

「厚生 の 指標」 抜 刷

第 27 卷 第 8 号

財団法人 厚生統計協会

地域類型化の方法論的考察

—保健所分類の1つの試み—



福富 和夫* 大隅 昇**
西岡 和男* 箕輪 真澄*
永井 正規*

1 はじめに

地域の保健・医療問題を考える際、まず、地域の社会的特性を正しく把握することが何より大切なことは、議論の余地がなからう。ところで、保健・医療問題に関連する社会的特性には、疾病の発生傾向、医療施設やマンパワー、人口の過疎・過密や高齢化、上下水道や交通網の整備状況など、実に様々なものがあり、これらを表わす指標の数は通常用いられているものだけでも100を下ることはあるまい。これらの情報は、コンピュータ技術の発展と相俟って着々と集積されつつあり、その意味では地域特性のきめ細い把握も可能になってきたといえる。しかし一方、過大な情報量はそのままでは取り扱いが不便であり、問題点の抽出も困難である。したがって、この中から有用な情報を抽出して簡略かつ明解な2次の指標を作ること、これを基に地域を典型的に分類してその全体像を浮き彫りにすること、これらが当面の問題となろう。

上述のように、いくつかの地域特性指標に着目し、互に相似る地域を同一の類にまとめることを、地域類型化とよぼう。残念ながらこの種の実例または方法を論じた文献はきわめて少ない。

ここでの狙いは、地域類型化の方法論について、主として統計学的な観点から検討することにある。もちろん、この種の地域類型化は本来衛生行政的ないし社会学的問題であり、それら

の面からの検討を欠かすことはできないが、それはまた別の機会にゆずることとする。

地域類型分類の具体例に、保健所の型別分類がある。これは保健所管轄地域を、人口と管轄地区面積に基き、格付けを行ったものであり(表1参照)、それなりに地域の社会的特性をよく表現している分類といえるが、保健・医療面の特性を表立って取り入れてはいない。

これら保健・医療に関する指標を導入し、類型分類を行うには、いかなる問題があるか、またどのような統計手法が用いられるか。次節で述べる。また実際の保健所データについて試行、検討した結果を3節で述べる。

表1 保健所型別人口規模別格付表

面積等 km ² 人口	政令 市特 別区						
	未 満	50 ~ 100	100 ~ 250	250 ~ 500	500 ~ 750	750 ~ 1000	1000 以上
I 30万以上 25 ~ 30						U R	
II 20 ~ 25 17.5 ~ 20	U						
III 15 ~ 17.5 12.5 ~ 15							
IV 10 ~ 12.5 7.5 ~ 10					R		
V 5 ~ 7.5 3 ~ 5						L	
3万以下			S				

U:都市型 UR:中間型 R:農山漁村型
L:人口稀薄な地域の型 S:支所型

* 国立公衆衛生院 ** 文部省統計数理研究所

2 地域類型化問題

1) 地域類型分類と特殊地域の指定

さきにも述べたとおり、地域類型化は主要な地域特性指標の値とそのパターンに基いて類型分類を行い、地域全体を構造的に把握するとともに個々の地域を位置づけることであるが、目的によっては単一もしくはごく少数の特殊な指標のみを取り上げ、群を作ることもある。これは類型化というより、特殊地域の指定ともいべきものである。

保健所分類の例では、保健所業務を遂行する上で特殊な配慮を要すると思われる地域として、1)医療機関の極度に少ない地域、2)離島、3)豪雪地帯、4)大都市中心部、5)公害発生地帯、6)観光地などが挙げられる。これらはいずれも、通常の社会指標による分類では、衛生行政の上で問題となる特殊性が埋没するおそれがあるものである。

地域類型化問題においては、類型的分類と特殊地域の指定とを組合せることを、工夫する必要があるだろう。

2) 二次的指標の作成

地域類型化を行う際の統計的問題は、(1)類型化の基礎データとしての社会指標を選択し、それらによる二次的指標を作成すること、(2)地域間の類似性を定義し類型群を形成することの2段階からなる。

