

オケージョナル・ペーパー No.25

位置情報を用いた調査票情報の情報価値 の拡張とその分析的意義について

2011年6月

法政大学

日本統計研究所

位置情報を用いた調査票情報の情報価値の拡張 とその分析的意義について

森 博美(法政大学・経済学部)

要 旨

分析材料としての統計資料に対する分析手段としての統計解析技術の発展は、現実についての統計的認識の進歩に有効に貢献しうる。しかし、いかに革新的な解析技術であっても、分析資料が潜在的に保有する情報も含め、それらが最大限に語ることのできる範囲を超えたいかなる知見も提供することはできない。その意味で実証分析の質と可能な到達範囲は、あくまでも分析に用いられる統計資料がデータ形態として本来的に保有する情報量によって自ずと制約される。

これまで政府統計は、新たな方式の統計調査の導入によって収集された調査個票情報を原情報として、次々と新たな形態のデータセットを分析資料として提供してきた。そしてこのことは、新たなデータ形態に適合した解析手法の開発と相俟って、旧来のデータ形態の分析資料によっては決して到達しえなかった現実に対する認識の前進に貢献してきた。

調査客体であると同時に分析の対象でもある個人や家計、企業や事業所といった社会を構成する要素は、いうまでもなく社会的存在である。これらは、その社会的性格の故に、当然のことながら時と場所の要因によって規定される。本稿では、これらのうち特に個体の場所的規定に注目することで、今日まで政府統計によって提供されてきた様々な形態のデータに共通に内在する問題点を明らかにする。そしてさらに、調査票情報からのいわば演繹的拡張として、場所的規定を内在させた新たなタイプのデータの可能性とその分析的意義についても考察する。

キーワード 調査票情報、位置情報、個体レコード、立地、GPS

はじめに

わが国においても、かつての実証分析は、集計結果表に基づくマクロ分析が中心であった。当時も、一応制度的には、個票情報の利活用への道が指定統計調査票情報のいわゆる目的外申請として開かれていた。そして、申請が許可され使用が承認された研究者等によって、調査個票情報を用いた非集計データからなる分析資料に基づく分析(マイクロ分析)が試みられ、疫学や経済その他の分野で、集計データからなる分析資料からは得ることのできない様々な知見が提供されてきた。

わが国では、統計調査に係る秘密保護重視の観点から、これまで非集計データである調査個票情報の提供に関する制度運用については慎重な取り扱いが行われてきた。特に、利用目的についても高度な公益性が要求されてきた。その結果、調査個票情報へのアクセスの壁は、分析者にとってこれまで極めて厚くしかも高いものであった。調査個票情報についてのこのような制度運用により、調査個票情報の利用は行政並びに一部の研究者に限られ、わが国では、長きにわたってマイクロ分析は、学術研究における支配的な研究様式とはなりえなかった。このような分析資料への

アクセス可能性において存在した明らかなデータデバイドは、研究面でのわが国の実証分析の立ち遅れを招いた。

近年、わが国においても、遅ればせながら、マイクロ分析への研究者の関心が高まっている。大容量のデータ処理を可能にした計算機の高性能化、非集計データの解析手法の整備、そして何よりも統計法改正による匿名標本データ(マイクロデータ)の提供システムの制度化がその背景にはある。なお、分散型統計制度を持つわが国では、この種のデータの提供は、原則として各統計を所管する府省によってそれぞれのチャンネルで行われており、しかも提供形態も、匿名データの提供、データのオンサイト利用、さらにはデータのオンデマンド処理と多様である。

ここで、データのタイプ論という観点から、分析資料のデータ形態論的特徴とそれが可能とする分析との間の関係を、簡単な例示によって整理しておく。

データの基本的形式としては、分析資料はまず集計データと非集計データとに大別できる。しかし、データに基づく分析の方向やその広がりやその規定するのは、集計、非集計という基本的なデータ形式の違いだけではない。なぜなら、仮に非集計データだけを例にとっても、例えば世帯内における世帯構成員の世代間あるいは世帯員間の関連を分析するためには、位階的(hierarchical)なデータ構造を持つ世帯データから世帯員コードを用いて構成員を相互にリンクさせ、それらを統合したデータセットを編成することが必要となるからである。また、個人や企業に対する政策の有効性を DID 分析(difference in difference analysis)を用いて評価するためには、個体ベースでの縦断面(longitudinal)データが不可欠である。

分析資料としてのデータのタイプ論的側面が、それに基づいて実行される分析の基本的方向ならびに内容を制約する。本稿では、このようなデータ論的視角から、特に調査客体の存在に関する位置情報に焦点をあて、位置情報を媒介にした個体レコードへの空間的特性を反映した変数の付加によるその情報価値の拡張可能性について論じるとともに、そのような拡張された個体レコードによって編成される新たなタイプのデータセットが持つ分析的意義について考察する。

本稿の構成は、大要、以下の通りである。第 1 節では、統計調査から得られる調査客体に関する調査票情報に関して、調査客体そのものの場所的存在にかかわる外性諸要素が存在する事実の確認を行う。第 2 節では、ドイツ社会統計学の代表的論者であるマイヤーの所説を手掛りに、統計における位置情報の意義について考察する。第 3 節では、社会的存在としての個体(個人、世帯、企業等)に対する場所的規定要因とそれらの類型化を試みる。第 4 節では、調査票情報の個体性に注目しつつ、位置情報の個体レコードへの付加の可能性を検討する。そしてさいごに、個体レコードへの場所に規定された諸変数の取り込みによる場所的要素対応型データセット構築の実証分析に対する意義について筆者なりの見解を提示することでむすびとしたい。

1. 実証分析のデータ制約

若者、女性、それに高齢者の就業行動、世帯類型別の家計の消費行動、さらには企業の投資行動や生産性等に関するこれまでの多くの実証分析には、共通するひとつの特徴がある。それは、回帰分析の場合、専ら個体の属性を示す諸変数が独立変数として使用されている点である。冒頭において指摘した点との関連で言えば、このような分析の方向は、基本的に分析者が分析資料として使用できるデータそのものの制約によるところが大きい。

もちろん、個人、家計、それに企業といった個体は、無機的な世界の中に存在しているわけではない。かつての社会統計学は明示的に意識していたことではあるが、それらはまさに社会的存在として様々な集団を形成しつつ社会的営みをしている。それらが社会的存在であるという事実は、その存在そのものが時間的ならびに場所的(時空間)制約の中に置かれていること、別な表現をすれば、時間的ならびに場所的規定は、それらを分析対象とする際の前提となる枠組的要素を構成していることを意味する。

社会的集団の存在に関わる要素のひとつである時間的規定については、データの形態としては時系列データがそれらの集計量としての時間的変化を捉える分析資料として存在する。さらに、時間軸方向での集団の個体ベースでの変化の追跡についても、1960年代以降、時間的存在にかかわる要素をデータ形式そのものの中に内在させた縦断面データ(パネルデータ)が構築されることで、分析資料が整備されることになる。

他方、場所的規定は、これまで主として地域別集計データに基づく地域比較分析との関係で取り上げられてきた。地域比較分析においては、多くの場合、わが国では都道府県あるいは市区町村といった行政区分によって表章された集計表がその分析資料として用いられてきた。ちなみに、工業統計(経済産業省所管)工業地区編は、工業地区という特に選定された地域別の結果表を掲げている〔経済産業省(2010)〕。現在、全国で233箇所の工業地区が選定^{〔注1〕}されているが、これらはいずれも単一ないしは隣接する複数の市区町村にまたがる地区として境域設定が行われているものである。工業統計における工業地区別集計は、基本的に市区町村という行政区分に基づく名寄せ再集計といった性格を持つ。その意味では、場所的規定のあり方としては、境域設定の点で既存の行政区分と必ずしも次元を異にするものとはいえない。

後述する地域メッシュ統計や町丁・字、基本単位区別結果表のように、行政区分とは異なる地域表章による集計結果表もその後提供されるようになった。しかしながら、わが国の場合、調査結果の地域表章は、これまで基本的に行政区分あるいはそれを基に編成される地域区分に従って行われてきた。場所的規定の取り扱いに関するこのような分析資料の提供形態を巡る事情が、公表された既存の分析資料に基づいて行われる分析の方向と内容を規定してきた。

ところで、例えば流通業における商圈、農業における土地利用、通勤・通学圏などからも明らかなように、場所的規定の中には、必ずしもその境域が行政区分によって形作られるそれと一致しない場合が少なくない。このような場合には、行政区分別の結果表は、必ずしも実態を反映しない統計的表章を与えることになる。このような現実から乖離した集計量からなる分析資料に基づく分析は、結果的には偏りを持つ解析結果をもたらすことになる。

それでは、場所的規定に関してこれまで統計学ではどのようにそれを取り扱ってきたのであろうか。

2. マイヤーにおける位置情報と統計地理法

(1) 統計の比較可能性と場所的規定

統計における比較可能性は、統計のまさに生命線である。その意味で比較可能性は、分析資料としての統計の価値を左右する最も重要な要件のひとつといえる。比較可能性は2つの方向を持つ。横断面では、同一時点における異種の情報源から得られるデータ間の比較可能性、また縦断

面では、同一項目に関する調査結果の時系列的比較可能性がそれである。統計の比較可能性は、統計で用いられる各種の定義や概念さらには分類の標準化によって担保される。

本稿は、統計における場所情報の取り扱いをその主たる検討テーマとしている。これまで統計では、通例、地域という面的広がりによって、統計情報の空間的な結果表章は行われてきた。統計調査が経常的な行政行為として反復実施されるようになるに従って、調査結果の地域表章については、比較可能要件を充足する地域区分として一般に行政区画が採用されることになる。ジー・ジェックが「場所的総括は、通例、政治行政区域・・・による」[Žižek (1922) 8 頁]というとき、そこには、すでに経常的行政活動として展開されている統計作成行為を背景にして、比較可能な調査結果の提供という官庁統計の社会的役割を強く意識した指導的統計家としての彼の姿が投影されている。本稿は、ジー・ジェックが単に「通例」として片付けている統計の地域表章における場所的規定の取り扱いに対する筆者の素朴な疑問から出発している。

