごとに、さらにそれらを総合して、評価を与えること（品質測定）と、統計生産工程において、品質管理を行うこと（品質管理）、利用者に対して、品質を明示すること（品質評価表、品質報告の公表）に及ばなければならない。

7.1 統計の品質の測定 デポト(Depoutot,R.1998c)は、品質測定の補助的方法として、利用者（異なった立場にあり、関心も違う）の満足度を通じて、自己評価を通じて、ベンチマーキングを通じての3つを指摘している。品質測定の主体について見ると、さらに、第三者機関、あるいは利用者と生産者、第三者をふくむ評価機関、等を

あげることができる。評価の具体的な手法としては、品質要素ごとに、①素点を入れる、②A/B/C/D等の幅をもったグレード、③点数化する、方法がありえ、さらに、全体として総合評価を与える、を考え得るだろう。

7.2 統計の品質管理 統計生産過程の品質管理は、統計生産物の企画から生産物の貯蔵・配布の全過程にわたってステップと、関連する品質要素とのマトリックスを作成し、このマトリックスに関与する各レベルの統計生産者を指定することが前提になる。その上で、品質要素ごとの評価を与え、さらに品質向上のために詳細な指針を作成し、実施結果を評価する、という往復の作業になる。この過程で、品質要素あるいは品質向上の進展度の要点を必要に応じて公表することが重要である。もちろん、全体としての品質向上のためには、特に統計制度の品質でふれた、職員教育、資源配備ほかの要素が必要である。

7.3 品質測定結果の公表 これまで見てきた統計生産物あるいは関連統計分野の品質測定結果は、

利用者、そして関係者・一般国民に対してどのような形で公表されるか。統計生産物にそえて、あるいは評価一覧として、さらに主要な統計生産物については、「品質報告書」の形で、印刷物あるいはインターネットへを通じて、公表されることになる。Eurostatは標準の『品質報告書』を提起している(Eurostat, 1998c)。1998年のメキシコでのIAOS会議で示された例を見ると、フレームと回収標本(母集団の当初リストと変化、標本数と非回収標本の理由別内訳、回答率)について詳細を示す(オーストラリア, Neiger, D.N. 1998)、フレーム誤差と無回答理由別標本数・率、拒否率を調査と地域別にインターネットで示す(イタリア, Brancato, G. 1998)、無回答標本数と率、標本誤差の他に品質指数を要点について示す(フィリピン, Virola, R. 1998)等がある。さらに調査の概要と背景等に関する説明とともにフレーム、標本回収等についての詳細情報を数年ごとに『品質報告書』を出版する例(アメリカ合衆国センサス局, SIPPのQuality Profile, Jabine R.B, 1990, U.S. Bureau of the Census, 1998)もある。しかし、現時点では、国際的動向においても、全体としては、統計の品質論でとりあげられている詳細な点をふくめての品質測定とその公表迄にはいたっていない。

8 むすび : 今後に向けて

統計の品質をめぐる論議が活発になったのは、1990年代に入ってからであり、論議は進行中である。品質の構成要素とそれに影響を与える統計分野、そして、これら論議を、より広く統計の品質評価として実際の統計生産の場に持ち込み、品質管理に役立て、その結果を公表するという点では、ようやく実践化の過程に入りはじめた

段階といえるだろう。統計生産物以外の分野をも、広く統計の品質とみなすこれらの考え方は、従来、科学性あるいは民主制といった見地からとりあげてきた論点を、「品質」というタームによって単に包括したにすぎない、とも受け取れる。しかし、この品質論議は、統計利用者を取りわけ重視しており、包括的であり、実践的である。筆者は、これを支持して、筆者の考えを織り込み、全体的な整理を試みた。今後に向けての課題は、第一に、本報告で示した、統計の品質論議の枠組み、構成要素、相互関連等については、なお重複や用語など充分には整理されていないので、さらに検討を深めること、第二に、生産過程、統計関係者と、品質構成要素のマトリックスを一覧的に整理すること (Nanopoulos, P. 1997 が若干の参考になる)、第三に、品質測定・評価の方法を実際の統計生産物にそくして具体化すること、第四に、品質関係の情報の公表を政府統計機関に求めつつ、品質公表の形を具体的に検討すること、等を一層追求していくことである。