社会指標の選択は当然ながら地域類型化の目的に依存する、保健・医療問題と関連する社会指標群を大別すると、おおよそ、つぎの3群になる。

- (1) 疾病の発生傾向、疾病情報の代用として死亡指標も含まれる。
- (2) 保健・医療の供給に関するもの。
- (3) その他の社会指標、人口学的指標、社会・経済学的指標など。

選択された社会指標の数が多数にのぼる場合、

何らかの方法により情報を縮約した二次的指標による表現が望まれるが、このような指標を作る統計的手法に、主成分分析や因子分析がある。すなわち、指標の相関行列にこれらの手法を適用し、相関の情報をできるだけ保存するように少数の主成分ないし因子を抽出してこれを二次的指標とする方法である。地域類型化に関する数少ない報告の中に、上記手法の適用例を2つほどみることができる。

北川¹⁾は、北海道と東京の市を除く、全国522の市について、出生率、死亡率など10の保健指標を求め、これに因子分析を適用して求めた3因子により、都市の類型分類を試みている。また、島尾²⁾は、人口30万以上の保健所区域について、結核のまん延状況に関する24の指標を求め、これより抽出した6因子を用いて地域の評点化を行っている。

これらの手法により作られた二次的指標（または因子）には、その具体的な意味づけの問題が残る。上に引用した2報告ではいずれも、因子について適正と考えられる解釈がなされているが、これは常に成功するとは限らない。

二次的指標の作成法として、恣意的な方法も考えられる。さきに社会指標を3つに大別したが、その各群ごとに主観的な重みづけをした指標を作ればよい。この場合、二次的指標の意味づけは容易であろうが、情報が有効に使われているかどうかは保証されない。

結局、主成分分析等で得られた指標に小幅な修正（恣意的な）を加え、分り易い指標を得ることができれば、これがもっとも実際的であろう。

3) 類型化の方法—クラスター分析

特性の類似度を与える距離関数を定義し、一定手順に従って類似するものどうしをまとめて類型群を形成する手法を、一般にクラスター分析とよんでいる。これには各種の手法が提案されているが〔文献3〕を参照されたい〕、大別すると階層的手法と非階層的手法とがある。前者は系統樹で表現されるような分類の階層的構造を

求めるもので、特性項目数は多くてもよいが、比較的少ない個体数からなる集団の分類に適するものである。後者の代表的なものに、分割最適化型手法があるが、これは少数個にまとめられた指標による多数の個体の類型分類に向けた手法で、地域類型化問題にはこちらを適用するのがよい。

この方法を要約すると、(1)あらかじめ、類別する群の個数を定め、その個数だけの代表点を与える(適宜にあるいはランダムに)。(2)各個体をもっとも近い代表点に配置する。近さは前もって定義された距離関数、たとえば、ユークリッド距離、の値で与えられる。(3)代表点の周りに配置された個体で群を形成し、その平均ベクトルを新たな代表点とする。(4)群のまとまりを測る尺度、たとえば、全平方和に対する群内平方和の比率、を計算する。この値が小さいほど群のまとまりは大きいとする。(5)この基準の改良がみられなくなるまで、(2)~(4)を繰返し、最後に得られた類型群を採用する。

この方法の1つの問題点は、恣意的に与える初期代表点の選び方により、結果の異なる場合が起りうることである。そのため、初期代表点を取り換えて数回試み、結果の安定性を調べることが大切である。

4) 類型分類の実際面からの改良

上記の方法により得られた分類は、指標パターンの類似性関数と、分割の最適化基準に基き機械的に作られるものであり、各類型群の性格(たとえば、「医療供給度大で死亡率低い」群、「人口過密でかつ医療施設少ない」群など)については、そこに含まれた地域の特徴を観察して判定しなければならない。