ところで、社会統計学の初期の理論家の中には、統計における場所的規定に関して、ジー・ジェックのような通例とは見解を異にする者がいないわけではない。

例えば、蜷川虎三は、彼の主著である『統計利用に於ける基本問題』の中で、「其の存在が社会的に規定せられた集団を・・・大量」(p.68)と定義し、大量が社会的存在であることを根拠に、「ゆゑに一定の存在の時と存在の場所を有つことは云うまでもない」(p.99)と指摘している。さらに彼は、存在の場所に関して、「場所は必ずしも行政的区画を意味するものでなく、要するに単位の一団が大量として限定せらるゝ空間的存在条件たればよい。従って生産米を市場により、或は又、交通路により、平野によつて大量として限定することも、かく規定せざるを得ざる必然性を有つならば、その方が寧ろ正しいのである」(p.160)といった立論を展開している。

このように蜷川は、統計における場所的規定に関して、ジー・ジェックとは異なり、都道府県や市区町村といった行政区画によって結果表章を行うことが必ずしも適切ではなく、統計的把握の対象となる調査事項によっては、それぞれに対応した独自の地域表章がありうることを示唆している。

しかしながら、蜷川の主たる理論的関心は、大量概念の導入による独自の集団論に基づく統計理論の体系化、すなわち、主としてドイツ社会統計学派によって展開されてきた悉皆大量観察を中心とした統計調査論と英米を中心に発展してきた統計解析法を内容とする数理統計学との統合に向けられていた。蜷川は、独自の集団論、すなわち大量という社会的に存在する集団と統計解析のために意識的に構成された集団との区別から出発し、集団論という基礎の上に演繹的な理論構築を行っている。そのため彼の知的関心は、理論構築の出発概念としての二種類の集団の特質の解明に主として向けられることになる。

集団論の見地に立つ蜷川にとって、大量の時間規定は一定の実質的意味を持つ。なぜならそれは、静大量と動大量という統計的把握において質的位相を全く異にする二種類の集団として、統計調査において区別された取り扱いを必要とし、調査論における重要な検討事項のひとつとなりうるからである。一方、場所的規定に関しては、集団の存在の場所が異なればそれらの社会的性格も当然異なるとの認識からか、蜷川においてはこの点について集団論にまで踏み込んだ特段の理論展開は見られない。

時間的規定と異なり、統計における場所的規定については、ジー・ジェックや蜷川に象徴されるように、これまでの統計学における関心はおしなべて希薄であった。

(2)マイヤーにおける場所的規定

ところで、ドイツ社会統計学の創始者のひとりとして知られている統計家にマイヤー (Georg von Mayr: 1841-1925) がいる。彼は、1866 年からバイエルン統計局長補佐、また 1869 年から 1879 年まで同局長の地位にあり、統計行政官としてバイエルン王国の政府統計行政に従事した。マイヤーの統計学が政府統計実務を背景に持つのは、彼のこのような社会的立場によるものである。

マイヤーは統計における場所的規定に関して、当時の他の諸家に見られない興味深い所説を展開している。そこで本節では、マイヤーの場所的規定に関する主張を今日的視点から捉え直すことによって、彼の所説を手掛りに統計における位置情報の意義について考察してみたい。

(i) 統計作成者の立場からの場所的規定

マイヤーも蜷川と同様に、その著書のひとつ *Statistik und Gesellschaftslehre* (大橋隆憲訳 (1943)『統計学の本質と方法』)の中で、「社会集団は、…国家および自治体区域とは無関係に、多種多様な分布状態で地表を蔽ふてゐる」[Mayr (1914) S.62、訳書 p.153]として、行政区分と集団の存在の空間的広がりや境域とが必ずしも一致しない可能性を指摘している。マイヤーは、統計調査による把握の対象となる社会集団がこのような空間的存在特性を持つという認識に立ち、統計における場所的規定に関して次のような議論を展開する。

マイヤーの場所的規定に関する立論は、統計作成にかかわる一般的規定と結果の分析に関わる具体的規定の二段階から構成されている。そのうち一般的規定にあたる部分は、統計調査の企画者、実施者、あるいは統計実務全体の統括者、すなわち指導的統計家としてのマイヤーの立場から展開されているものである。

マイヤーは、調査が実施される空間的範囲としての「一般的な場所の意義」について、「調査機関が活動の手を差し上げ、観察すべき集団の要素をその範囲内で捉へねばならぬ調査区域の規定と限定にとっても重要である」[Mayr (1914) SS.84-5、訳書 p.205]とする。調査の実施が一般に政府による多大な財政的支出と人員の投入を必要とすることから、「統計調査は行政の権力手段…を用いずには実施し得ない」。その結果、「観察分野全体を限定することにとっては公法的構成体の事実上の場所区画が重要な意義」[Mayr (1914) S.62、訳書 p.153]を持つことになる。

行政情報あるいは個々の行政単位によって予め要約された資料から編成される二次統計の場合には、それぞれの資料を記録している行政区画が、統計的把握における場所となる。一方、大規模調査の場合には、「調査機関の観察分野の場所限定によって、観察すべき場所区域が特別に構成されねばならない(調査区)」[Mayr (1914) S.85、訳書 p.206]。

このような、調査の実施範囲、さらには実査によって調査客体を漏れなくしかも重複なく把握するために排反的に領域設定される調査区という統計の場所的規定は、あくまでも調査実施者の立場から捉えられたものである。従って、このような「観察分野全体の確定」は、場所的規定としては「単に社会集団についての最初の最も一般的な場所的限定をしたに過ぎない」[Mayr (1914) S.62、訳書 p.154]のものである。

(ii) 統計分析者の立場からの場所的規定

バイエルン統計局の指導的統計家であったマイヤーは、同局に在職当時、ミュンヘン大学統計学部の私講師(後に員外教授)も兼務していた。彼は、全体で 85 の節から構成される *Statistik und Gesellschaftslehre* (第 1 版)の第 57 節から 63 節を、調査結果の科学的利用、特に社会生活における合法則性の検出に充てている。ここには、指導的統計家であると同時に社会学者としての一面を併せ持つマイヤーの独特な立場が投影されている。

このようなマイヤーの立場の二重性は、彼の場所的規定を巡る議論にも反映されている。すなわちマイヤーが「社会集団につき具体的な場所的限定を考へねばならぬ」[Mayr(1914)S.62、訳書 p.153]、なぜなら、それを限定することが調査結果の総括や説明に重要な意味を持つからである、と主張するとき、そこには、統計作成の責任者としての立場を超えたひとりの社会学者としてのマイヤーの場所的規定の一端をわれわれは読み取ることができる。

マイヤーは、すでに紹介したように、社会集団の空間的ありようについて、調査さらには調査結果に基づく分析と関連づけて、「観察目的よりするも、またその後の科学的研究目的よりするも、観察分野内の集団の要素を完璧に更に場所的に区分する必要が極めて大となるのである。蓋し事実上、社会集団は多種多様な仕方で観察地域に場所的に分布してゐるからである」[Mayr(1914)S.62、訳書 p.154]と指摘している。にもかかわらず、実際に作成されている統計の中にはこのような社会集団の実態に配慮を欠くものが少なくないとして彼は次のように批判する。すなわち、「大なる県、大なる州や国、かかるものについての平均事情を与へただけの統計値図表は、ほとんど価値がない。否、そればかりか悪影響すらあるのである。けだし、それは広範なる地域にわたり諸現象の斉一性が全然ありもしないのに、あたかもその斉一性があるかのごとく見せかけるからである。」[Mayr(1914)S.183、訳書 p.433]、と。

それでは、マイヤーが考える社会集団の実態を反映したしかるべき調査結果の地域表章とは、一体どのようなものであろうか。大規模な地域区分を持つ結果表であっても、「縮めた数値をその社会集団の状態や現象に就ての地理的統計的細目を認識せしめうるがごとく注意が払われてゐる場合」には、その限りで「その現実姿態を完全に体现してゐるものとして真の意義を持ちうる」[Mayr(1914)S.142、訳書 p.337]。すなわち、地域区分が社会集団の空間的ありように照応したものである場合には、仮に大規模な地域区分による集計結果表(場所的総括)であっても実質的意味を持つのである。

社会集団の実態を反映した有効な地域区分の基準としてマイヤーは、地理学的地域区分という概念を導入する。すなわち、「総括表を作成する際に地理学的な地域や地帯の実際の境界を精密に遵守する程、総括表の作成はより完全に行はれる。」[Mayr(1914)S.143、訳書 p.339]、と。この地理学的地域区分という調査結果の地域表章の基準を彼が提起するその理論的背景には、地理学あるいは統計地理学と呼ぶ彼の独自の分析手法がある。そこで、以下にそれらの内容を概観しておく。

(a)マイヤーの場所的規定と地理学、統計地理学

マイヤーは、調査結果の地域表章を行う際には、安易に行政区画によるのではなく、「地理学上異なる(異質のものとして区別される—引用者)地域乃至地帯を標準として観察結果をグループ分け」[Mayr(1914)S.143、訳書 p.338]することを提案する。なお、地域区分にあたってマイヤーは、「風土学上の、水勢学上の、地球構造学上の、その他等々の、地域構成が問題とされる」[Mayr(1914)S.143、訳書 p.338]という地理学的視点に立って、ライン平野やシュワルツワルトといった既存の政治的・行政的地域区分とは異なる地域表章の可能性を挙げている。

彼がこのような政治的・行政的地域区分と異なる地域表章を提案することの背景には、地理学の知見に基づく地域の設定とそれを分析資料とする解析という二つの作業工程からなる「統計地理学」という彼の独特の社会科学分析方法論がある。

統計地理学の前半の工程は、調査結果の地域集計に基づいて、このような科学的考察法が適用できる分析資料を準備することにある。そのために何よりも必要なのは、「一般行政区画によつて