謝辞 IAOS(1998年)とISI(1999年)においてこのテーマをめぐって報告されたEurostatのDepoutot, Raoul氏にはEurostatの文献の入手と翻訳許可等でお世話になった。感謝したい。

文献 (今後の研究の参考のため、1999年8月時点で、筆者が未見のものもふくめている)

- 蛭川虎三(1932)『統計利用の基本問題』岩波書店
大屋祐雪(1995)『統計情報論』九州大学出版会、に所収
内海庫一郎(1975)『社会統計学の基本問題－現代統計学批判』北海道大学図書刊行会
伊藤陽一(1980)『統計学』法政大学通信教育部
伊藤陽一(1988)「覚え書き：政府統計の理解・批判の視角について」
法政大学日本統計研究所 Working Paper No. A-2
伊藤陽一編著(1994)『女性と統計－ジェンダー統計論序説』梓出版社
伊藤陽一(1997)「日本におけるジェンダー統計－現状、問題、克服の方向－」
『研究紀要』創刊号(国立婦人教育会館)
趙彦雲(1998)「中国政府統計の質について」『統計学』第75号
Arondel P. and Depoutot R.(1997), "Quality and Comparability of international Statistics", paper presented at the Comparative Analysis of Enterprise Data Seminar, Bergamo, Italy, Dec 1997, Eurostat Working Paper 9/1998/A/4
Arondel P. and Depoutot, R.(1998), 'On overview of Quality Issues when dealing with socio-economic products in an international environment', XXXth ASU Meeting, Rennes, France, 25-29 May 1998

- Bonnen, J. T., Duncan, J. W., Goldstein, H. and Hagen, R. L. (1980), "Policy Relevance and the Integrity of Statistics", *Statistical Reporter*, Jan. 1980. 法政大学日本統計研究所訳(1981)『統計の自立性と政策適合性』統計研究参考資料 No.6 に収録。
- Brancato, G., D'Angionili, G., and Signore (1998), "Building up the Quality Profile of ISTAT Surveys", Materials distributed at 1998 IASS/IAOS Joint Conference, Mexico
- Beekman, M. and Struijs, P. (1993), "The Quality of Economic Concepts and Definitions". *Statistical Journal of the United Nations ECE* 10, 1-15. IOS Press.
- Depoutot, R. (1998a) Quality definition and evaluation. paper presented at the Fourth Seminar on the Future of European Social Statistics / Harmonization on Social Statistics and Quality, 26-27 March, 1998..
- Depoutot, R. (1998b), "Quality of Statistical Norms", presented at the conference New Tools and Technology for Statistics 98, Sorrento, Italy, 4-6/11/98
- Depoutot, R. (1998c), "Quality of Statistical Output in an International Statistical Organization" paper presented at the 1998 IASS/ IAOS Conference Mexico
- Depoutot, R. and Huighen, R. (1998), 'A quality model for business surveys management', paper presented at the 1998 IASS/IAOS Conference, Aguascalientes, Mexico
- Elvers E. and Rosen B. (1997), 'Quality Concept for Official Statistics', the Encyclopedia of Statistical Sciences, Wiley & Sons, Statistics Sweden R&D Rpt.
- Eurostat (1998a), 'Glossary on Quality of Statistics', Doc. A4/98/Quality/General/Glossary
- Eurostat (1998b), 'Definition of Quality in Statistics', Document A4/98/General/Definition
- Eurostat (1998c), 'Standard Quality Report' in statistics', Document A4/Quality/98/General/Standard Report/Proposal.
- Eurostat (1998d), 'Report on the assessment of the quality of statistical norms', Qualistat Internal report
- Eurostat (1998e), 'Assessment of Comparability', doc Eurostat/A4/98/Quality/General/Comparability rev.1
- Eurostat (1998f), 'Assessment of Coherence', doc Eurostat/A4/98/Quality/General/Comparability rev.1
- Eurostat (1998g), Sub-group of the Quality Assurance Group, Quality of Norms. Final Report
- Franchet Y. (1998), 'Improving the quality of the European Statistical System', Discussion paper for the DGINS Seminar, Stockholm, May 1998.
- Grais B., (1998), 'Statistical Harmonisation and Quality: The case of social Statistics', Paper presented at the 4th Mondorf Seminar on 'Harmonisation on social Statistics and quality'
- Jabine, T. B. (1990), SIPP Quality Profile, Bureau of the Census, U.S. Dept of Commerce.
- Nanopoulos, P. (1993), "Measuring the Quality of Official Statistics", Seminar on "Quality Control of Statistics", Athens, Greece, January 27-29

