

全国における苗字の空間的偏在とその変化 —失われつつある地域性—

林 利充, 大澤 義明, 小林 隆史

日本は世界で最も多い30万種類の苗字を持つ国である。苗字は地域の歴史・風土の影響を受けているため、その世帯数には地域差が存在し地域性の一指標となりうる。本稿では、2005年電話帳データに基づいて、苗字の空間的偏在を把握し、その偏在と人口流動との関係性について報告する。なお、その地域の苗字構成比率と全国の苗字構成比率との比較を用いて偏在性を示し、人口移動データと地利値の概念を用いて人口流動を表した。

キーワード：苗字、地域性、人口移動、ジニ係数

1. 苗字

1.1 目的

日本では現在から約130年前の1875年、平民苗字必称義務令によって全国民が苗字を持った。日本には約30万種類の苗字があるといわれており、世界で最も多い。苗字はその土地の歴史や風土の影響を強く受けている[1]~[3]。したがって、地域により住民の苗字が異なり地域差がある[4][5]。そこで、本研究では苗字の空間的分布の偏在を計量的に把握する。近年では人口流動性の高まりによりこの偏在性が弱くなりつつあると想像できるので、人口流動の影響についても考察する。まちづくりの基本に地域の個性の活用と地域間交流や連携の促進がある。苗字という側面から地域差がどの程度あるのか、さらに人口移動からどの程度地域差が失われるのかを、将来予測も踏まえて分析する。

本研究では、全国の電話帳を電子化した日本ソフト販売の「写録宝夢巢 Ver. 9」を使用する。データは2005年3月時点のものである。このソフトでは読み方の区別がされていないため、字体が違う場合のみ別の苗字として扱う。ファイル出力ができないため、すべて手入力によって分析対象地域のデータベースを作成した。このソフトに掲載されている全件数を全世帯数(2005年国勢調査)で除すことで2005年時点にお

ける電話帳掲載率を算出した。26,629,957÷49,566,305≒0.537と全世帯数の半数以上を対象としている。

図1は都道府県別に電話帳掲載件数上位3位までを示す。香川では大西(全国100位)が1位、宮崎では黒木(同313位)、甲斐(同310位)が上位2位、沖縄では比嘉(同437位)、金城(同410位)、大城(同403位)が上位3位を占め地域色が見られる。一方で、宮城、東京では順番は異なるが全国と同じである。このように、地域と苗字とはある程度関連している。

1.2 全国上位の苗字

全国上位20位の電話帳掲載件数、全件数に占める構成比、推定人口を表1に示す。最も多い苗字は佐藤であり、以下鈴木、高橋、田中と続く。佐藤の掲載件数は412,169件、全体に占める構成比は1.55%であ