定められた大地域の概括的結果を利用する代りに、一つの統計的に観察された具体的事実現象について、独自の地理的分布状態なくその等級別の総括分布状態〔Mayr(1914)S.144、訳書 p.340〕を提示することにある。その際に、彼によって「部分地域限定の基礎」として使用されるのが、「地理的要素」〔Mayr(1914)S.143、訳書 p.338〕である。

地理学からの知見を手掛りにして、ある共通の特性を持つひとつの境域が編成される。その境域は必ずしも行政的区分とは一致しない。また、設定される境域は、関心事項が異なれば、自ずと違ったものとなりうる。

このような、調査結果に基づいて設定される境域を彼は「統計州」と呼ぶ。なお、この点についてマイヤーは、「斉一の(また斉一と定められた)統計的現象形態を持つ小さな基礎的区画から形成された新しい複合地域は、普通の意味で固定した地理的地区ではなく、具体的な統計の対象が異なるに応じて変化する地理的分布に関する統計的地区である。それは、・・・当該大地域についての、云わば新たに生じた事実上の地理的に総合された分布を明示するものであり、おそらくは「統計州」とでもいふに相応しいものであらう」〔Mayr(1914)S.144、訳書 p.341〕と指摘している。

調査結果に基づいてこのような統計州が適切に設定できた場合、すなわち、「統計地理法による結果が、事実として明確な「統計州」の存在を示す場合には、所与の事実現象の等級別を表はす特殊の統計地区を基礎として、他の仕方によった統計的観察材料をグループ別けする場合には、この方法は統計的吟味法として顕著な効果を挙げることができる」〔Mayr(1914)S.144、訳書 p.341〕ことになる。

マイヤーは、統計地理法の前半の作業工程の具体的な進め方について、「ある小地区について総括を行ひ、それに基いて比率計算を行はふと思ふならば、この手続を行った際に、隣接する多数の場所・・・に同一または近似の等級へ属させ得るとき斉一的な配置が生ずるか否かを吟味する必要がある」〔Mayr(1914)S.144、訳書 p.340〕として、具体的に児童死亡率に関する境域設定〔Mayr(1914)S.145、訳書 p.342〕を行うとともに、同様の手法は人口密度、出生数、私生児や死産といった人口動態現象の頻度分析にも適用できるとしている。そして、「実際に斉一な配置が発見」されることで、統計地理法は有効な吟味法たることが立証される」〔Mayr(1914)S.144、訳書 p.340〕ことになるのである。

社会生活における規則性(合法則性)の検出への科学的接近を試みる際に、利用可能な集計結果における表章形態の制約から「事実観察だけでは因果性推断のための問題とすべき質的区別が見つからぬ場合」〔Mayr(1914)S.64、訳書 p.157〕には、変数同士を直接的に関係づけて分析することはできない。このような場合の代替的方法として彼が提案するのが、地域比較による統計地理法である。

彼は、統計地理法の適用例示として、宗教の宗派別の自殺率に関する結果表が存在しない場合の分析を取り上げ、次のような方法を紹介している。すなわち、「諸宗派と自殺者との関係は示されておないが、諸宗派と人口一般との関係が示されてある場合、各種地域、すなわち、同一宗派だけの諸地域あるひは種々の割合で宗派の混合してある諸地域につき、無差別の一般的な自殺頻度を宗派別に研究するならば、自殺頻度と宗派との間に存する推測上の関係についての近似的な観念を得ることができる」〔Mayr(1914)S.64、訳書 p.157〕、と。このような地域的システムによる間接計算に基づく比較は、マイヤーにとって「因果性発見の重要な手がかり」〔Mayr(1914)S.145、訳書 p.342〕を与えるものである。

(b)統計地理法の観点から見た地域結果表章のあり方

社会集団の地域における分布状態は多様である。その分布が地域的に集中している場合には、集団の所在地の個々の特性を明らかにすることが重要であり、結果表章の形態としては、「観察結果は場所または土地の特性によって別々に示されねばならない」〔Mayr (1914) S.63、訳書 p.155〕。

これに対して、社会集団の構成単位である個体が場所的に広範に分布している場合マイヤーは、「もし理想的な完全さを期さんと欲するならば、・・・統計調査に際し、できる限り個々の場所的個別性を重視すべき」〔Mayr (1914) S.63、訳書 p.155〕としながらも、統計技術上の制約という見地から、その処理は自ずと現実的とならざるをえないとして、「一切の捉へきれない場所的個別性を残らず区別しておくことは、第一次調査にとっても・・・困難至極であるが、さらにその調査結果の整理の段となつては全く不可能」〔Mayr (1914) S.63、訳書 p.155〕であると考えている。

調査における場所的規定の取り扱いにおいて、個々の場所的個別性の把握といった「理想的な完全さ」の追求が統計技術的に不可能な場合には、多かれ少なかれ一定の空間的広がりを持つものとして社会集団の構成要素の場所的存在を把握する必要がある。その際に、集計作業上の便宜からすれば、区域の規模は広域であるほど望ましい。しかし、編成される統計州の境界が仮に既存の行政区分と整合的でない場合には、行政区による地域表章結果から社会集団の現実の存在を反映した適切な統計州を編成することはできない。

そのような場合の次善の方策としてマイヤーは、「統計州の構成に当っては、一般に現実問題として、計算労力の過大となることを避け」（訳書 p.341）、しかも「具体的な社会関係を明確に認識するに最も役立つが如く何時なりと結果の地域別を編み直しうる様に予めしておくこと」〔Mayr (1914) S.63、訳書 p.156〕が大事であるとして、「社会集団を適当な下級小区画へ場所的に総括」〔Mayr (1914) S.63、訳書 p.155〕することを提案する。なぜなら、実務作業的に現実的な小区画を最も基礎的な地域単位としておけば、「何時なりとも・・・観察客体を前とは異なったより大なる範囲の場所的総括へ類集することができる」〔Mayr (1914) S.63、訳書 p.156〕からである。

積み上げにより統計州を編成する際の基礎的地域単位としての小区画は、調査結果の「第一次整理」の単位を形成する。なお、地域単位と関連してマイヤーは、今日のメッシュ統計にあたる物理的に構成される区画設定についてもすでに言及している。彼は、この種の地域設定について、「純幾何学的に規定された斉一的な面積単位（たとへば1平方キロメートル毎といったが如き）はあるが、これを社会集団の場所区画の第一標準となすことは不適當である」〔Mayr (1914) S.63、訳書 p.155〕としてそれを排除している。

このようにマイヤーは、既存の州のような大規模行政区画も、そしてまたメッシュのような行政区画とは異質の物理的区分も地域単位としては不適切であるとして退ける。そして彼は、「観察の・・・第一次整理は所与の最下級小行政区画（特に市町村の小地区）と結びつけられねばならぬであろう」〔Mayr (1914) S.63、訳書 p.155〕として、最終的に「小行政地区（県、地方裁判所管区等）による区画」〔Mayr (1914) S.144、訳書 p.340〕をありうべき地域単位として提案する。

マイヤーが最終的に県や地方裁判所管区といった小行政地区を調査結果の第一次整理の地域的単位として採用したのには、いくつか理由が考えられる。その1は、統計調査の実施体制との関連である。すなわち、事後的に編成される統計州の境域が行政区画のそれと必ずしも整合性が持てない場合があるとはいえ、彼自身、統計調査の実施も含めた行政権の行使という点で、行政区画の持つ社会的意味を無視し得なかつたものと推察される。

第 2 は、統計州編成における自由度という視点である。このことは、「社会集団をかくの如く小地域別に限定しておくならば、ひとは随時その欲するところにしたがひ現存統計資料を観察地域の地事情とむすびつけることができ……これは等しく科学的に意義のあるところである」〔Mayr (1914) S.64、訳書 p.156〕という彼の指摘からも分かる。その意味では、「できるだけ小さな場所的区画を基準として総括を行はねばならない」〔Mayr (1914) S.144、訳書 p.340〕と同時に「小さすぎない場所区画」である必要がある。彼の場所区画設定に関する第 3 の論点は、整理結果の安定性への配慮である。彼は、最終的な小行政地区の提案に先立つ文章で、「この区画の大きさは、信用における比率計算を行ひうるだけの充分大きな基本数値を保持しておきたいといふ願望のために、制約されざるを得ない」と指摘し、その点からも「小行政地区(県、地方裁判所管区等)による区画が地理的要素として利用」〔Mayr (1914) S.144、訳書 p.340〕できるとしている。

(3) 小括 マイヤーの所説の今日的意味

大量あるいは社会集団の場所的規定に関して、それらの空間的広がりが必要しも行政区画によって編成される境域とは整合的でないと認識では、蜷川とマイヤーは共通の立場に立っている。しかしながら、蜷川が場所的規定そのものにそれ以上の関心を払っていないのとは対照的に、マイヤーには、統計地理法という空間と関連付けられた集計量(マクロ統計)をベースにした分析方法論との関連で、場所的規定に関して今日的観点からみても示唆的内容を持ついくつかの論点指摘が見られる。

ところで、蜷川の代表的著書『統計利用に於ける基本問題』は、それを今日的用法で表現するならば、政府統計の二次利用に際して踏まえるべき、分析資料と解析手法の吟味批判に関する手引書であるといえる。一方、マイヤーの *Statistik und Gesellschaftslehre*(第 1 版)は、統計の調査から集計過程、さらには結果の利用による社会生活における合法則性の検出に至る一連の過程を論じたものである。マイヤー自身が指導的統計家として統計作成に直接的に責任をもつ立場にあったことから、彼には立場としての政府統計の二次的利用者という明示的な意識はない。

しかし、マイヤーが場所的規定を巡って展開している所説の中には、既存の結果表からは直接は読み取ることのできない事実に対して、統計地理法という集計量(マクロ)ベースでの分析手法を導入し地域別に表章された既存結果を比較分析することによって、そこに潜む合法則性あるいは因果関係に対して、いわば疫学的に接近を試みるという立論が見られる。これは、見方によれば、場所的規定に着目した既存結果の二次的利用とみなすことができよう。