- Nanopoulos,P.(1997),”Quality of European Business Statistics”, in Z.Kenessey ed.
Economic Statistics-Accuracy, Timeliness, and Relevance.U.S.Dept of Commerce, BEA
- Neiger,D(1998),” Survey Quality-Measuring it and Communicating it to the Users”, material
 Distributed at 1998 IASS/IAOS Joint Conference, Aguascalientes, Mexico.
- Picard,Hugens(1993) ”Measuring the Quality of Official Statistics”,Seminar on
 “Quality Control of Statistics”, Athens,Greece, January 27-29
- Statistics Canada (1992),’Policy on Informing Users of Data Quality and Methodology
 Policy Manual’, Statistics Canada, April 7.
- Statistics Sweden (1994),’ Quality definition and recommendation for quality declaration
 of official statistics’ (authors: Chris Denell,Eva Elvers and Bengt Rosen)
- United Nations (1983), ’ Guideline for quality presentations that are prepared for users
 of statistics’, Conference of European Statisticians, Meeting on Statistical
 Methodology,21-24 November 1983.
- U.S.Office of Federal Statistical Policy and Standards (1978), ’ An error profile: Employment
 as Measured by the Current Population Survey’. Statistical Policy Working Paper 3
- U.S.Bureau of the Census (1998), *SIPP Quality Profile* 3rd edition, Draft.
- Virola,R.A,(1998)” Quality of Official Statistics in the Philippines”, material distributed at
 1998 IASS/IAOS Joint Conference, Aguascalientes, Mexico.

訳者（執筆者）あとがき

1. 本資料は、統計の品質をめぐる Eurostat（ヨーロッパ連合統計局）関係の6つの重要文書、すなわち、

1. Eurostat (1998) *Standard Quality Report*, (Eurostat/A4/Quality/98/General/Standard Report/Proposal),

2. Eurostat (1998) *Definition of quality in statistics*,
(Eurostat/A4/Quality/98/General/Definition)

3 Eurostat (1998) *Glossary on Quality of Statistics*,
(Eurostat/A4/Quality/98/General/Glossary)

4 Raoul Depoutot (1998) "Quality of Statistical Norms" presented at the conference
New Tools and Technology for Statistics 98, Sorrento, Italy, 4-6/11/98

5. Eurostat Sub-Group (1998) *Quality of Norms: Final Reprt*

6. Raoul Depoutot (1998) "Quality of Statistical Output in an International Statistical
Organization" presented at the 1998 IASS/IAOS Conference, Aguascalientes, Mexico
の翻訳と、これらの動向をふまえた伊藤の論文（1999年9月）を収録している。