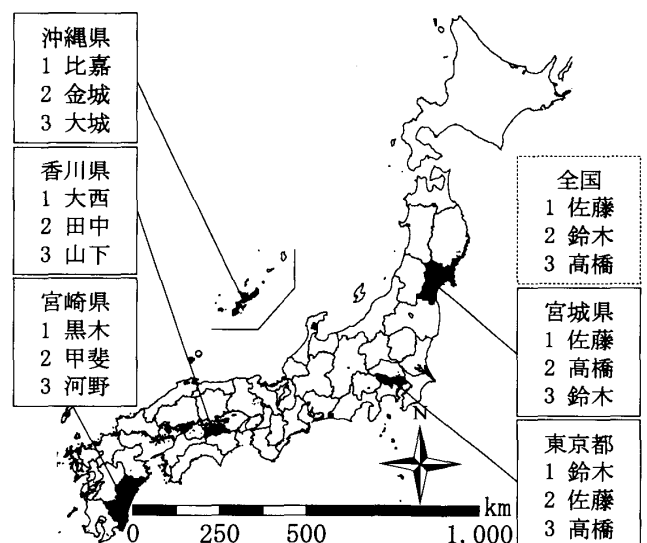


図1 上位3位の苗字

はやし としみつ, おおさわ よしあき,
こばやし たかふみ
筑波大学 システム情報工学研究科
〒305-8573 つくば市天王台1-1-1

表1 全国ランキングの上位10位

順位	苗字	電話帳掲載件数	構成比 (%)	人口 (万人)	同規模の都市	ジニ係数
1	佐藤	412169	1.55	198	名古屋	0.48
2	鈴木	363848	1.37	175	札幌	0.49
3	高橋	302062	1.13	145	京都	0.39
4	田中	281586	1.06	135	福岡	0.20
5	伊藤	230840	0.87	111	広島	0.38
6	山本	228328	0.86	110	〃	0.29
7	渡辺	227931	0.86	109	〃	0.34
8	中村	223020	0.84	107	仙台	0.18
9	小林	217705	0.82	104	〃	0.35
10	加藤	183053	0.69	88	千葉	0.38
11	吉田	176633	0.66	85	堺	0.19
12	山田	173513	0.65	83	〃	0.21
13	佐々木	152851	0.57	73	静岡	0.49
14	山口	136139	0.51	65	相模原	0.20
15	松本	133083	0.50	64	〃	0.26
16	井上	127959	0.48	61	鹿児島	0.29
17	木村	122851	0.46	59	〃	0.23
18	斎藤	118949	0.45	57	船橋	0.58
19	林	116484	0.44	56	〃	0.28
20	清水	111443	0.42	54	姫路	0.27

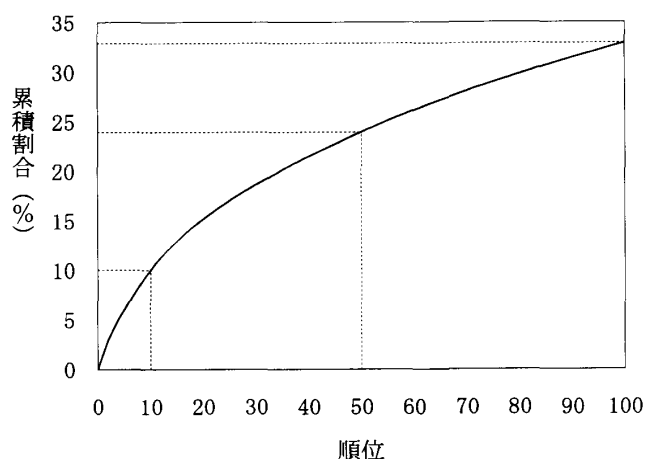


図2 全国100位までの累積割合

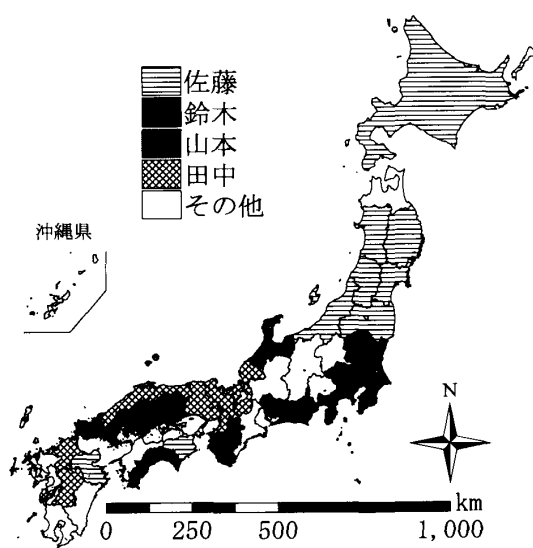


図3 各都道府県の1位

る、2005年国勢調査の全世帯数と全人口の比から算出すると、佐藤の推計人口は198万人となり名古屋市人口にほぼ匹敵する。同様の計算から、鈴木は札幌市、高橋は京都市、田中は福岡市と同規模となり、100万都市に対応する苗字が9種類もあることが分かる。

全国1位の佐藤から100位の大西までの構成比を、順に累積合計したものが図2である。この結果から、日本人の10人に1人(10.0%)が10位以内、4人に1人(24.0%)が50位以内、3人に1人(32.9%)が100位以内の苗字であることが分かる。また、グラフ形状から上位でも傾きが小さくなく、苗字の偏りはロングテールな構成であることが読み取れる。