マイヤーはまた、統計地理法と関連して、数量的に表現された地域の型にも言及しており、「地域区画の場所的個別性は区分された統計的総体の担い手(引用者)として、それが含む質的に異った統計要素の混合物の程度によりこの数量関係の型を現はすものである」〔Mayr (1914) S.64、訳書 p.157〕としている。このマイヤーの指摘に関して特筆すべきは、あくまでも地域区画という一定の空間的広がりを持ってではあるが、地域区画が統計的総体の担い手として位置づけられている点である。このことを今日のデータ論的な観点から敷衍すれば、地域区画がプラットフォームとして、集計量としてその地域を特徴づける統計をオーバーレイしているという統計の捉え方を示唆しているとの解釈も成り立つように思われる。

すでに指摘したように、マイヤーは今日のメッシュ統計にあたる地域表章の形態そのものについては否定的である。とはいえ彼は、今日のメッシュ統計に相当する地域表章だけでなく、「場合によっては幾何学的見地(たとへば同心地帯設定)にしたがった総括の方が必要であり、あるひは合

目的々である」[Mayr(1914)S.63、訳書 p.156]として、今日 GIS において有力な手法として広く用いられているバッファリング的領域設定の可能性にも言及するなど、彼の場所的規定に関する議論の中には、いくつか時代先取りの指摘も散見される。

その一方でマイヤーは、統計をあくまでも集計量として捉える統計観に立脚している。統計地理法が依拠する地域単位として彼があくまでも小行政地区にこだわり続けたのは、それぞれの地域単位が安定的な統計的総括を持つために必要な調査客体の確保が不可欠であったからである。その点で彼にとっては、あくまでも場所的規定は「小さすぎない」ことに意味があったのである。マイヤーは、調査における場所情報の取り扱いにおいて、個々の場所的個別性の把握といった「理想的な完全さ」を意識し、また個々の関心項目に関して様々なポリゴンの形を取りうる統計州さらには今日のバッファリング的な発想を持つ一方、地域単位としてはあくまでも一定の空間的広がりを持つものとして場所的規定を捉えている。

3. 社会的存在としての個体の場所的規定

(1) 個体の社会的性格とその場所的規定

統計調査において調査客体となる個人や世帯、それに企業者や事業所は、決して無機的な世界における孤立した存在ではない。マイヤーや蜷川の指摘を俟つまでもなく、それらは社会的存在である。それらは、個体としてあるいは集団の一構成要素として固有の属性を持ち、時空間の中で他の個体や集団と様々な関係を取り結ぶ中で自らの意識を形成、変容させ、また活動を行う存在である。これらの関係の総体が、社会現象として統計が究極的にその反映を行わなければならない対象に他ならない。

本稿の冒頭でもすでに簡単に触れたように、これまで政府統計は、その歴史的発展の過程で、様々な形態の統計を分析資料として提供してきた。それらは、それぞれのデータ形態に適合した解析手法の開発ともあいまって、現実に対する統計的認識の進歩に貢献してきた。しかしながら、統計データの形態的展開をデータ論という視角から捉えた場合、これまで種々の形態を持って展開されてきた統計資料には、その分析面での有効性に関して、ひとつ決定的な難点が内在しているように思われる。

マイヤーや蜷川は、集団の社会的側面に注目することで、時間的規定と場所的規定を、統計がその把握対象とする集団の存在を制約する基本的要素として位置づけている。しかし前節でも述べたように、本稿が検討の対象としている場所的規定については、マイヤーにおいて統計地理法と関連しての若干の議論は見られるものの、蜷川の場合はこの点に関して特に立ち入った指摘はない。専ら分析資料の解析手法の学問として発展してきた数理統計学の場合はいうまでもなく、統計学を現実に対する認識方法論と位置づける論者達の間でも、場所的規定については、「通例」に倣い、多くの場合、行政区画による地域表章の問題として捉えられてきた。その後、情報技術の発展を受けて、わが国でもセンサスを中心にメッシュ統計や様々な小地域集計が提供されるようになる。しかしこれらについても、その意味するところは、基本的に集計量(マクロ)ベースでの地域表章という形での表章形態の外延的拡張に他ならない。

統計ならびに統計学がこれまで場所的規定をこのように取り扱ってきたこと背景には、どのような事情が考えられるであろうか。前節で紹介したようなマイヤーの先見の問題提起にもかかわらず、統計における場所的規定を巡る問題は、その後の統計学の理論展開の中では後景へと退く。そ

の結果、社会集団に関して、統計が把握すべきその存在に関わる社会的要素は、単に社会を構成する主体が織り成す関係、あるいはせいぜい時間的経過の中でのそれとして捉えられてきたように思われる。そこでは、地勢等の自然的条件だけでなく、環境汚染等の負の外部環境も含めた事後的に人為的に加えられた社会インフラの整備等の改変、さらには個体の有機的結合から醸成される社会的条件という外部環境から制約を受ける存在という個体の社会的側面が完全に脱落している。

(2) 調査客体に対する場所的規定

(i) 場所的規定と個人、家計

経済活動の主体としての個人や家計は、具体的な存在の場を持つ。ここで、わが国のマイクロ実証分析の有力な推進分野のひとつとして知られる労働経済における就業行動の分析を例に取り上げてみよう。分析で用いられるモデルの形式にはいろいろな創意工夫は認められるものの、それらは基本的に、就業、非就業という二値からなる従属変数を、性や年齢、学歴、その他の個人属性、親子関係といった世帯内の関連属性、さらには職業訓練等の経験の有無といった個体が獲得した事後的属性に対応する一連の独立変数によって説明するという点では共通している。

ここで、この種のモデル分析について、素朴な疑問がある。それは、仮に共通の個体属性を持つ個人は、果たしてその者の就業行動に対して単なる確率的な差異しか持ち得ないのかという疑問である。なぜなら、この種のモデル構築にあたっては、各個体の存在にかかわる具体的な位置特性は全く考慮されていないからである。この種のモデルは、実質的には行政区コード情報以外には何らの場所的規定も有しない空間的にはあたかも無機的世界の中に存在する個体観を前提に構築されている。言い換えればこの種のモデルでは、例えばその者が享受する利便性その他に他の個体と本質的相違があったとしても、それらの要因は彼らの就業行動に対しては独立であるとの前提に立っている。

家計の消費行動についても同様である。すなわち、商業施設の隣接地域に居住する世帯と遠隔地に立地する世帯とではおそらく家計消費行動のパターンが異なり、また世帯が日常的に利用する商業施設間の競合関係の有無等の立地特性の差異もまた価格等を介して家計の消費に影響をおよぼしうると考えられる。

これまでの個人や家計の就労や消費といった経済行動分析は、行動主体が具体的に置かれた場所的規定を完全に考察の埒外に置くという点で、共通にこの種の問題を内在させている。

(ii) 場所的規定と企業、事業所

企業についても同様に、様々なマクロあるいはマイクロ解析手法を用いた実証分析が広範に展開されている。

ここで、企業の生産性の分析を例に取り上げてみよう。物的生産性にせよ付加価値生産性にせよ、生産性は基本的に投入される労働や資本に対する生産物の価額によって評価される。また生産指数と投入指数との差によって定義される全要素生産性(TFP)も、長期的には生産組織も含めた技術水準の進歩、短期的には従業員の技能向上などとして説明されている。

現行の生産性分析に共通する特徴として、個々の企業の立地環境特性の相違が全く考慮されていない点がある。仮に、同一の業種に属し、同等の量と質の資本、労働の投入によって生産活動を営む企業であっても、それぞれが置かれた外部環境としての立地条件が異なることで、両者の生産性には自ずと違いが生じるものと考えられる。ここで広い意味で立地条件に関わる要素^[注2]と

しては、競合状態の有無、交通アクセス(港湾、空港、高速道路のICからの距離)、周囲の人口特性(販路としての市場人口、確保可能な従業員の質と量における難易度の違い)、関連業種の集積状況、さらには制度的要因としての経済特区としての指定の有無や誘致自治体による種々の助成措置の有無、といったものが考えられる。

立地環境特性としてのこのような外部要因については、ある業種での主要な要因が他の業種では副次的要因となる場合も当然考えられる。しかしいずれにせよ、こういった外部要因は、企業の場所的規定に関係付けられた立地環境特性として、企業活動に少なからず影響を及ぼしていると考えられる。なお、従来の生産性分析は、少なくともモデル論としては、このような立地環境特性の作用についても、それを技術進歩などとともに、事実上、全要素生産性として一括して取り扱ってきた。

(iii) 価格と場所的規定

個人や世帯、それに企業や事業所といった通常、調査客体として取り上げられる個体の他にも、外部環境要因によって有意に影響を受けるものがある。取引される財やサービスの価格がそれである。公示地価や路線価、実勢価格といった地価は、都市計画法における地域地区や傾斜方位といったその土地が置かれた制度的、地勢的条件だけでなく、公共交通機関・公共施設・商業施設等へのアクセスの利便性等によっても影響を受けるものと考えられる。また、土地以外の財やサービスの価格についても、それを扱う店舗等の業態類型や店舗規模といった供給側の個体属性に起因する差異だけでなく、所在都市の人口規模、さらには競合店の有無をはじめとする立地環境特性が価格決定に及ぼす影響も無視できない。

4. 場所的規定の個体レコードへの付加可能性

これまで経済分野で行われてきた実証分析は、モデル構築に際して、場所的規定に関連した外部要因を基本的に考慮の対象外としてきた。しかし、分析者の名誉のためにいえば、これは分析者がそのことの意義を決して理解していなかったのではなく、むしろ、そのような分析ができるデータセットがこれまで整備されてこなかったことによると考えるべきであろう。そこで本節では、場所的規定の個体レコードへの付加による新たなデータ形式の可能性について検討してみたい。