4,6の著者 Raoul Depoutot 氏はこのテーマを Eurostat で担当されている。

2. 伊藤論文でふれているが、ここにいう「統計の品質」論とは、統計資料の品質をめぐる、特に1990年代に国際的に集中的にとりあげられている論議である。

統計データの品質と言えば、さしあたり、誤差（標本誤差と非標本誤差）についての論議と受け取る向きがあるかも知れない。しかし、現在の国際的論議は、論議の枠をはるかに拡大し、実に多面的に進められている。これを一瞥的に概観するためには、本資料の伊藤論文に示した図と表（82-83ページ）を参照するのが便宜的かと思う。

3. これら文書を訳出しようと思いついた経過を、収録した伊藤論文の背景説明や若干のポイントを説明する中で示すことにしたい。

伊藤の論文は、Eurostat を中心とする論議の弱さを指摘し、これら論議をふまえて問題全体を整理し、この先の展望を示すべく、1999年9月にまとめて、10月10-11日に中国の嘉興市で開催された第3回日中経済統計学会議で発表・配布したものである。その日本語版は、他に公表しておらず、翻訳した Eurostat の論議を位置づけ、評価する上でも役に立つと考え、本資料に収録し、読者の参考にするとともに、ご批判を仰ごうと考えたのである。

詳しく立ち入ることは避けるが、統計の品質論の一部領域は、かつて蜷川虎三氏が「統計の理解・吟味・批判」として提起して以来、日本の社会統計学でも、論議されてきた点と重なる。政府統計機関では実は、日々の統計活動で突きつけられている問題として、また日本統計学会等でも「標本設計・標本誤差」論、そして「非標本誤差」論として取り上げられてきた。その後、大屋祐雪氏が、統計学に統計制度論的視

角を取り入れるべきことを提起された。この主張は、統計研究と政府統計活動の実際との乖離・遊離を強く戒める指摘として重要であった。

伊藤は、大屋氏からの刺激を受けて、1980年代に、いわゆる方法論的視角と統計制度論的角度の両方が客観的に存在するとみた。そして、特に、国民一般（広い非専門的統計利用者であり、被調査者であり、ごく一部が自ら統計生産に乗りだし、情報公開・プライバシー保護問題に関わり、統計生産・配布をふくめて多くの統計要求を持つ）の側から、政府統計活動を、方法と制度の両側面をふくめて検討するというスタンスで研究を進めてきた【『統計学』法政大学通信教育部（1980）、「統計の理解・吟味・批判」日本統計研究所ワーキングペーパー（1989）】。1980年代の合衆国での、今に言う「統計の品質」論をも念頭におきながら、上記1989年のワーキング・ペーパーでは、多くの未整理部分を持ちつつ、制度論をふくむ新たな論点を取り入れて、統計の真実性あるいは統計の理解・吟味・批判論議にひとまずの整理を与えた積もりであった。その後、1990年前後から、北欧や合衆国で、政府統計活動の中に一段と利用者（顧客）重視が進んでいることに共感を覚えたのであるが、この問題には立ち入らないでいた。

4. そして、1998年のメキシコ・アグアスカリエントスのIAOSの、「統計の品質」セッションにおいて、今日的論議の状況を把握することができた。このセッションを傍聴することが、実はメキシコ迄でかける強い動機の1つであった。

ここでは、政府統計機関が、自ら生産した統計データ・加工データの品質を評価し、ランク付け・評点を与えて、国民に対して広く公表するところまで論議が進み、一部実施されていること、さらにEurostatが、統計の品質をめぐる、非常に包括的な論議を進め、順次報告書にまとめていることが示された。この会議の直後にEurostatに立ち寄り、（そこでは会えなかったが）会議でEurostatの動向を示されたRaul Depoutot氏から、関連報告書を送っていただいた。本資料で訳出したのは、1998年段階で入手したものである。