2. 47 都道府県間比較

2.1 各都道府県の1位

図3は各都道府県の1位を示したものである。北海道、青森以外の東北地方、新潟では佐藤が1位である。これらは地域的に固まっている。群馬を除く関東および

東海地方では鈴木が1位である。一方、西日本では山本を取り囲むように田中が分布している。例外的に徳島、大分の1位が佐藤である。このように、図3から、各都道府県の1位の苗字は連結しており、地理的相関がかなり強いことが読み取れる。

佐藤、田中が9県、鈴木、山本が8県で1位となった。それら以外の苗字が1位になっている県は、群馬(高橋)、山梨(渡辺)、長野(小林)、岐阜(加藤)といった内陸県や青森(工藤：全国64位)、三重(伊藤)、香川(大西)、愛媛(高橋)、佐賀(山口)、長崎(山口)、宮崎(黒木)、鹿児島(中村)、沖縄(比嘉)など列島端部に位置している。これらは、他地域との交流が生まれにくく、独自の苗字が残存していると推測される。

2.2 全国の分布と相関係数

図1では、上位3位の構成が全国に似ている都県と

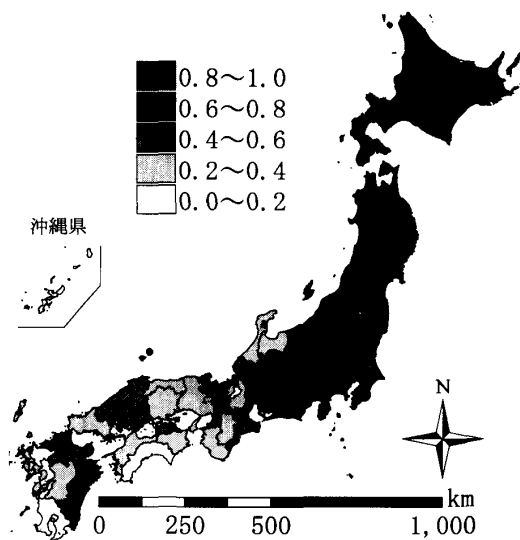


図4 各都道府県の相関係数

そうでない都県とがあった。そこで、全国と各都道府県の類似傾向を相関係数で評価する。データ入力作業が膨大であることから、全国における上位50位の苗字だけに限定する。そして、これら苗字の都道府県内での構成比率と全国の構成比率を求め、それらの相関係数を計算する。

結果を図4に示す。相関係数は東京が最も高く0.963であり、全国の縮図だといえる。これは全国各地との転出入が多いことが要因だと考えられる。次いで、近隣である神奈川(0.938)、埼玉(0.908)、千葉(0.873)の値が高い。そして、東京圏とは別に北海道(0.870)が次に高い。これは明治時代における全国各地からの移住がその要因だと考えられる。逆に、沖縄(0.065)、佐賀(0.171)、高知(0.191)、鹿児島(0.194)、長崎(0.258)、熊本(0.301)と九州地方での相関係数が低い。特に、沖縄は全国の傾向とは無関係だと言えよう。

また、全体的な傾向としては東日本で全国との類似性が高く、西日本で低いことが分かる。これは全国と最も似ている東京との人口移動が西日本より東日本の道県で多いためだと推測される。

2.3 東京までの直線距離との関係

図4で図示した全国との相関係数の空間的分布から、相関係数は東京までの直線距離が近いほど高いように見える。そこで、それらの関係を図5に示す。横軸に東京までの直線距離、縦軸が全国との相関係数であり、47都道府県の値がプロットされている。全体的的に右下がりであり、東京に近いほど全国との相関係数が高く、東京から離れるにつれて相関係数が小さくなる傾

