

○倉科 周介, 津野 正朗, 大橋 誠 東京都立衛生研究所

国家は一個の有機体である。国家を人体のごとき生物の個体にたとえれば、細胞に相当する構造要素は国民である。国家という擬似生物と人体との最大の差異は、構造と運用の原理が先験的に個体に内在するのではなく、構造要素である国民の中の特定の時空間的成分によって、後天的、人為的に形成されたものだという点にある。現在の公称個体数169（本邦承認数、国連加盟国数は159）のこの擬似生物を適切に認識する方法を提案する。

### 1. 国勢科学の概念

生物の本質的特性は、有限な存在期間における非休止性不可逆的な変化の場としての連続性にある。こうした対象については、場の時空間的構造特性と、その構造要素の時空間的分布特性との関連を観測することが必要になる。認識の分野としては時空間構造科学（time-space structure science）と呼ぶべきであろう。対象が生物ならば、加齢の過程と死を含めた広義の発生生物学（developmental biology）である。そして国家を対象とするとき、これを国勢科学もしくは国家科学（state science, civil science）と呼ぶ。

日本はいわゆる先進諸国の中でもきわめて民族的単一性の高い国であり、それだけに擬似生物としての個体性も濃厚といえる。さらに国勢科学の基礎資源となる官庁統計資料の蓄積も長い歴史をもつ。ゆえにわが国はいわば巨大な時空間試験管のごとく、国勢科学の好個の研究対象と考えられる。

### 2. 知識生成の一般論

外界の対象を直接観測して得られる知識を、一次知識（primary knowledge）と呼ぶ。一次知識の構造特性は、当該知識の生成にかかわる個々の認識行為の特性によって規定される。時空間的に広範な展開規模を持つ客体を対象とする場合、個々の観測は客体のごく一部の情報を与えるにとどまる。したがって一次知識に依存するだけでは時空間構造科学に必要な認識は得られない。

この制約を克服するために、二次知識（secondary knowledge）の生成を考える。それには個別の直接観測によって得られた多数の一次知識を蒐集し、これを組合せて実在の世界（実世界, real world）における客体の状態を復元した上で、改めてその像を観測する。これは大規模データの構造化と多目的利用の定式化である。この過程を迅速化し、蓄積性と発展性を与えて信頼性の高い二次知識の生産を図るため、われわれはSAGEと呼ぶ知識装置を開発中である。このシステムは国勢科学の推進に寄与するところが大きいと思われる。

### 3. SAGEの構築と基本構造

#### 3-1. 一次知識の蒐集

対象とする客体に関する一次知識を網羅的に蒐集する。蒐集範囲は実世界における客体とそれに関与する要因の推定分布域とする。蒐集した一次知識（データ）は、できる限り原型のまま電子媒体に搭載する。この段階では、内容の修正などは原則として一切行わない（primary data file）。

#### 3-2. 仮想空間の設定

コンピュータの中に実世界の時空間と同型の位置座標系を設定する。これは物理的な時間および空間軸によって定義される。実世界に存在した事象はすべてこの座標系上で固有の位置を占める（real world framework）。さらにこの座標系を基準にして、客体の状態特性の位置を定義する座標系を組込む。人口現象、疾病現象、社会経済現象などを客体とする場合は、生物学的座標系（性別、世代、年齢、傷病類型など）、社会的座標系（職業、組織、その他各種の社会的条件）をデータの特性に応じて選択する。

#### 3-3. データの配置

primary data file についてデータの正誤を照合調整して、運用のためのデータファイルを作成する（revised data file）。これから個別データを抽出して real world framework 中の該当する位置に配置する（real world data file）。この段階では、データの記述単位は一次知識のそれと一致する。

#### 3-4. 実世界の復元

real world data file のデータ構造をもとに、個々のデータを分解、再配置して過去の実世界の像を復元する（real world database）。さらにデータの欠損部分は近傍部分との連続性と全体構造の中での整合性の両面に配慮しつつ推計補填する。復元に用いるデータは一次データではなく、それに基づいて作成された人工的なもの（artificial data）であることに注意する。

#### 3-5. 実世界の観測

こうして復元された実世界の像を新たな客体として、一次知識の生成の場合と同様な観測を行う。像（image）を介しての実世界の間接観測であるところから、これを二次観測、または知識観測と呼ぶ。

### 4. 疾病構造解析システム（SAGE-D）

疾病現象は一種の社会的自然史ともいべきもので、すこぶる広範にして長期かつ多彩な展開と経過を持つため、個別で断片的な観測によって適切な認識を得ることは困難である。また、いわゆる疾病構造の変貌に伴って、厚生行政に対する社会需要も、その規模と構造が大きく変化するとみられる。こうした領域は国勢科学の研究とその応用としての行政科学の実践に好適な対象である。そこで疾病および死亡現象を人口や社会の構造との関係で認識できるデータベースを構築し、これを用いて死亡の時間的、地理的、生物学的、社会的な分布特性と地域社会の時空間構造との関連性の分析や疾病の動向の将来予測などを行う。これを疾病構造解析システム（SAGE-D）と呼ぶ。

