

博 士 論 文

確率モデルの変数制御性を利用した人口分布の再現

指導教官 林 幸雄 准教授

北陸先端科学技術大学院大学
知識科学研究科

富田 真治

2008 年 3 月

Abstract

The congestion and depopulation are serious problems in the modern Japanese society. A specific area develops greatly, and why does not another area do it? The evolution of the population seems to change with time due to various factors such as economy, environment, and policy, etc. The mechanisms of population dynamics have been studied in demographic and economic contexts, and there are international comparisons. The advantage of these studies is to be able to argue about concrete factors on the basis of geographical and social conditions. However, it is difficult to show a unified view, because individual circumstances in each country are different. Now, therefore, the present conditions have many hypotheses, but there is little inspection. In this study, we treat these factors as stochastic variables and discuss the perspective of population dynamics by statistical-mechanical approach.

The purpose of this study is twofold. First, we explain the perspective of the population distribution. We have investigated the cumulative population distribution of (all) municipalities and prefectures, and they seem not to have changed so much in these 26 years. This appears similar to (1) the double Pareto distribution, which is closely matched to the head of a log-normal distribution, and the tail of a power law distribution or (2) the single power law distribution. Second, we investigate the origin of the distribution by using simple stochastic models (Random Multiplicative Process (RMP) and Preferential Urn Model (PUM)). These stochastic models can generate a power law distribution. We found the parametric control methods of these models. Using this controllability, we explained the origin of the cumulative population distributions. In these models, randomness plays an important role in generating a power law distribution. In the RMP, we assume the growth rate and unexpected comings and goings of each area are uniform. This means all areas are equality. In the PUM, we assume each area has different attractiveness and people move depending on them. This means people can move freely anywhere. In other words, discrepancy in population inter-areas naturally arises in the society where free growth and movement were guaranteed. It is very interesting to be able to reproduce the population distribution by such simple mechanisms.

This controllability of RMP and PUM can be applied to a power law distribution in various other fields related to complex systems and social networks.

目 次

第 1 章

序論	1
1.1 研究の背景と目的	2
1.2 本論文の構成	12

第 2 章

人口に関する研究	13
2.1 はじめに	14
2.2 人口に関する研究の俯瞰	14
2.2.1 人口問題の認識	14
2.2.2 人口学の研究領域	15
2.2.3 人口移動のモデリング	17
2.3 還元論に基づくアプローチの限界	20
2.4 本研究の戦略：統計力学的アプローチ	21
2.5 まとめ	25

第 3 章

自然界と社会に見られる共通の分布	27
3.1 はじめに	28
3.2 サイズ分布とは	28
3.2.1 確率分布	28
3.2.2 累積分布	31
3.2.3 累積分布とランクサイズプロット	33
3.3 これまでに報告された事例	36

3.4 データの再発掘	38
3.5 まとめ	45

第4章

人口データの分析	47
4.1 はじめに	48
4.2 人口統計データ	48
4.3 自治体のサイズ分布	49
4.3.1 都道府県の人口サイズ分布	50
4.3.2 市町村の人口サイズ分布	52
4.4 移動量のサイズ分布	54
4.5 まとめ	56

第5章

成長モデル	58
5.1 はじめに	59
5.2 成長率とジブラ則	60
5.2.1 自治体の成長	60
5.2.2 ジブラ過程	63
5.3 モデル:加算ノイズを伴った乗算的確率過程	65
5.3.1 初期状態と成長過程	67
5.3.2 定常状態	71
5.4 変数の制御性	71
5.4.1 べき指数 β の操作による分布形の制御	71
5.4.2 ノイズ操作による左右への分布シフト	73
5.4.3 サイト数変更による分布幅の拡大縮小	75
5.5 データフィッティング	76

5.5.1	フィッティングの手続き	76
5.5.2	シミュレーション結果と実データのフィッティング	78
5.5.3	変数間のトレードオフ関係	81
5.5.4	ノイズの組み合わせと実データの比較	82
5.6	まとめ	86

第 6 章

移動モデル	87
6.1 はじめに	88
6.2 モデル: 壺モデル	88
6.2.1 初期状態	91
6.2.2 定常状態	91
6.2.3 変数の制御性	92
6.3 データフィッティング	93
6.3.1 フィッティングの手続き	93
6.3.2 シミュレーション結果と実データのフィッティング	95
6.4 定常状態に収束した後の移動量の比較	97
6.5 まとめ	101

第 7 章

結論	102
7.1 本研究のまとめ	103
7.2 今後の研究課題	106

付録	107
----------	------------

付録1 コーホート要因法	108
--------------------	-----

付録 2 地域人口分布の推移	110
付録 3 人口移動ネットワークの可視化	113
3.1 影響力の地図	113
3.2 移動空間地図	115
付録 4 市町村のサイズ分布	120
謝 辞	121
参考文献	122
研究業績	128

第 1 章

序論

1.1 研究の背景と目的

世界人口白書¹によると, 2007 年現在, 世界には 66 億の人々が暮らしている. 図 1.1 は, 世界人口の推移である. 横軸は西暦年, 縦軸は人口²を表している.