このためには、分類に用いる2次的指標の意味の明確さが要求される。主成分分析などにより得られたものを加工して分かり易い指標とし、しかも、ほぼ同一の類型分類を与えるものを探索すべきである。

最後に、実際的な見地による分類の手直しがなされるべきであろう。2・1)で述べた地域の特

性による修正、行政的観点からの細分制や合併など、種々の側面からの検討を経て実用的なものに近づけることが望ましい。

3. 統計手法による保健所類型分類の試み

1) 資料

本研究で用いた基礎データは、厚生省統計情報部衛生統計課において編集された「昭和52年度保健所別ファイル」による。このファイルには、全国862保健所の医療指標47項目、疾病・死亡指標14項目、人口、管轄面積、保健所コード、保健所型区分コードなどの情報が盛り込まれている(小笠原支庁も含まれているが、これには型区分コードが付けられていない)。実際の解析にあたっては、特に重要と思われるもの、21項目を取り出し、医療施設数、ヘルスマンパワー数、患者延数、死亡数、出生数などは人口あたりの比率を、乳児死亡数、周産期死亡数は通常の定義どおり出生数あたりの比率を、また、人口と管轄面積からは人口密度を算出して用いた(表2を参照)。

2) 2次的指標の作成

21項目の指標から2次的な指標を作成する方法として、ここでは相関行列について主成分分析を行い、固有値が1を越えた4個の主成分を取り出した。表2に各主成分の因子負荷量(この語は因子分析において用いるのが正しいが、主成分分析においても便宜的に用いることがある)、固有値、累積寄与率を示してある。累積寄与率をみると、第1、第2主成分で56.6%と高い値を示し、第4主成分まで含めた場合、71.2%に達している。4主成分で21指標の持つ情報の7割をカバーできるわけだから充分実用にあえうるとみてよからう。

つぎに、因子負荷量の値を観察すると、第1主成分は医療施設規模、患者数、ヘルスマンパワーに関連するものであることが分る。第2主成分は死亡、出生、人口に関するもので、死

表2 21社会指標の因子負荷量(相関行列の主成分分析による)

社会指標	第1主成分	第2主成分	第3主成分	第4主成分
1 一般病院数*	0.682	0.258	-0.257	-0.073
2 病院の結核病床数*	0.312	0.144	-0.466	-0.269
3 病院の一般病床数*	0.954	0.108	-0.084	-0.010
4 一般診療所病床数*	0.155	0.412	-0.475	-0.120
5 医師数(常勤)*	0.900	0.010	0.211	0.152
6 医師数(非常勤)*	0.622	-0.011	0.308	0.199
7 歯科医師数(常勤)*	0.605	0.080	0.261	0.126
8 薬剤師数*	0.963	0.012	0.144	0.072
9 助産婦数*	0.822	0.002	0.136	0.026
10 看護婦数*	0.946	0.043	0.001	0.024
11 准看護婦数*	0.899	0.166	-0.187	-0.049
12 在院患者延数*	0.830	0.134	-0.383	-0.131
13 総死亡率	-0.186	0.949	0.058	-0.068
14 脳血管疾患死亡率	-0.226	0.849	0.064	-0.071
15 がん死亡率	-0.016	0.826	0.172	-0.155
16 心疾患死亡率	-0.164	0.862	0.047	-0.094
17 乳児死亡率	-0.144	0.284	-0.306	0.757
18 周産期死亡率	-0.121	0.370	-0.236	0.726
19 出生率	-0.121	-0.612	-0.477	0.015
20 人口	0.172	-0.656	-0.316	-0.021
21 人口密度	0.250	-0.424	0.516	-0.078
固有値	7.309	4.576	1.727	1.347
累積寄与率(%)	34.8	56.6	64.8	71.2