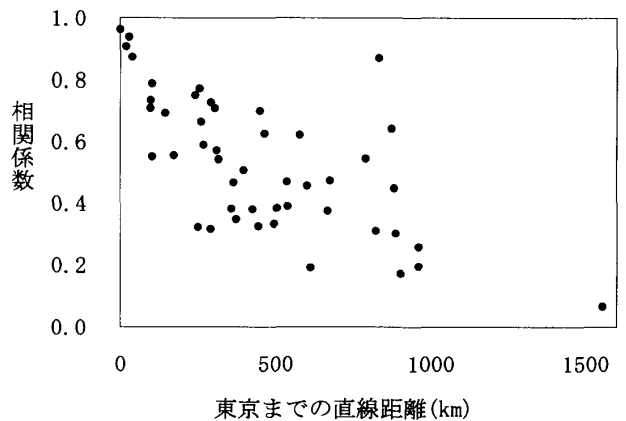


図5 相関係数と東京までの距離

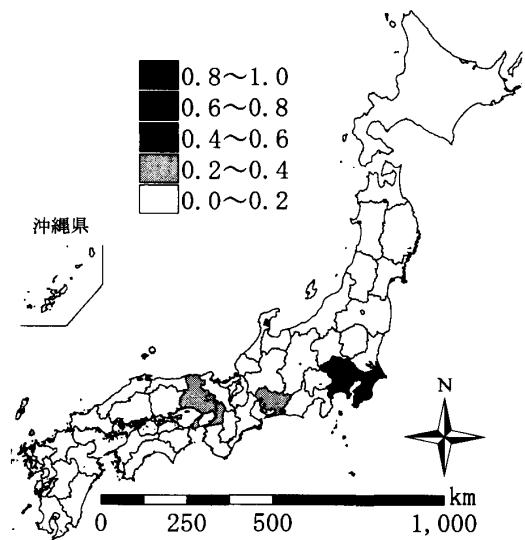


図6 各都道府県の地利値

向をプロット図からも確認できる。

2.4 地利値との関係

各都道府県の人口移動データを用いて、流動性の度合いと全国との類似傾向との関係を見る。ここでは流動性を地利値の概念を用いて評価する[6]~[8]。地利値を用いることで、東京のみならず日本全体から見た、交流の強さを表現できる。地域*i*の地利値を*x_i*としよう。地域*i*と地域*j*との地域間人口移動者数を*T_{ij}*とし、正数λを用いて地利値*x_i*を次の式で表す。

$$\lambda x_i = \sum_j T_{ij} x_j$$

式から人口移動が多いほど地利値*x_i*は大きくなるし、地利値が高い都道府県との人口移動が多ければ地利値*x_i*は大きくなる。また、自身の値は他の都道府県の値にも影響を与えるので、地利値*x_i*は、すべての都道府県間との調整が取れているという意味では均衡条件を満たす。数学的にはλは固有値、*x_i*は固有ベクトルの要素となる。

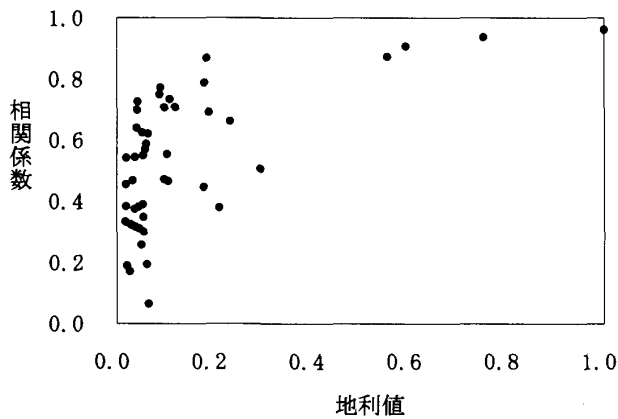


図7 地利値と全国との相関係数

2000年から2005年の都道府県間人口移動者数から算出した各都道府県における地利値を図6に示す。地利値が最大となった東京を基準化しその地利値を1.00とする。神奈川(0.75), 埼玉(0.59), 千葉(0.56)となり東京が突出し, 東京と隣接する県が続く。これに次いで, 大阪(0.30), 愛知(0.23), 兵庫(0.21)など大都市圏を抱える府県の値が高くなっている。人口規模の大きい大阪より南関東3県の地利値が高くなっているのは, 人口移動が最も活発な東京との交流が深いためであろう。

地利値と全国との相関係数の関係を図7に示す。横軸が地利値, 縦軸が全国との相関係数であり, 47都道府県の結果をプロットしている。この図から, 地利値が高いほど, 全国との相関係数が高い傾向がみとれる。このように, 地利値が高い, つまり流動性の高い地域では, 苗字の構成割合は全国に近づくことが理解できる。