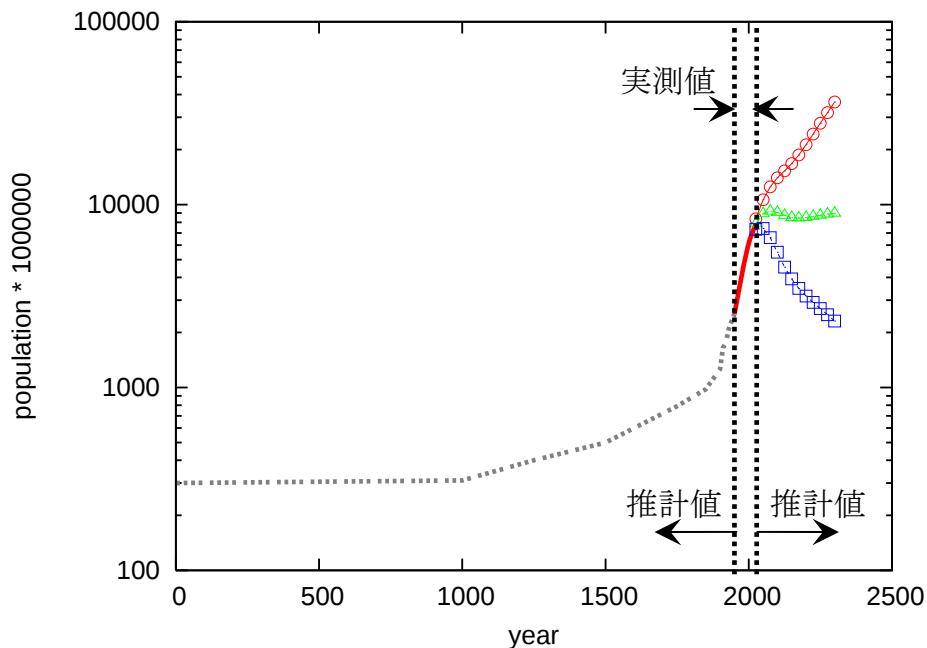


図 1.1 世界人口の推移³

(...):推計, (—):実測値, (⊕):推計(高), (△):推計(中), (□):推計(低)

¹ <http://www.unfpa.or.jp/4-1.html> (2007.10.27 参照)

² 「人口 (population)」は, 常識的には人間の数ということになるが, 人口研究者の定義にも様々なものがあり, 統一された定義はない. 舘は人口を「人間の集合体, あるいは集団である」と定義している[舘 1960]. また, 山口は「何らかの標識によってとらえられる人間の集団」としている[山口 1989]. この「標識」を客観的または経験的に与えることにより具体的な人間の集団を規定できる. たとえば, 「2007 年(時間)に日本(地域)にいる」という標識によって「日本人口」が規定できる. 同様にして, 都道府県人口, 市町村人口が規定できる.

³ 国連資料をもとに作成.

過去の推定および近年の実際の人口:

<http://www.census.gov/ipc/www/worldhis.html> (2007.10.27 参照)

2025 年以降の将来推計:

<http://www.un.org/esa/population/publications/longrange2/WorldPop2300final.pdf>
(2007.10.27 参照)

1950年以前は推計値、1950年から2005年は実測値、2025から2300年は、高位、中位、低位の将来推計値である。中位推計によると、2050年には89億人になると予測されている。西暦1年頃に約3億だった人口は、1950年に25億になった。つまり1950年間かけて8倍に増えたことになる。それが、1950年から2007年のわずか57年間で3倍も増えている。21世紀の人口の増加は、「人口爆発」に例えられている。

この人口の増加は、今後もしばらく続くと予想されているが、先進地域⁴と途上地域⁵では様相がまったく異なる。途上地域には、人口爆発と呼ばれるほどの人口増加が続いている国⁶もある。一方で、多くの先進地域では出生率の低下による少子・高齢化が大きな問題になっている。図1.2は、先進地域と途上地域の人口割合の推移である。横軸は西暦年、縦軸は割合を表している。