注* 人口あたりの数

亡率は正, 出生率と人口は負の負荷量になっている。これは死亡指標として粗死亡率を用いたため, 死亡傾向と人口の年齢構成による影響とが混在していることによるものと思われる。人口が少ない地区は若年層の割合も低く, 粗死亡率が上昇し出生率が低下することになる。第3主成分は人口密度が正で, 出生率, 診療所病床数が負で絶対値が大きい。診療所病床数が多い地域はどちらかといえば病院の少ない郡部にみられ, これが人口密度と逆に位置することは分かるが, 出生率の方は解釈が困難である。第4主成分は明らかに乳児死亡, 周産期死亡に関するものである。

また, 図1-1は第1主成分と第2主成分について, 21指標の布置を示したもので, 各点に付せられた番号は社会指標を(表2の番号を参照のこと), また, 輪で囲まれた点は互に類似する指標群を表わしている。図1-2, 図1-3はそれぞれ, 第1主成分と第3主成分, 第2主成分と第3主成分の布置図である。ここで,

G₁…主として一般病院の施設に関するもの, 在院患者延数も含まれる。

G₂…ヘルスマンパワーに関するもの

G₃…総死亡率および主要死因の死亡率

G₄…一般診療所病床数と結核の病床数

図1-1. 社会指標21項目の主成分分布図-1軸と2軸-

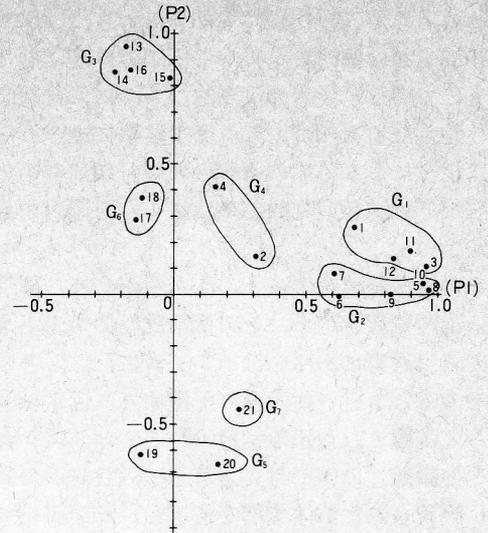


図1-2 社会指標21項目の主成分分布図-1軸と3軸-

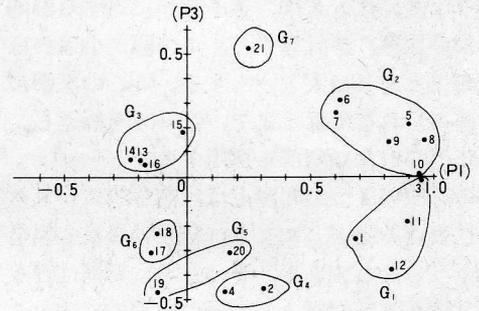
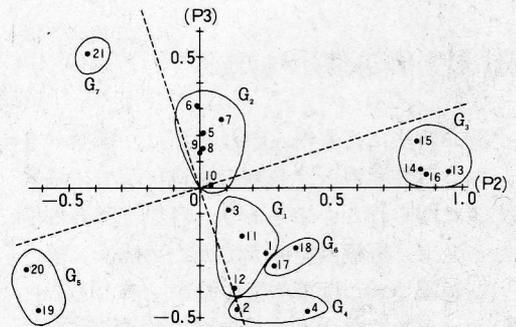


図1-3 社会指標21項目の主成分分布図-2軸と3軸-



G₅…人口と出生率

G₆…乳児死亡率および周産期死亡率

G₇…人口密度

のごとく, まとめられている。

3) クラスタ分析の結果

上記のように求めた2次の指標(第4までの

主成分) について、2.3) 節で述べた分割最適化型手法により(ここで用いたのは k-means 法、これについては文献3) 参照のこと) 862 保健所区域を10の類型群に分類してみた。すでに述べたように、この方法は初期代表点の与え方に問題があるので、初期代表点を換えて5回分類を行い比較検討したが、分類パターンはおおよそ類似したものが得られている。