3. 苗字の偏在度

3.1 佐藤, 田中の分布

北海道, 東北地方で1位が多い佐藤と西日本で1位が多い田中の各都道府県における掲載件数を図8および図9にそれぞれ示す。佐藤は東日本で掲載件数が多く, 最も多い北海道での掲載件数は40,716件, 次いで宮城, 東京となっている。西日本で掲載件数は少なく, 大分を除いた西日本のどの府県でも1万件以下である。このように佐藤は東日本に偏った分布といえる。西日本に位置する大分で件数が多いのは戦国時代の関東地方からの移住が影響していると推測される。

これに対して, 田中は大阪で最も多く22,203件, 次いで東京, 福岡となっている。この他には北海道, 南関東3県, 愛知, 兵庫など人口の多い都道府県での

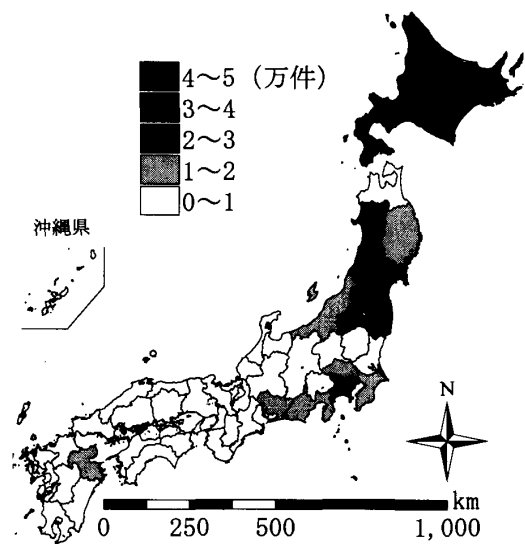


図8 佐藤の掲載件数

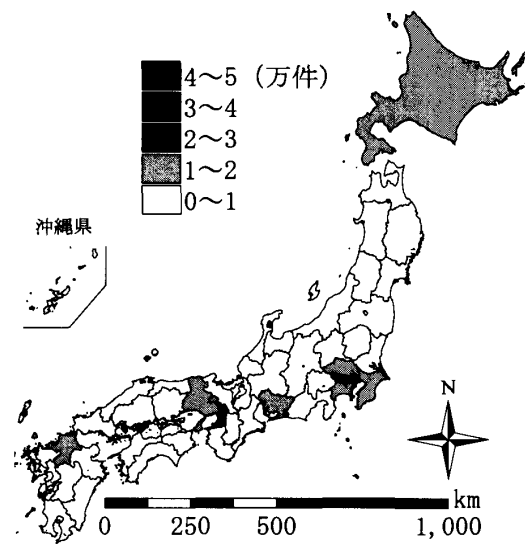


図9 田中の掲載件数

掲載件数が多い。東日本に偏在する佐藤と比較すると極端な疎密がなく, 田中は人口の重みに対して均一に分布していることがみとれる。

3.2 ジニ係数

産業が地域間で偏在する度合いを表す指標として, ローレンツ曲線やジニ係数が経済地理学で頻繁に使われる[9][10]。そこで, 苗字が全国でどれほど偏在しているかをこれらによって評価する。

図10は佐藤, 田中, 比嘉の掲載件数について, ローレンツ曲線を求めたものである。縦軸に掲載件数の累積割合, 横軸に各都道府県の全件数の累積割合をとり, 都道府県単位で傾きの小さいものから並べた。ジニ係数はローレンツ曲線と45度線との間の面積に対応し, その値が大きければ偏在度合いが大きく, 逆に