(1) 個票調査の歴史的意義と位置情報からみた情報喪失

表式調査を主要な調査形態とする黎明期の統計調査と異なり、近代統計調査は個票による統計単位情報の収集を調査形態上の特徴とする。表式調査に対する個票調査の調査技術面での優位性は、統計作成の出発点となる統計単位情報の形態にある。表式調査の場合、調査事項はもともと集計量として把握される。このように最初から集計された計数を原情報として行政単位のレベルに従って市区町村→市郡→都道府県→全国へと積み上げることで統計が作成される。これに対して、個票調査では、調査対象となる個々の調査客体に関する個体情報が統計単位情報として収集される。

統計作成の出発点となる原情報のこのような形態の違いが持つ意味は決定的である。なぜなら、個票調査の場合、原情報としてそれに続く統計作成過程の出発点となるのは、統計的把握の最小単位であり現実存在としての調査客体と1対1で対応づけられた各変数(調査事項)に対応する

まさに統計単位情報だからである。統計作成の出発点となる原情報がデータの形式としては個体ベースのレコードとして与えられることから、理論的には調査票情報が持つ変数間の任意の組み合わせからなる集計結果表の作成が可能となる。近年の情報処理技術の飛躍的發展によって初めて可能となった多様かつ詳細な集計結果表の提供は、ある意味では個票調査の導入以来もともと調査票情報に内在していた情報の潜在的な利用可能性を単に現実のものにしたに過ぎない。

ところで、これまでの個票調査では、場所情報は、都道府県、市区町村、あるいは調査区によって与えられる一定の空間的広がりを持つ境域として取り扱われてきた。その後、作成、提供される集計結果表が本質的に拡充され、特に国勢調査や企業・事業所統計といったセンサスについては、従来からの地域表章に加えて、基本単位区や町丁・字といった様々な小地域集計が提供されるようになる。行政区、調査区、基本単位区、町丁・字、さらにはメッシュ統計等によって区分される小地域に対応した固有のコードが付与され、それらは調査票情報の欄外記載事項あるいは事後的に付与される場所的コード情報として、調査結果として得られる諸変数値と関係づけられてきた。言い換えれば、これらの地域情報は、それぞれの地域コードを data carrier として持ち、data body を構成する諸変数が相互にリンク可能なリレーショナルなデータ構造を持っている。(統計個票情報における data carrier と data body については[森(2009)]参照)

こういった様々な地域表章に共通する注目すべき点として、該当区域内の調査対象単位の数の多寡こそあれ、複数の単位の位置情報が一個の地域コードによって統合的に表示されていることがある。このことは、データ構造としては、ひとつの地域コードに対して複数の個体に関わる data body 情報が同時にリンクされるという1対nの対応関係が成立していることを意味する。要するにここでは、地域コードという data carrier が、data body として複数の個体から構成される集計値さらにはそれから導出された統計指標群を多次的にレイヤーの形で背負っている。

いうまでもなく調査区とは、調査の全ての対象地域を重複なくしかも漏れなく排反な小地域に区分することで実査精度の担保を図るために設定されたあくまでも調査実施技術上の仕組みである。それは、調査票情報に反映される調査客体が有する場所的規定とは本来的には異次元のものである。

現実には、例えば、世帯、事業所、あるいは企業といった調査客体は、それぞれ特定の存在の場所を持つ。そのような調査客体に係る個体情報が現行の統計では基本的に調査区コード(あるいは基本単位区コード)として統合され、それに属する調査対象単位が一括して積み上げられることで、市区町村あるいは都道府県といった各地域レベルの表章が行われている。調査区は一般に既存の行政区画と整合性を持って設定されることから、調査区からの積上げ集計によって行政区別表章を行う際には、表章の境域について、特段非整合的な問題は発生しない。

近年、様々な小地域統計が整備され、地域についての統計の結果表章面での解像度は飛躍的に改善した。しかし、調査区という場所的規定はあくまでも場所情報を一定の面として捉えたものである。従って、統計客体が本来的に持つ統計単位情報と調査個票に写し取られた調査票情報との間には、調査地点情報についての明らかな情報の乖離が存在する。すなわち、統計の情報特性から見た場合、調査区の設定という実査のための調査技術上の仕組みが、場所的規定に関してその情報量を明らかに減衰させた形で調査票情報に投影させているのである。そこでは、統計対象単位が場所に関して本来的に保有する各地点情報は、一定の広がりを持つ空間(面)情報として、アグリゲートされた形で記載されている。言い換えれば、調査客体の場所的規定が調査区コードや基本単位区コードあるいは町丁・字コードによって集合的に代置された瞬間に、場所的規

定に関して個体が本来的に有する情報からの情報喪失が発生している。

(2)GPS 情報の統計への利用可能性

天頂衛星を含む観測可能な衛星の増加、携帯電話基地局情報や無線 LAN ターミナル等の包括的活用によって、民生用の位置情報システム(Global Positioning System: GPS)についても観測精度の限界とされてきた「最後の 10 メートル問題」が克服されるなど、近年、位置座標情報の観測精度は飛躍的に改善されつつある。さらに GPS 機能付携帯をはじめとして、低廉な GPS 観測端末の普及も著しい。このような技術面での展開をふまえ、社会経済の様々な分野で GPS による位置情報を活用した新ビジネスが簇生している。

現在の GPS は、少なくとも住戸、事業所、企業、それに例えば地価が関連づけられる土地区画(筆)について、平面上の空間位置情報としてはすでに十分な観測精度を有しているといえる。端末機器による GPS 情報の直接取得あるいは住所情報を用いたアドレスマッチングのいずれの方法にせよ、調査区あるいは基本単位区内に所在する調査客体の所在地情報は、多くの場合、GPS 座標として一意的に与えることができる。なお、集合住宅あるいは複合商業施設等の場合、同一の所在地表示に対して複数の住戸や事業所等が該当する可能性があるが、この場合には、一組の GPS 座標情報に複数の住戸等が対応することになる。

このように、調査客体が所在する地点の GPS 座標情報を個々の調査票情報に直接関連づけることによって、個票調査方式の導入以来、調査票情報が潜在的可能性としては持ちながらも技術的には克服できていなかった調査区という場所的規定に関係するくびきから調査票情報をようやく解放することができたのである。なお、個体レコードへの GPS 座標情報の付与のデータ論的な意義について筆者は、GPS 情報の活用による調査票情報の潜在的情報価値の開拓との関連ですでにいくつかの観点から論じている〔森(2010、2011a、2011b、2011c、2011d、2011e、2011f)〕。

(3)位置情報を介した場所的規定の個体レコードへの付加

実査によって調査客体から入手された個々の調査票情報は、それをデータ形式という側面から見れば、個体レコードとして捉えることができる。調査票情報は調査客体の属性情報をはじめとしてその者に直接帰属する情報を持つだけでなく、実際にも実査に先立ってあるいは事後的に様々な情報を付加されている。プレコードされた事業所や企業の個体識別コードや行政区コード、それに調査区コードの他にも、産業や職業等の分類コードが、調査票への記載内容を参照しつつ事後的な格付けによって付与される。この他にも、例えば総務省統計局が所管している住宅・土地統計調査では、調査員による外観確認により、住宅の種類、建て方、構造、敷地に接している道路の幅員等が建物調査票に記入される。さらに、これらの他にも、調査区情報として、都市計画の地域区分、調査区の建ぺい率、容積率、公共下水道の有無、最寄りの交通機関、緊急避難場所、老人デイサービスセンター、医療機関、郵便局・銀行、コンビニエンスストア、保育所、小学校、中学校までの距離等の情報が、調査対象となった市区町村において指定調査区情報一覧表に記入される。調査票情報のいわば外延部分を構成するこれらの変数は、いずれも調査過程で地方自治体職員あるいは調査員といった調査従事者によって把握される情報である。

調査客体からは得られないこういった情報もまた、調査票情報として個体レコードの一部を構成する。それをデータのレコード形式としてみた場合、これらの情報は、調査客体からの報告によって得られる本来の個体レコードのいわば外延部分に相当し、両者は全体として広義の個体レコード

を構成する。言い換えれば、それらは、本来の調査票情報としての個体レコードへの追加的変数として位置づけることができる。個体レコードへの追加的な情報の付加は、これまでの調査においても行われてきた。そして、こういった追加的な情報もまた、調査客体から得られた情報とともに集計され、分析資料として提供されてきた。

これらの情報は、多くの場合、調査の実施業務と直接関連して個体レコードに付加されるが、実査とは独立に事後的に新たな変数の追加が行われる場合もある。例えば、経済産業省所管の商業統計調査の立地環境特性格別統計編では、小売業について、都市計画法に規定された商業地域及び近隣商業地域等の用途地域をそれぞれの地域の立地環境特性を考慮して独自に再区分〔本稿末の付表1参照〕することで、調査結果の表章が行われている〔経済産業省(2007)〕。また、総務省統計局所管の全国物価統計調査の結果報告書には、業態別の価格差や割引・特典サービスの有無別の価格差、価格の週変化などとともに、立地環境特性格別(「商業集積地区」(駅周辺型、市街地型、住宅地背景型、ロードサイド型、その他)、「オフィス街地区」、「住宅地区」、「工業地区」及び「その他」)、さらには競合店舗の有無別の調査結果が掲げられている〔総務省統計局(2007) 579-632 頁〕。商業統計にせよ全国物価統計にせよ、地域の立地環境特性によるゾーニングは、現在のところ立地環境特性格別集計を目的に行われているものである。その意味では、集計表の作成、公表を統計行政上の任務とするわが国の統計作成機関にとって、それ自体提供する調査結果の拡充という点で意欲的試みとして評価できる。

ところで、全国物価統計調査での調査対象店舗の立地特性区分としても用いられている商業統計調査における各商業店舗の立地環境特性づけは、1999(平成 11)年実施の調査までは調査区ベースで行われていた。このため調査区内の店舗は全て同一の立地環境特性に分類区分されてきた。このような調査区ベースでの立地環境特性への同定方式は、2002(平成 14)年調査から改められ、その後の店舗の同定作業は、個々の商業店舗、すなわち事業所毎に行われている〔経済産業省(2007)7 頁〕。