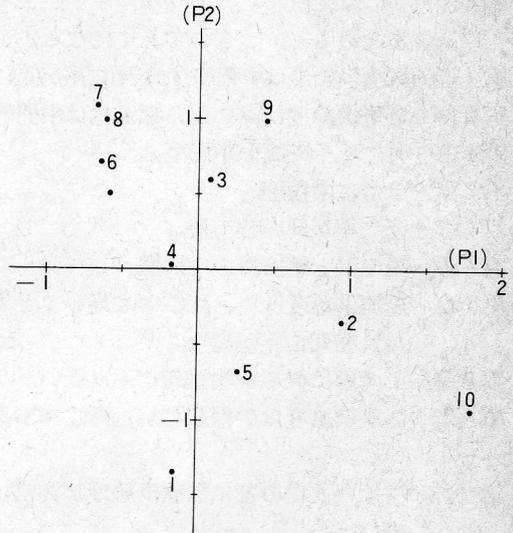
その一例について、従来の保健所型別分類とのクロス表を、表3に示す。これによるとR₄、R₅、Lなど、人口の少ない農村型の地域が、新分類では第3群、第4群、第6～9群に分散され、S型(人口3万以下の地区)も都心にある2保健所地域は、農山村の過疎地域と分離され、U₄、U₅の一部とともに第10群を形成するなど、異なる趣きを見せた面もあるが、第5群のように(ほとんどU型のみからなる)従来の類型によく一致したものもある。

表4は、この10類型群における、4主成分得点の平均を示したものである。また、図2は各類型群の第1主成分と第2主成分の平均をプロットしたものである。先に述べたように、第1主成分は医療施設およびヘルスマンパワーなど医療の供給に関するもので、この値が大きいほど、充足度は高い(これを仮りに医療供給度とよぶことにする)。第10群に属する9地区は全く特殊な地域で、この平均が7.034と極端に高く、図2で第10群の点は右方(はるか先)にはみ出している。

表4 10類型群における各主成分得点の平均

類型群	第1主成分	第2主成分	第3主成分	第4主成分
1	-0.157	-1.355	-0.183	0.174
2	0.942	-0.350	-1.418	-0.361
3	0.070	0.599	-0.869	0.595
4	-0.173	0.008	-0.004	-0.424
5	0.293	-0.663	1.380	-0.036
6	-0.633	0.706	0.315	0.508
7	-0.677	1.080	-0.361	2.334
8	-0.616	0.988	1.041	-1.235
9	0.437	0.989	-0.674	-1.321
10	7.034	-0.940	2.786	1.660

図2 10類型群の布置図—第1主成分と第2主成分の平均値による



第2主成分は主として粗死亡率を表わすものであり、人口老齢化の情報も含んでいる。第1群は第2、第5群と異なり、医療供給度が高くないにもかかわらず、死亡率の低いのが特徴と

表3 従来の保健所型別分類と10類型分類とのクロス表

保健所型	総数	1	2	3	4	5	6	7	8	9	10
総数	861	166	75	87	142	97	115	41	69	60	9
1 U1	54	43	6	1	1	3	0	0	0	0	0
2 U2	72	29	13	1	1	27	0	0	0	1	0
3 U3	67	29	8	1	2	26	0	0	0	1	0
4 U4	52	15	3	1	3	26	2	0	0	0	2
5 U5	23	1	1	1	0	13	0	1	0	1	5
6 UR1	41	9	26	2	4	0	0	0	0	0	0
7 UR2	35	15	4	4	12	0	0	0	0	0	0
8 UR3	23	8	4	2	6	0	0	1	0	2	0
9 R2	11	0	4	3	4	0	0	0	0	0	0
10 R3	48	3	5	6	23	0	4	1	1	5	0
11 R4	170	8	1	37	50	0	32	6	11	25	0
12 R5	95	3	0	12	16	1	29	7	14	13	0
13 L4	19	0	0	5	4	0	6	0	3	1	0
14 L5	112	3	0	11	15	1	32	18	21	11	0
15 S	39	0	0	0	1	0	10	7	19	0	2