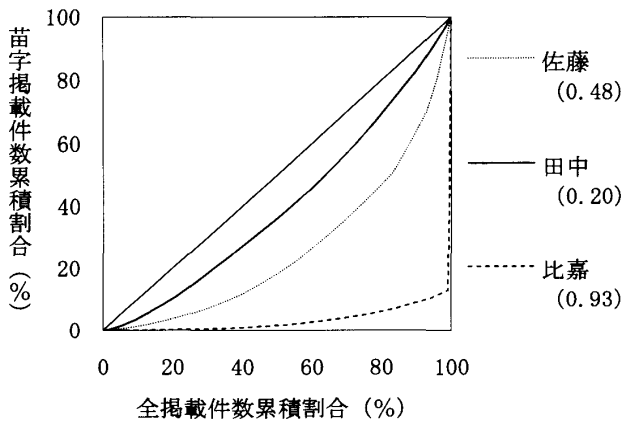


図10 佐藤, 田中, 比嘉のローレンツ曲線

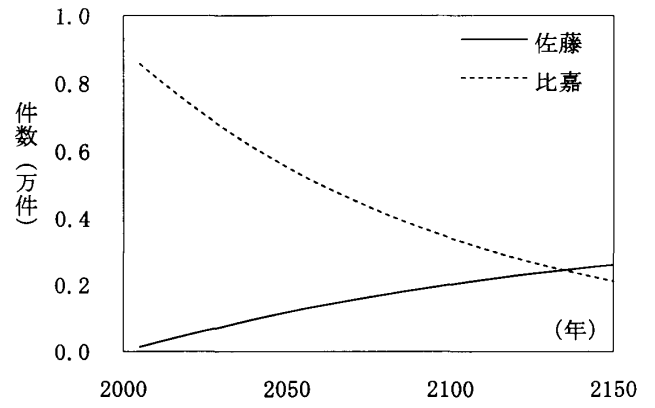


図12 沖縄県内の比嘉と佐藤の推定件数

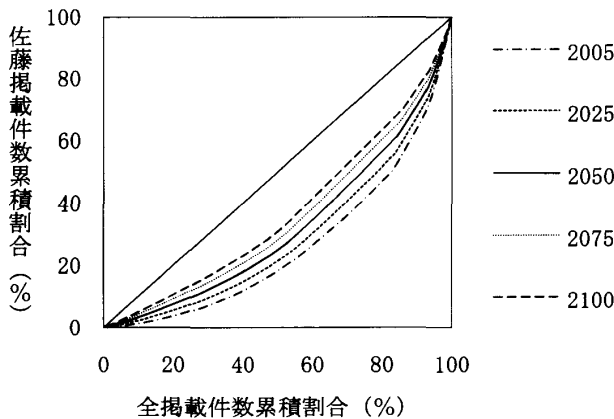


図11 佐藤のローレンツ曲線予測

小さければ偏在度合いが小さいといえる。佐藤は中程度の偏在度合、田中はほぼ偏在がない均一的な分布状態、比嘉は大きく偏っていることが確認できる。

表1にジニ係数も併記する。この表から、ジニ係数の高い順に、斎藤、佐々木、鈴木、佐藤、逆にジニ係数の低い順に、中村、吉田、田中となる。特に、斎藤は0.58であり偏在度合いが強いことが分かる。

3.3 将来予測

2000年から2005年までの都道府県間の国勢調査人口移動データを基に移動率を計算し、移動パターンが苗字とは独立に生じるという前提で、都道府県別の将来の苗字分布が推定できる。このようにして求めた都道府県別苗字分布からジニ係数を計算しよう。

図11は、2100年までの佐藤のローレンツ曲線である。25年単位の変化ではあるが、少しずつ45度線に近づく様子が読み取れる。ジニ係数は2025年には0.42(高橋と同程度)、2050年には0.36(小林と同程度)、2075年には0.31(井上と同程度)、2100年には0.27(清水と同程度)となり、時間とともに、佐藤の偏在が少なくなる様子が確認できる。

図12に同様の計算を用いて推定した沖縄県での佐藤の件数を実線で、比嘉の件数を点線で示す。2005年時点の沖縄では、比嘉は佐藤の63.6倍もいるのだが、佐藤は全国1位の苗字である。人口移動によって将来佐藤が比嘉を逆転すると予想できる。極めて単純な分析ではあるが、その年は実線と点線とが交わる120年後の2130年であることが計算から求まった。人口移動が苗字という地域性を奪う様子が観察できる。