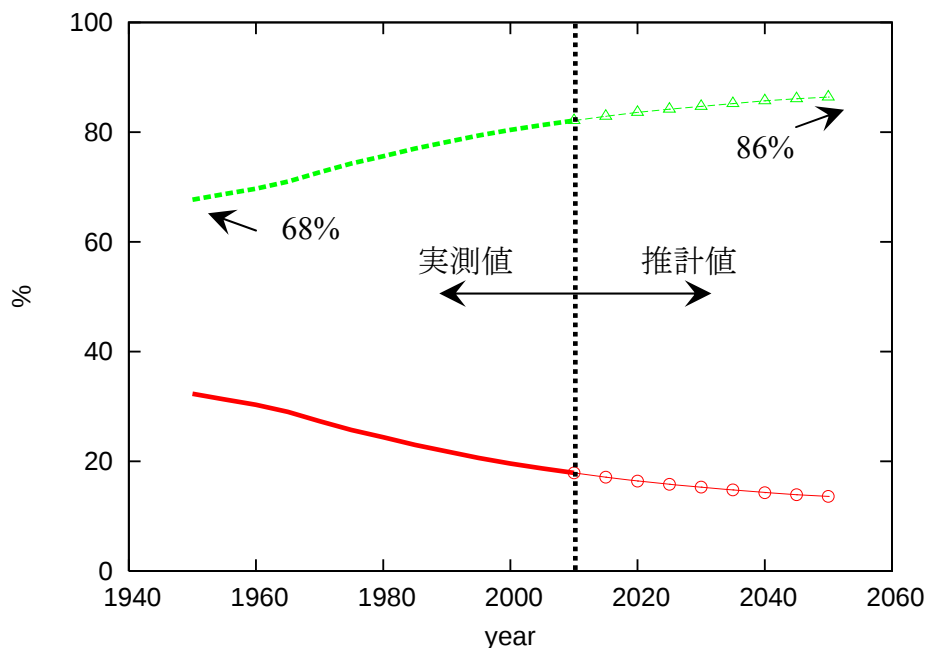


図 1.2 先進地域と途上地域の人口割合の推移⁷

(○) : 先進地域, (△) : 途上地域

⁴ 「先進地域」は、北部アメリカ、日本、ヨーロッパ、およびオーストラリア、ニュージーランドを含む。

⁵ 「途上地域」は、アフリカ、ラテンアメリカのすべての国々、日本を除いたアジア、そしてオーストラリア、ニュージーランドを除いたオセアニアを含む。

⁶ インド、ソマリア、エチオピア、東南アジア諸国など。

⁷ 総務省統計局 2-1 世界人口の推移(1950～2050年)をもとに作成。

<http://www.stat.go.jp/data/sekai/02.htm> (2007.10.27 参照)

1950年の段階で途上地域の人口は17億1000万人で、世界人口に占める割合は68%であったが、2000年には48億9000万人と飛躍的に増大し、80%を超えた。2050年には、さらに増大して86%になる見込みである。これに対して、先進地域の人口は緩慢にしか増加せず、世界人口に対する割合は次第に減少していくと予想されている。

表1.1は、世界の人口上位10カ国である[United Nations 1998]。1950年では、1位中国、2位インドだが、3位アメリカ、4位ロシア、5位日本、7位ドイツ、9位イギリス、10位イタリアと先進地域の国が6つも入っていた。しかし1998年では、途上地域の国7つに対し、先進地域の国はアメリカ、ロシア、日本の3つである。2050年には、先進地域の国はアメリカただ1つで、他の9つはすべて途上地域の国になる。日本は17位になると予想されている。

表 1.1 人口上位 10 カ国(国連中位推計)⁸ (単位:1000 人)

1950 年			1998 年			2050 年		
順位	国 名	人口	順位	国 名	人口	順位	国 名	人口
1	中 国	554760	1	中 国	1255698	1	インド	1528853
2	インド	357561	2	インド	982223	2	中 国	1477730
3	アメリカ	157813	3	アメリカ	274028	3	アメリカ	349318
4	ロシア	102192	4	インドネシア	206338	4	パキスタン	345484
5	日 本	83625	5	ブラジル	165851	5	インドネシア	311857
6	インドネシア	79538	6	パキスタン	148166	6	ナイジェリア	244311
7	ドイツ	68376	7	ロシア	147434	7	ブラジル	244230
8	ブラジル	53975	8	日 本	126681	8	バングラディシュ	212495
9	イギリス	50616	9	バングラディシュ	124774	9	エチオピア	169446
10	イタリア	47104	10	ナイジェリア	106409	10	コンゴ	160360
						⋮		
						17	日 本	104921