注:総数が861であるのは、小笠原支庁を除いてあるため。

いえる。その逆が第9群で、医療供給度が高く、かつ、死亡率も高い。第7群と第8群を分離するには、第4主成分、すなわち、乳児・周産期死亡率をみなければならない。前者は乳児死亡率などが高いが、後者は低い部類に属する。同じく、第2群と第5群も第3主成分により分離される。これはやや解釈の難しい主成分であるが、都市化の指標に近いものと思われる。実際、第5群がほとんどU型のみからなる反面、第2群はUR型が主で、R型も含んでいる(表3参照)。

以上をまとめると、つきのごとくなる。

- 第1群…医療供給度は中程度だが、死亡率が低い。
- 第2群…医療供給度は高いが、都会化は中間的。
- 第3群…死亡率やや高い中間型。
- 第4群…全てに中程度。
- 第5群…死亡率が低い都市型。
- 第6群…第7群と第8群の中間型。
- 第7群…医療供給度低く、死亡率も高い、とくに乳児死亡率が高い。
- 第8群…第7群に似るが乳児死亡率は低い。
- 第9群…医療供給度は中程度だが、死亡率が高い。
- 第10群…少ない人口の割に医療供給度が非常に高い。

4) 問題点の考察

以上の保健所情報による地域類型化の試みについて、つきに問題点を考察しよう。

- (1) 今回導入しなかった社会指標のうち、とくに重要と思われるものとして、人口の年齢構成、年齢訂正した死亡率や有病率、各種指標の年次傾向、水道普及率などが挙げられる。この解析では、粗死亡率を用いたため、死亡傾向と人口の年齢構成が分離できず、二次的指標の解釈が明解とはいかなかった。また、人口および人口構成の年次傾向を導入することにより、過疎化、過密化、老齢化の速さという特性をみることができよう。
- (2) 類型分類に用いる2次的指標のもつ意味を

明瞭にするため、主成分を直交回転することが考えられる。実際、図1-3で軸を回転させた場合(点線のごとく)、各点は多少軸に近づくであろう。それより、人口、疾病量、医療供給などの指標を、別々に作成するのが実際のであろう。また、初めから少数個の重要な社会指標に絞る手もある。ここでは、一般病床数、常勤医師数(いずれも対人口比)、総死亡率、乳児死亡率、人口、人口密度の6項目を用いた解析も試みたが、結果は従来の型別分類と10類型群の分類との中間的なものが得られている。

(3) 類型分類は、本来、地域の指標パターンによる類別をいうが、用いている指標が数量的であれば、各地域の数量的な位置づけ、もしくは順位づけも可能である。施策の数量的評価が要求される場合もあり、つぎの検討課題といえる。

4 むすび

社会指標による地域類型化問題について、方法論、とくに、統計手法の面から考察し、実例として保健所データを用い、その類型分類を行って検討した。実際面からの検討をさらに加えなければ実用に供し得ないが、手法としてかなり有用であると考えられる。

おわりに、この研究を進めるにあたり、データを提供して下さい、また、有益な御助言を下さった、厚生省大臣官房統計情報部衛生統計課の方々に心から感謝の意を表します。

付記

この報告は、昭和54年度における厚生科学研究費補助金による研究結果の一部をまとめたものである。

参考文献

- 1) 北川定謙：保健指標の因子分析と都市の分類，日本公衛誌，14(7)，889，1967
- 2) 島尾忠男：結核のまん延状況と管理を評価する試み，都市保健所問題研究，2号，75，1972
- 3) 大隅昇：クラスター分析はどう使われるか，数理科学，190号，26，1979

厚生指標

8

昭和55年8月
第27巻第8号

★衛生と福祉と保険の統計★

寄贈



財団法人
厚生統計協会