4. 関東449市区町村間比較

4.1 相関係数

都道府県単位での分析と同様に、相関係数を用いて全国との類似性を市区町村単位で評価する[11]。対象地域は2005年時点の関東地方449市区町村とした。全国における上位50位の苗字構成比率を市区町村内で求め、それらと全国とを比較する。図13に相関係数を示す。濃い色の自治体ほど、そこでの苗字分布が全国の苗字分布と相関が高いことを意味する。なお、苗字の偏りは、交通インフラ整備の影響を受けることも予測できるため、鉄道網をも同時併記する。図13から都心部において高く、郊外になるほど低くなることが読み取れる。同時に、茨城、栃木、群馬においては鉄道路線に沿って相関が高い市町村が並んでいることが確認された。つまり、鉄道の有無は苗字の偏在に影響を与えているのである。

また、表2に相関係数の高い10自治体、逆に低い10自治体を抽出して示す。中間に位置する茨城県内の10自治体も加えた。表2から東京都、千葉市区部、横浜市区部、川崎市区部での相関係数が高いことが読み取れる。逆に、栃木、群馬、埼玉の山間部での相関係数が低いことが分かる。ただし、低い方から8番目である川里町の0.197は、前述した沖縄、佐賀、高知、

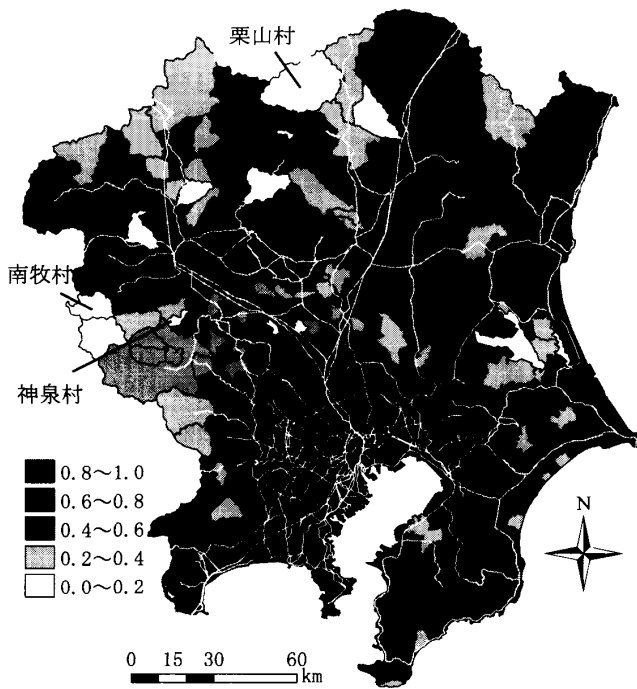


図13 各市区町村の相関係数と鉄道網

表2 全国との相関係数

順位	都県	市区町村	相関係数	順位	都県	市区町村	相関係数
1	千葉	美浜区	0.971	245	茨城	石岡市	0.654
2	東京	多摩市	0.970	280	茨城	常陸太田市	0.600
3	東京	府中市	0.969	291	茨城	水海道市	0.586
4	東京	杉並区	0.965	330	茨城	鉾田町	0.513
5	神奈川	保土ヶ谷区	0.963	421	茨城	坂東市	0.280
6	神奈川	宮前区	0.963	:			
7	東京	町田市	0.961	440	群馬	東村	0.202
8	神奈川	青葉区	0.961	441	埼玉	大里町	0.202
9	東京	日野市	0.960	442	埼玉	川里町	0.197
10	東京	世田谷区	0.959	443	群馬	榛名町	0.148
:				444	栃木	塩谷町	0.142
125	茨城	取手市	0.874	445	群馬	赤城村	0.139
138	茨城	守谷市	0.852	446	群馬	上野村	0.132
147	茨城	土浦市	0.830	447	群馬	南牧村	0.093
182	茨城	水戸市	0.759	448	埼玉	神泉村	0.088
189	茨城	つくば市	0.746	449	栃木	栗山村	0.079