## 5. SAGE-Dの資料およびデータ

### 5-1. 基礎資料

#### a. 人口動態関係；出生数，死因別死亡数

日本帝国人口動態統計；1899-1905

日本帝国人口動態統計および日本帝国死因統計；1906-1938

人口動態統計；1939-1943, 1947-

#### b. 人口静態関係；人口数

日本帝国人口静態統計；1903, 1908, 1913, 1918

国勢調査；1920, 1925, 1930, 1935, 1940, 1947, 1950, 以後5年間隔で継続

人口調査；1944, 1945, 1946

#### c. 海外領土人口静態関係

海外領土における国勢調査（台湾，樺太，関東州，朝鮮，南洋群島）；1920, 1925, 1930, 1935

満洲国国勢調査；1935（康德2），1940（康德7）

### 5-2. 基礎データ

性別出生数（年次別）；1872-

性別5歳年齢階級別死因別死亡数（年次別）；1899-

性別1歳年齢階級別人口数；1903, 1908, 1913, 1918, 1920, 1925, 1930, 1935, 1940, 1944, 1945, 1946, 1947, 1950以後5年間隔で1980まで

### 5-3. 欠損，不明，または調査未了部分

全国；1944-46は，出生数不確実，死亡数不明

沖縄；出生数，死亡数，人口数とも1944-70が不明または調査未了

海外在住者，国籍取得者，戦争に関連する人口移動；調査未了

## 6. 死亡とその関連事象の real world image

時間域を1872年以降，空間域をこの時間域の日本国領土とそれに準ずる地域，状態空間の定義座標軸を出生世代（生まれ年）とする real world framework を設定し，この中に性別の生存数や死因別死亡数などの個別のデータを配置する。ただし対象時間域は主として1950年以降とし，死因別死亡数は以下の死因を対象とした（総死亡，自殺，事故，老衰，結核，脳血管疾患，虚血性心疾患，高血圧性心疾患，心不全，糖尿病，肝硬変，腎疾患，全癌，その他の全死因）。また全癌についてはさらに主要臓器別の展開を行う（食道，胃，結腸，直腸，肝，胆道系，膵，肺，腎，膀胱，子宮，乳房，前立腺，ホジキン病，その他の全癌）。以上の各死因による死亡数は総死亡の約70%に達し，数量的に重要な死因はすべて網羅した。なお戦前については総死亡のみを対象とした。つぎに毎年5歳階級で記述されている死亡数を，spline補間法を応用して1歳階級に分割し，所定の位置に再配置する。データの欠損部分はその近傍の状況から推計補填する。

## 7. real world database を用いた疾病現象の巨視的観測

仮想空間内に復元された過去の死亡現象の像（image）をもとに，各種の射影

(projection) を作ってその形状を観察する。

X軸を暦年，Y軸を出生世代（計測単位は出生年）とする二次元位置座標系を定義し，この上にreal world database内の死亡現象の像を射影する．これを世代マップ（generation map）と呼ぶ．これにより対象とする時空間域における死亡現象の共時的（synchronismic）および通時的（diachronismic）な全数観測が可能になる．また特定の世代について表頭を時刻，表側を死因とする二次元分割表を作れば，当該世代が各種の死因によって個体数を減少させつつ，その生涯を閉じてゆく状況が観察できる．これを死因マップ（mortality map）と呼ぶ．

## 8. 知識の表現法

数値表は上述のマップ上の単位区域内の事象の度数を，二次元分割表として表示する．区域相互の量的関係の直観的認識には不便だが，実在の数量を正確に把握できる上，認識の分解能はもっとも高い．

X軸に時刻，年齢などをとりY軸に死亡事象の度数をとってグラフを作れば，実世界におけるそれぞれの位置座標軸上での事象の分布状況を，パターンとして把握できる．このパターンは，特定の死因について，時代，世代，年齢の分布特性に関する特異点を，視覚的に抽出するために有効である．

また疾病の巨視的な病像を適切に認識するためには，地理的分布だけでなく，多様な時空間分布特性の表現法を考案せねばならない．世代マップ上の各領域における死亡数を高さと考え，同水準の高さの領域を結べば，一種の地形図ができる．これを死亡地形図（mortality contour map）と呼び，人口集団内の死因別死亡の発生状況を構造的に観察する手段として有効である．

## 9. 展望

健康を保ち生命を永らえるには人体とその集団の時空間的構造を適切に認識することが最大の要件である．これと同じく，変転常ならぬ国際関係と国内事情の変貌の中に在って国家の安寧と健全な発展を保証するために，国家像の理性的な認識が不可欠であることは論をまたない．国勢科学の必要とされる所以である．そして時空間構造科学の核心は，大数観測の記述統計に基づく既往の実世界像の正確精密な復元によって，万人に共通の認識基盤を確立するところにある．

SAGEは目下開発の途上にある．データベースの内容が人口構造と疾病現象に限られる，用手的処理に依存する部分が多い，など独立した知識生成支援システムとしての成熟は今後の課題として残る．しかしデータベースの記述構造は普遍性が高く，ひろく人間社会全般の構造と機能の記述と分析に適用可能と思われる．また日本を近似的な閉空間と見なせば，その内部構造の位相時空間的分布の記述群は，世界各国や国内諸地域の構造と機能の記述と分析における標準値の役割を果すことも期待できよう．

なおSAGEのデータベース資料の大半は半世紀以上を経過した印刷記録であり，摩耗や損壊によって保存の限界に近付きつつあるものが少なくない．国家的見地から保護と利用の対策が講ぜられることを切望する次第である．