商業店舗の各立地環境特性への同定方式の転換には、二つの意味で統計における次代の展開に関わる画期的な契機が内包されているように思われる。そのひとつは、商業店舗の立地環境特性づけが、それまでの調査区から個々の事業所ベースへと改められた点である。この転換は本質的である。なぜなら、個票調査の導入以来、わが国でもすでに1世紀以上が経つが、この間、調査票情報の場所的規定については、基本的に行政的区分によるものであった。それは、ある意味では調査区からの積上げによる集計の必然的帰結でもあった。情報処理技術の飛躍的発展の中で、メッシュ統計をはじめ多様な小地域集計がその後提供されるようになるが、それらも基本的に調査区という場所的規定についての情報枠組みを前提したものであった。メッシュ統計作成の際に避けることのできない膨大な同定作業はこれを象徴するものである。

それまでの調査区ベースからアドレスマッチングによる個々の事業所ベースへの各立地環境特性への同定方式の変更は、基本的に調査実施の制度的枠組みとしての性格を持つ調査区が、場所的規定に関して統計の集計・利用過程までも支配してきたこれまでの統計作成慣行に対する質的変革をもたらす契機を内包している。なぜなら、新たな処理方式によれば、調査区の境域とは独立に、文字通り地域の特性の空間的広がりの実情に応じて領域設定を行うことができるようになるからである。

第2の論点は、立地環境特性情報の個体レコードへの変数としての付加可能性である。

事後的に獲得あるいは作成した情報の個体レコードへの追加変数としての付加というここでの筆

者の理解は、英国における匿名標本データ(Samples of Anonymised Records: SARs)で行われている変数次元の拡張にその着想を得たものである。

英国では、1991年人口センサス以来、センサスの調査票情報に基づいて匿名標本データ(SARs)が作成され、学術研究や民間利用に供されている。現在、SARsデータの提供窓口機関となっているのが、マンチェスター大学センサス・調査研究センター(Centre for Census and Survey Research)である。同センターでは、センサスの匿名個体レコードに多くの派生変数(derived variables)^[注3]を付加し、追加的な派生変数を伴った個体レコードからなるデータセットが利用者に提供されている[森(1997)]。

このような事後的に付加される派生変数という観点から個々の事業所の場所的規定とその場所の立地環境特性情報の関係を捉え直してみよう。もし個々の事業所を調査区という一定の空間的(面的)広がり我代表するコード情報とではなくまさに個々の位置(所在地)情報と1対1で対応させることができれば、その場所を含む地域の立地環境特性を表現する変数値を、いわば広義の調査票情報として個々の事業所の個体レコードに付加することが可能となる。もちろん、個々の事業所の調査票情報を転写した個体レコードが固有の位置情報を変数として持つ場合、事後的に編成される個々の立地環境区域の境界は調査区界と整合的である必要はない。調査区と立地環境特性による境域が非整合的な場合にも問題なく対処できるという自由度が、個体ベースによる場所的規定の情報面での最大の優位性である。もちろん、複合商業施設のような場所に複数の事業所が集合立地するような場合、調査票情報と位置情報との間にn対1の関係が成り立つケースもありうる。しかし、その場合にも、単に個々の調査票情報に対応するn個の個体レコードに共通の立地環境特性変数値が付与されるだけである。

(4) 付加すべき変数のタイプ

本稿末尾に掲げた付表1からもわかるように、立地環境特性として提案されている区分は、いずれも質的特徴によるものである。このため、この場合、商業店舗事業所の個体レコードに付加する立地環境特性に関する情報は、質的変数として取り扱うのが適当であると考えられる。

個人、家計、地価観測地点等の個体レコードに地点の利便性の指標のひとつとして公共交通機関や商業施設等からの距離を変数として付加する場合にも、厳密な時間距離の測定は、必ずしも容易ではない。このため、実質的にはこれらについても、多くの場合、質的変数として取り扱うのが適当である。

以上のことから、調査客体の位置情報を介して立地環境特性等の要因を変数化して個体レコードに付加することを想定した場合、レコード形式としては、一連の質的変数を付加することによる個体レコードの変数次元の拡張ということになる。図1は、世帯・個人、企業・事業所個体レコードとその拡張レコード形式を模式的に示したものである。

拡張個体レコード(世帯・個人)																					
個体レコード(世帯・個人)													位置境特性変数(派生変数)								
				世帯属性			属性			統計把握項目											
調査 アイ デント	調査 年月	地域 符号	世帯 一連 番号	世帯 員番 号	一 般・ 単身 の別	家 族類 型	.	.	年 齢	性 別	.	.	調査員記入欄		調 査 項 目 1	調 査 項 目 2	.	.	変 数 1	変 数 2	(質的変数)
													項 目 1	項 目 2							

拡張個体レコード(企業・事業所)																					
個体レコード(企業・事業所)													立地環境特性変数(派生変数)								
				企業(事業所)属性			属性			統計把握項目											
調査 アイ デント	調査 年月	地域 符号	企 業 コ ー ド	事 業 所 コ ー ド	設 立 年	資 本 金	産 業 部 門	調査員記入欄		調 査 項 目 1	調 査 項 目 2	.	.	変 数 1	変 数 2	(質的変数)
													項 目 1	項 目 2							

図1 個体レコードとその拡張レコードの模式図

(5) 識別可能性とデータの処理システム

集計量に比べて個体レコードでは、当然のことながら調査客体が特定されるリスクは大きい。そのため、個人や世帯を調査客体とした個体レコードの場合、わが国でも、個体レコードに地域や職業等の分類区分の統合あるいはトップコーディングといった様々な匿名化措置を施した上で学術研究利用に供されている。企業や事業所を調査客体とした個体レコードの場合、匿名化は個人や世帯データに較べて実際には極めて困難である。

立地環境特性を構成する個々の要因の作用の及びうる範囲は、要因ごとに様ではない。このため、調査客体に関わる立地環境特性が仮にダミー変数の形で拡張された個体レコードに付加されたとしても、個々の地点情報にオーバーレイされる変数がとりうるパターンは変数次元の拡張とともに多様でありうる。このため、行政区や調査区コード、さらには住所情報を元にアドレスマッチングによって取得された位置情報を仮に個体レコード情報から削除したとしても、個体の匿名性は必ずしも担保されない。なぜなら、ダミー変数のパターンによっては、それから個体が特定されるケースもありうるからである。企業や事業所ほどではないにしても、場所的規定としては住戸単位で把握される常住者としての個人や家計についても、識別化に関する事情は同様である。

これらの事情を考慮すれば、位置情報を介して拡張された個体レコードから編成されるデータセットについては、統計上の秘密保護との両立を図るためにも、オンデマンド型の解析サービスなど、その使用にあたっては何らかの制度的仕組みの構築が必要であると考えられる。

むすび

本稿では、調査票情報におけるこれまでの調査客体の場所的規定を巡る取り扱いに焦点をあて、マイヤーの統計地理法等の所説も部分的に参照しつつ、個票の導入によって調査票情報が

本来個体がその場所的規定として潜在的に保有してきたものとは異なる、いわば集計量に変換された形で不当に取り扱われてきたことを論じた。そして、調査客体の存在の場によって規定されたまさに個々の場所的規定を媒介することで、調査票情報への立地等の場所と関係づけられた特性変数の事後的な付加による情報価値の拡張可能性について考察してきた。

近代統計の黎明期において支配的な調査方式であった表式調査は、その後、次第に個票調査にとってかわられる。統計調査で用いられる調査票の表式から個票への変化は、単なる調査票の形式の変化以上の意味を持つ。なぜならそこには、統計作成の出発点となる統計単位情報のあり方の本質的転換が見られるからである。

表式調査に対する個票調査の調査方法上の優越について例えば鮫島龍行は、集計組織が未発達段階において支配的であった表式調査の場合、予め集計結果表の表形式が定められており、既存の結果表以外に「調査項目を各種の目的に応じて自由に製表することはできないし、調査上の誤りを修正することもできない」[相原・鮫島(1971) 343 頁]と指摘している。作成、提供される集計結果表のその後の飛躍的拡大、統計の質の改善は、個票調査の調査方式としての優位性を雄弁に物語っている。この点に限っても、個票の導入は、統計調査の史的展開におけるひとつの画期をなすものである。

しかしながら、その後の1世紀以上にわたる個票調査に基づく統計作成の実際を見た場合、個票調査は、それが本来有する意義に照らしてあまりにも過小に評価されてきたように思われる。個票調査を表式調査から本質的に区別する核心的要素は、個票の導入を契機にそれまでの集計量から個体情報が統計の原単位情報となったことである。原情報が各個体と直接関係づけられたことによって、調査客体が持つ場所的規定についてもそれを調査票情報と少なくとも形式的には直接関連づける条件が整ったことになる。しかしながら、実際の統計作成においては、調査区という調査技術上の要件が調査結果の形式までも支配し続けた結果、時間的規定とともに社会集団が本来的に有する場所的規定については、専ら行政区画による結果表章として取り扱われてきた。

統計における場所的規定は、それが潜在的に有する可能性に比して不相応な取り扱いを受けてきたのには理由がある。集計処理に関わる情報技術の未発達もさることながら、調査客体の場所的規定を個々の位置情報として十分な精度で観測できる技術の欠如が決定的であった。逆に言えば、近年のGPSの観測精度の飛躍的な向上によって、個票という調査方式がその導入以来内在させていた可能性を現実のものとしてはじめて顕在化させたのである。

さいごに、GPSが与える位置情報をキー変数として立地環境特性等の特徴づける諸変数を付加することによる個体レコードの変数次元の拡張が持つ認識論的な意義についての若干の指摘を行うことにより本稿のむすびとしたい。

個体レコードに新たな変数を追加することによって、既存の調査データの利用可能性の外延的拡張が可能となる。そのことによって利用者は、集計表あるいは回帰分析において、新たに付加した諸変数を既存の変数とクロスオーバさせることによって、それまでの統計資料では実現できなかった新たな分析を行うことができる。