鹿児島値を超えており、関東と九州ではかなり傾向が異なることが再確認できた。また、茨城県内では、県南や人口の多い市での相関係数が高く、逆にその他の市町では低い。

4.2 人口密度との関係

図14の縦軸に全国の構成比率に関する相関係数、

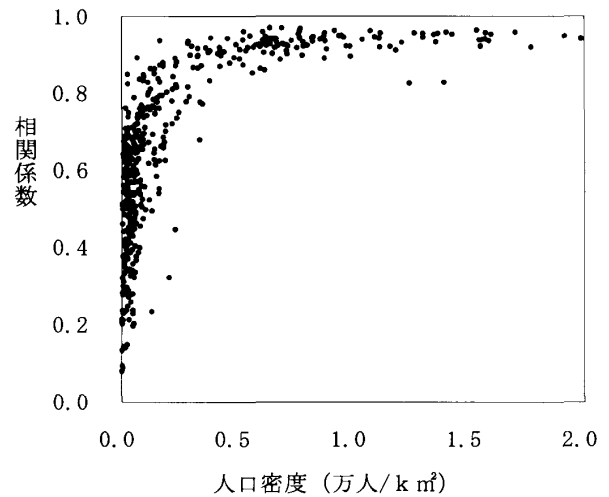


図14 全国との相関係数と人口密度

横軸に各自治体の人口密度をとり、449市区町村の値をプロットする。図14からデータ間には正の関係があることが読み取れる。特に、密度が5,000人/km²を超えるとほぼ相関係数0.8を超える。したがって、人口密度の高い市区町村では全国と似た傾向を持つと考えられる。これは、人口移動の結果だと推察できる。

5. まとめ

本研究では2005年電話帳データを用いて、苗字の空間的偏在と人口流動性との関係を考察した。大雑把で素朴な分析ではあるが、以下の知見を得た。

第一に、日本人の1割が10位以内、4分の1が50位以内、3分の1が100位以内の苗字であることを示した。第二に、日本全体の苗字上位の構成比の比較から、東京がほぼ全国の縮図であること、東京の隣接県に加え北海道でも全国傾向が強いことが分かった。逆に、沖縄、佐賀、高知、鹿児島といった九州ではかなり低く、これらの傾向が東京からの距離や地利値で説明できることを示した。第三に、関東市区町村単位の分析から、都心に近いほど全国との類似性が高く、郊外になるほど低いことが分かった。鉄道沿線の人口密度の高い市区町村ほど全国との類似性が高いということも確認できた。

謝辞 南山大学で開催された「都市のOR 2006」の出席者、腰塚武志先生、糸井川栄一先生、斎尾直子先生(筑波大学)を含め多くの方からコメントを頂戴しました。また、本研究の一部は科学研究費(萌芽研究:20651041)の援助を受けました。記して御礼申し上げます。

参考文献

- [1] 丹羽基二：日本の苗字読み書き事典, 柏書房, 1994.
- [2] 丹羽基二：日本人の苗字, 光文社, 2002.
- [3] 森岡浩：名字の謎, 新潮社, 2002.
- [4] 森岡浩, 村山忠重：名字の新聞, 宝島社, 2005.
- [5] なるほど知図帳 2007, 昭文社, 2007.
- [6] 鶴飼孝盛, 栗田治：交通網により生成される都市平面上の地利値分布, 都市計画論文集, 38-3, pp. 163-168, 2003.
- [7] 鶴飼孝盛, 栗田治：地域間の相互作用に依存した都市活動分布の理論と連続平面への一般化, 都市計画論文集, 40-3, pp. 103-108, 2005.
- [8] 大澤義明, 林利充：隣接グラフと地利値最大化, 日本建築学会計画系論文集, 633, pp. 2417-2424, 2008.
- [9] P. Krugman: *Geography and Trade*, MIT Press, 1991.
- [10] 尾崎尚也, 大澤義明, 歳森敦：公共図書館の空間的不平等生徒利用可能圏の把握, 地域施設計画研究, 19, pp. 107-112.
- [11] 林利充, 大澤義明, 小林隆史 (2007)：苗字の空間的分布に関する研究, 日本オペレーションズ・リサーチ学会 2007 年秋季アブストラクト集, pp. 186-187.