これらの資料を以てしても、国際的な論議と統計活動は、日本でこれまで行われてきた品質問題論議の枠を大きく越えて進んでいることがわかっていく。これらの動向について日本では関心が向けられているのかが不明であるし、なお多くの論議は、誤差論あるいは真実性論というかつてのフレーム内の論議にとどまっているのではないかと思う。統計活動の実際においても、ここ2年ほどの間に、日本の統計機関におけるホームページの充実によって、統計データと関連情報は一般国民に身近になったが、統計の品質情報をわかりやすく直接公開する迄には至っていない。本資料を急いだ所以である。

5. 中国での発表にあたっては、時間的制約から、この論文の中身をそのまま、語るわけにはいかなかったので、前置きと収録された説明図を中心にポイントのみを示した。その導入は、お客が、例えば、食料品の購入のために店に来た場合、あるいはレストランにでかけた場合を考えてみればよい、というところからはじめた。食料

品の場合には、価格との相対関係の下に、食料品そのものが良質であることが基本であるが、その組成や添加物、産地、賞味期限、が明示されていること、さらに店舗の商品配列からはじまる店づくり、従業員のサービス等々が、買い物客をして購入を決心させることになる。その場合、良質な食料品を生み出す産地で品質管理にはじまり、流通過程、そして店舗での販売における管理も関わる。気持ちの良いレストランでは、特に、雰囲気や従業員の対応も関わる。統計の品質論議のひとつのポイントは、統計生産物をこのような客向けの生産物とみなし、いわばハードとソフト、をふくめて、客へのサービス (customer service) を徹底するという見地から、考えてみよう、ということにある。統計生産物の有用性は、社会認識の材料として、対象反映性を持つことがカナメとなる、すなわち品質の核の部分になるが、決してそれだけではない多様な側面から、この生産物を最終的に消費するまでを考えることができるのである。

国際的な品質論議は、論文でもふれたが、統計機関が自分たちの統計生産物について品質に関して客観的評価を与え、これを消費者である統計利用者に公表するという方向で動いている。伊藤の論文は、この点を当然のこととして歓迎しながら、国際的論議においては品質構成要素が、並列的・羅列的に示されていると見て、大きな整理枠を与えようとした。またさらに、Eurostatとは異なって、費用も品質の構成要素に組み込まれるべきとみなし、統計生産者 (統計家) の労働条件等も考慮すべきなど、国際論議ではとりあげられていない点も指摘した。

とはいえ、品質構成要素の分類の大枠は試論的なものであるし、品質構成要素の相互関係などについては、概括的論議にとどまっている。細部の検討・展開の中で、不足な点もでてくるかと思う。

6. 翻訳に関しては、特にテクニカル・タームについて一定の検討はしたが、訳語が適切かについては自信があるわけではない。時間をかけて吟味・選択することもありえたのだが、この問題に関しては、目下の動向を早くに伝えることが重要だと考え、公表を急いだ。早くに出版する積もりが、他に時間を取られて素訳を棚上げする状態がひたすら続いてしまったからである。訳語の不始末についてはご批判・ご指摘を仰ぎ、改めて完成版に持ち込みたい。

7. Raul Depoutot氏は、1999年のフィンランドでのISIでも品質問題の報告をされており、そこで直接話を交わす機会を得た。ここでの資料によって、Eurostatの活動のかなりを把握出来るが、すべてではない。引き続きフォローする必要がある。もちろんEurostat以外の諸国での、品質問題の論議・まとめ、そして統計活動への導入の状況も、しっかりフォローする必要がある。

最後に、Eurostatと氏自身の関連報告・論文を送付していただき、それらの翻訳について許可され、また本資料での翻訳にあたっての幾つかの問い合わせに、丁寧にお答えいただいたRaul Depoutot氏に深く感謝したい。

伊藤陽一 1999年12月

統計研究参考資料（最近発行分）

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

統計研究参考資料 No.61

「統計の品質」をめぐる
—翻訳と論文—

1999年12月20日

発行所 法政大学日本統計研究所

194-02東京都町田市相原町4342

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

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

発行人 伊藤 陽一