このようなレコードの変数次元の拡張は、分析の質に対しても有意な影響を及ぼす。ここで回帰分析のモードに依拠すれば、従属変数の全変動(TSS)は、回帰二乗和(ESS)と残差二乗和(RSS)との和、すなわち、

$$TSS = ESS + RSS$$

によって与えられる。上述の生産性分析において、生産性のうち労働と資本の要素生産性によっ

て説明されない生産性の要素がすべて全要素生産性に含まれるのと同様に、回帰分析でも、用いられた独立変数によって説明されなかった変動部分は、すべて残差として取り扱われることになる。

個体さらにはそれから構成される集団は、社会的存在として特定の存在の場所的規定を持つ。しかしながら、上述のようなモデルでは個体の場所的規定要因が独立変数に加えられていないことから、本来、それらの要因によって説明されるべき変動部分も、単なる残差の一部を形成しているに過ぎない。このような場合、場所的規定要因が何らかの説明力を持っていれば、結果的に残差は系統的偏りを持つことになる。

ここで、場所的規定要因を独立変数として追加してみよう。この場合、場所的規定要因によって新たに説明されることになる変動部分を ESSr とすれば、それまで残差として一括されていた場所的規定要因の作用に係る部分を残差から分離することができる。場所に規定された構造的要因を除去後の残差を RSSm とすれば、それまで残差 RSS として一括されていた変動要因は、ESSr と RSSm とに区分される。その結果、全変動は、

$$TSS = ESS + ESSr + RSSm$$

という3つの変動要素に分解されることになる。

ところで、場所的規定要因の独立変数としての追加は、単に RSS から構造的変動要因を除去する以上の効果を持つ。RSS がこの種の要因を内在させていることは、結果的にもとの回帰推計パラメータにバイアスをもたらすことになる。その意味で新たな変数の追加は、それまでの独立変数によって説明されていた部分に対しても、その変更をもたらすことになる。すなわち、ESS そのものも新たに ESSm へと、すなわち、

$$TSS = ESSm + ESSr + RSSm$$

変化するのである。

現行の実証分析に用いられているモデルは、多くの場合場所的規定要因を明示的に取り込んではいない。このことは、逆に言えば、現行の実証分析は、その多くが、場所的規定要因をモデルに取り込んでいないことに起因する推計バイアスを伴っていることを意味する。

実証分析の可能性、さらには分析の質は、第一義的に利用可能な分析資料に依存する。内外を問わず、これまで統計は、個票調査の導入によって理論的には可能であった個体の場所的規定を、現実のものとして統計作成過程に反映してこなかった。このことが、結果的には利用可能な分析資料を通して実証分析の質を制約してきたことになる。

現実に対するより内容豊かで質の高い分析のための分析資料を確保するためには、一方で地理学等の隣接学問領域とも連携しつつ、空間的広がりを持って分布するいわばマイヤーにおける「統計州」のような領域設定を適切に行うとともに、位置情報を媒介することで様々な地域特性情報と関連づけられた拡張された調査票情報に基づく新たな分析的価値を持つ統計の整備が求められる。

[注]

1. 現行の工業地区(233 地区)は、①経済産業省が平成 14 年に実施した工場適地調査の対象地域のうち、事業所数 200 以上の工業地区、および②工業再配置促進法(昭和 47 年法律第 73 号)で移転促進地域に指定された地域のうち東京 23 区および大阪市、からなる。なお、工業地区の見直しは 10 年毎に行われている〔経済産業省(2010)5頁〕。

2. 企業が新たな工場等の事業所の進出地点を決定する場合、進出候補地域に関して様々な情報が収集され、比較検討が行われる。その際にどのような立地条件を判断材料として決定を行っているかは深く経営判断に関わることからそれらが開示されるケースは稀である。山本健児は、BMW 社がドイツ国内外の自治体から工場誘致情報を募り、候補地の選定に当たった際に開示した立地条件の詳細なリストを「BMW 社が提示した新工場のための立地条件」として一覧、紹介している〔山本(2003)〕。そこでは、1. 用地の位置と規模、2. 地形条件、3. 電力・用水・通信の調達と廃棄物処理条件、4. 交通アクセス、5. 近隣の建築状況、6. 最寄り空港、7. 用地の地質、建築阻害条件、8. 建築法上の基準等、9. 地域の人口構成等の雇用環境、10. 学校や病院等の生活環境、11. 用地取得の可能性、12. 部品供給業者の有無といった各項目について、詳細な報告事項が列挙されている。この資料は、自動車産業という製造業の一業種ではあるが、工場事業所の立地と関連する外部環境要因を検討する上でも極めて示唆的であるように思われる〔本稿末の付表2参照〕。

3. 2001 年 SAR の個人、世帯レコードにはそれぞれ以下のような派生変数が付加されている。

個人レコード	教育水準非充足 (E、W、NI)、雇用条件非充足 (E、W、NI)、健康・身体条件非充足 (E、W、NI)、居住水準非充足 (E、W、NI)、世代指標
世帯レコード	世帯内のふだんの居住者数、世帯内の 65 歳以上の世帯員数、世帯内の介護者数、世帯内の虚弱健康者数、世帯内で長期疾病者数、世帯内の成人雇用者数、学期中遊学者世帯、複合民族世帯指標、世帯主の社会階級、世帯内の家族数、家族類型、家族内の被扶養子供数、家族主の性別、家族主の経済的地位、家族主の社会経済統計分類、1 室当りの人数、世帯の居住室占有率、One Number Census による補定世帯員数、補定によるデータ付与数、補定レコード指標、自治体総合指標

E: イングランド、W: ウェールズ、NI: 北アイルランド

(資料) http://www.statistics.gov.uk/census2001/sar_update.asp (2011.6.3 現在)

このように、SARs の派生変数の中には、住戸の立地環境特性を反映した変数はまだ含まれていない。

〔文献〕

Mayr, Georg(1914) *Statistik und Gesellschaftslehre* 大橋隆憲訳 (1943)『統計学の本質と方法』

Žižek, Franz(1922) *Fünf Hauptprobleme der statistischen Methodenlehre*, München und Leipzig, Verlag von Ducker & Humbolt「統計方法論の5つの主要問題」有田・足利・松井訳『フランクフルト学派の統計学』晃洋書房 1987 年

Žižek, Franz(1937) *Wie statistische Zahlen entstehn. Die entscheidended methdischen Vorgänge*. Leipzig, Hans Buske

蜷川虎三(1931)『統計学研究 I』岩波書店

蜷川虎三(1933)『統計利用に於ける基本問題』岩波書店

蜷川虎三(1934)『統計学概論』岩波書店

相原茂・鮫島龍行(1971)『統計日本経済』筑摩書房

森 博美(1997)「イギリスにおけるセンサスマイクロデータの提供」法政大学日本統計研究所『統計研究参考資料』 No.50

山本健兒(2003)「BMW による新規向上立地選択プロセス」『経済地理学年報』第 49 巻第 4 号

総務省統計局(2007)『平成 19 年全国物価統計調査報告』第 3 巻店舗価格編 その1 小売店舗価格に関する結果

経済産業省経済産業政策局調査統計部編(2007)『商業統計表』立地環境特性別統計編(小売業)

森 博美(2009)「統計個票情報の情報特性について」『経済志林』法政大学経済学会 第 76 巻第 4 号

経済産業省経済産業政策局調査統計部編(2010)『平成 20 年工業統計表』工業地区編

Mori, Hiromi(2010)“Exploring Usability of GPSed Records – a data typological approach”, paper presented in Workshop “statistical innovation: use of GPS and GSM data and integration”, Statistics Netherlands, Heerlen, 6 September 2010

森 博美(2011a)「統計調査における地点情報の把握による統計の情報価値の新たな展開可能性について」『経済志林』法政大学経済学部学会 第 78 巻第 3 号

Mori Hiromi(2011b)“Constraints in Use of the Data Due to the Insufficient Obtaining of Location Information and a Breakthrough in Statistics” 『経済志林』法政大学経済学部学会 第 78 巻第 4 号

Mori Hiromi(2011c)“Exploring the Usability of GPSed Records: A data typological approach”『統計学』第 100 号

Mori Hiromi(2011d)“The GPSed Datasets and the Possibility of Exploring the Micro-based Concept of Regional Potentiality”『経済志林』法政大学経済学部学会 第 79 巻第 1 号

森 博美(2011e)「統計を規定する諸要因との関連から見た時空間個体データベースの可能性について」法政大学日本統計研究所『オケージョナルペーパー』No.24

森 博美(2011f)「データ統合の観点から見たデータキャリア情報の統計的利用可能性について」『経済志林』法政大学経済学部学会 第 79 巻第 2 号(印刷中)

本論文は、平成 23 年度日本学術振興会科学研究費補助金 挑戦的萌芽研究「GPS 情報の活用による公的統計の新たな展開可能性に関する多角的研究」(課題番号 23653060)による研究成果の一部である。

付表1 立地環境特性の区分及び定義

特性番号及び区分 商業集積地区細分	定 義
10 商業集積地区	主に都市計画法第8条に定める「用途地域」のうち、商業地域及び近隣商業地域であって、商店街を形成している地区をいう。 概ね一つの商店街を一つの商業集積地区とする。一つの商店街とは、小売店、飲食店及びサービス業を営む事業所が近接して30店舗以上あるものをいう。また、「一つの商店街」の定義に該当するショッピングセンターや多事業所ビル（駅ビル、寄合百貨店等）は、原則として一つの商業集積地区とする。
11 駅周辺型商業集積地区 うち、駅改札内事業所	JRや私鉄などの駅周辺に立地する商業集積地区をいう。ただし、原則として地下鉄や路面電車の駅周辺に立地する地域は除く。
12 市街地型商業集積地区	都市の中心部（駅周辺を除く）にある繁華街やオフィス街に立地する商業集積地区をいう。
13 住宅地背景型商業集積地区	住宅地又は住宅団地を後背地として、主にそれらに居住する人々が消費者である商業集積地区をいう。
14 ロードサイド型商業集積地区	国道あるいはこれに準ずる主要道路の沿線を中心に立地している商業集積地区をいう（都市の中心地にあるものを除く）。
15 その他の商業集積地区	上記「駅周辺型商業集積地区」～「ロードサイド型商業集積地区」までの区分に特性付けされない商業集積地区をいい、観光地や神社・仏閣周辺などにある商店街なども含まれる。
20 オフィス街地区	主に都市計画法第8条に定める「用途地域」のうち、商業地域及び近隣商業地域であって、上記「10商業集積地区」の対象にならない地区をいう。
30 住宅地区	主に都市計画法第8条に定める「用途地域」のうち、第一種・第二種低層住宅専用地域、第一種・第二種中高層住宅専用地域、第一種・第二種住宅地域及び準住居地域をいう。
40 工業地区	主に都市計画法第8条に定める「用途地域」のうち、工業専用地域、準工業専用地域及び工業地域をいう。
50 その他地区 うち、有料道路内事業所	都市計画法第7条に定める市街化調整区域及び上記「10商業集積地区」～「40工業地区」までの区分に特性付けされない地域をいう。

(出所) 経済産業省『平成19年商業統計表 立地環境特性別統計編(小売業)』13頁

付表2 BMW 社が提示した新工場のための立地条件

大項目	具体的条件
1 用地の位置と規模	200～250 haのまとまりを持つ用地
2 地形条件	最高標高、最低標高、道路の標高、最寄りの水面(川、湖)の標高
3 技術的な調達、廃棄処理条件	最寄りの電力(110 kV/40 MW)調達可能地点までの距離 最寄りのガス(6600 m ³ /h)調達地点までの距離 最寄りの工業用水(450 m ³ /h)調達地点までの距離 最寄りの通信(2×各 60AL の PMA、12×光回線、各 34MB/sec)接続地点までの距離 最寄りの工業排水溝(250 m ³ /h)までの距離、その標高 雨水処理施設(下水溝、自然水面)までの距離、標高 廃棄物処理(固形物 2000t/a、汚泥、油脂 1500t/a、薄め液 95t/a)
4 交通	最寄りの鉄道駅、最寄りの鉄道線路までの距離、標高 工場用地に接する道路 最寄りのアウトバーンの特徴 アウトバーンまでの道路距離(物理的距離、時間(分))
5 近隣の建築状況	最寄りの住宅地まで最低限 800m 以上離れていること 外部からの煤煙、粉塵その他の汚染物質が工場用地に来る可能性 半径5km以内で災害が発生する可能性
6 最寄り空港	国内空港、国際空港までの距離(1 時間以内が条件)
7 用地の地質、建築阻害条件	地盤の固さ 建築物を支えうる地盤の深さ 地下水面の高さ 地下に空洞部分や充填部分があるか否か 考古学的に価値あるものがあるか否か 汚染物質が蓄積されているか否か 地震の危険性 上空や地下に構築物(電線、管など)があるか否か 建築物や記念建築物があるか否か
8 建築法(権利)	土地利用計画による指定状況 建設計画 建築基準の規制(高さ 30m以上まで建物を建設できること)、例外条件(煙突の高さなど) 建築許可交付までの期間(2002 年初めには工場建設を開始できること) エコロジー、自然保護、樹木などの条件
9 雇用条件(半径 50 km 以内の地域)	地域名 人口の年齢構成(0-5 歳、6-15 歳、16-25 歳、26-45 歳、46-65 歳、66 歳以上) 人口動態(出生数、1980 年、1990～1999 年、死亡数、自然人口動態 1980 年、1990～1999 年) 住民の学歴(卒業資格なし、基幹学校卒、実業学校卒、大学入学資格保持者、学業継続中) 住民の職業教育受講状況(なし、職業教育、マイスター・テクニカー、大学卒、職業教育受講中) 学校種類別の生徒数 学校卒業者数(1980 年、1990～1999 年)(総数、基幹学校卒資格なし、基幹学校、実業学校、ギムナジウム、職業学校) 大学生(総合大学、専門大学)

	<p>大学生の専攻分野(自然科学および数学、工学)</p> <p>就業データ</p> <p>失業者、失業率</p> <p>金属、電機分野での規定労働時間数(週、有給休暇日数、祝日日数)</p> <p>平均病欠日数(年間、従業員数当り、労働日当り、工業分野)</p> <p>労働争議による操業不可能日数(年間、従業員千人当り、工業分野)</p> <p>産業分野別就業者数(農業、工業、商業、運輸、サービス)</p> <p>工業分野における平均時間給</p> <p>工業分野における平均月給</p> <p>金属・電機部門における付加的な人件費(法律に基づくものと協約に基づくもの。実労働時間当り)</p>
10 生活環境(半径 50 km 以内の地域)	<p>基礎学校、基幹学校、上級学校のある都市までの距離</p> <p>大学のある都市までの距離</p> <p>最寄りの病院までの距離</p> <p>最寄りの中規模都市、大規模都市までの距離</p> <p>医療インフラ(1999 年)</p> <p>人口 10 万人当り医師数</p> <p>人口 10 万人当り歯科医師数</p> <p>人口 10 万人当り薬局数</p> <p>病院病床数</p> <p>犯罪発生率(最新の統計で人口 10 万人当り犯罪件数)</p> <p>ドイツ学校、国際学校の有無(外国の場合)</p>
11 用地取得の可能性(権利を持つ第三者)	<p>用地を一人の土地所有者から取得できるか否か</p> <p>基礎的サービスを請け負うところがあるか否か</p> <p>土地返還要求権利者がありうるか否か</p> <p>用地内の耕区数</p> <p>用地の土地所有者数(個人、農民、企業、公的機関)</p>
12 年間販売額1千万マルク以上の供給企業	<p>半径 10 km 以内における供給企業名、所在地</p> <p>半径 10~50 km 以内における供給企業名、所在地</p> <p>半径 50 km 以内で供給企業が拡大できる見通し</p> <p>半径 50 km 以内に供給企業が新規立地する見通し</p>

(出所)山本健児(2003)「BMW による新規向上立地選択プロセス」『経済地理学年報』第 49 巻第 4 号 65 頁第 1 表(一部修正)

Expansion of Information Potential by Using Location Information and its Analytical Importance

Hiromi MORI

Abstract

Statistical materials and applied statistical methods substantially govern the direction, scope and quality of the analyses. Development of statistical methods applied to respective types of datasets has tapped frontiers of uncultivated information and contributed to the better understanding of the phenomenon of the universe. No innovation achieved in the statistical methods, however, is qualified to bring about new findings beyond the scope allowed by the extent of information potential inherent in the statistical materials. Put differently, ultimate quality and levels which every empirical analysis can achieve are basically governed by information potential that statistical materials applied for them inherently possess especially in terms of their data types.

Official statistics has supplied one after another new types of datasets based on the return questionnaire information collected through surveys in response to the expanding analytical needs. With the aid of newly invented methods developed primarily to process new types of datasets, analysts could to step forward into the arena of cognition that analytical endeavor based on traditional types of datasets could never achieve.

Entities which compose the elements of the society such as individuals, households, enterprises and establishments are the objects in survey taking practices and at the same time the subject of analytical quests. Since they wear social nature for their own sake, their existence is inevitably related to time and special location. By highlighting the discussion on the location constraints that govern each surveyed objects, this paper tried to elucidate the substantial informational deficiency in terms of location information inherent in every type of datasets which government statistics has hitherto offered as analytical materials.

The discussion suggests possible biases caused mainly by such data constraints which are supposed to exist in most empirical studies today. As the natural deductive expansion of the discussion from individual return questionnaire information, this paper also brings forward a new concept of datasets loaded with a set of location-constrained variables and discusses its meaning with regard to the possibilities of less biased estimates.

オケージョナル・ペーパー(既刊一覧)

号	タイトル	刊行年月
1	EC 標準産業分類(NACE)	1990.07
2	On Numerical Calculation Programs of American-type Options Using GAUSS Codes	1998.09
3	わが国の統計体系の現状と課題(I)	2000.03
4	ICD10 における自動車事故による死亡者数の試算	2000.04
5	人口動態統計における交通事故死亡統計の特徴について	2000.05
6	Trends in U.S. Working Hours since the 1970s	2001.07
7	わが国における外国人の国籍別出生率について	2001.09
8	東京の消費構造—東京都生計分析調査	2002.10
9	Wide Variations in Statistics Data Sets on the Same Subjects—Reconsidering the Report of the Indian National Statistical Commission	2003.12
10	日中 1995 年産業別購買力平価の推計	2004.04
11	日本における「統計法」の成立	2005.06
12	「統計法」と法の目的	2005.07
13	諸外国におけるマイクロデータ関連法規の整備状況とデータ提供の現状	2005.09
14	統計に係る個人情報の秘密保護について	2006.08
15	若年層における雇用状況と就業形態の動向—『就業構造基本調査』のマイクロ データによる実証分析	2006.12
16	社会生活行動から見た若年層の不安定就業化・無業化の分析	2008.03
17	国勢調査による従業地把握の展開と従業地別就業データの意義	2009.06
18	無償労働の評価と世帯生産サテライト勘定	2009.10
19	エンゲルとザクセン王国統計	2009.12
20	第一次統計基本計画と政府統計の直面する課題	2010.01
21	エンゲルとプロイセン統計改革	2010.02
22	エンゲルと 1875 年ドイツ帝国営業調査	2010.03
23	調査形態論再論	2011.03
24	統計を規定する諸要因との関連から見た時空間個体データベースの可能性 について	2011.04

オケージョナル・ペーパー No.25

2011 年 6 月 15 日

発行所 法政大学日本統計研究所
〒194-0298 東京都町田市相原 4342
Tel 042-783-2325、2326
Fax 042-783-2332
jsri@adm.hosei.ac.jp
発行人 森 博美