⁸ 資料: United Nations 1998

実際、先進地域の一つとされる日本の人口は、2004年の1億2783万8000人をピークとして、2005年から減少に転じた。そして、かつてない少子高齢化社会を迎えようとしている。このような急速な人口構造の変化は、未だどの国においても経験のないことである。図1.3は、日本総人口の推移である。横軸は西暦年、縦軸は人口を表している。1947年から2005年までが実測値で、2006年から2055年までが推計値である。この推計値には、高い出生と低い死亡を仮定した最も楽観的なシナリオ（出生高位死亡低位）、低い出生と高い死亡を仮定した最も悲観的なシナリオ（出生低位死亡高位）、その中間に位置するシナリオ（出生中位死亡中位）という3パターンをプロットした。今後数十年間は、出生死亡についてどのような仮定をしても、人口減少を免れることはできないことが分かる。

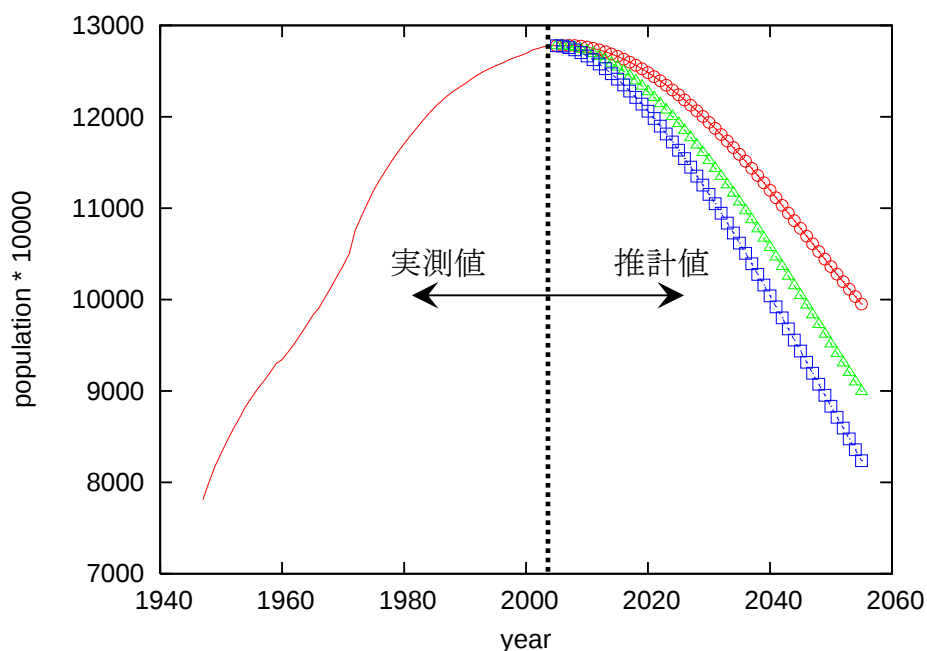


図1.3 日本総人口の推移(1947～2055年)⁹

(—):実測値, (○):出生高位死亡低位, (△):出生中位死亡中位, (□):出生低位死亡高位

⁹ 国立社会保障・人口問題研究所の将来推計人口をもとに作成。出生率、寿命、人口の流入・流出に仮定値を設け、将来の人口を推計している。

<http://www.ipss.go.jp/syoushika/tohkei/suikei07/suikei.html> (2007.10.27 参照)

ところで、人口減少は、地方、都道府県や市町村、さらには集落といった多様な空間的スケールで、それぞれ複雑な様相を呈して進行するものと考えられる。特に、中山間地域の小規模自治体は、平成の大合併によってその数が急減し、集落消滅の可能性さえ高まっており[菅沼 2005]、過疎問題¹⁰や、それに伴う地域間の格差問題¹¹が注目を集めている。このような状況を背景として、わが国における人口移動やそれに伴う人口の地域分布変化の背後にある基本メカニズムを解明し、それらに強く影響を与える要因を明らかにすることが、今後の国土政策や地域政策を策定する上で、非常に重要な課題であると考えられている[丹保 2004]。人口変動が人々の関心を引き、その原因についての研究や、結果として起こるさまざまな社会的、経済的諸問題の研究の必要性が高まっているのは、それが私たちの生活に不都合をもたらし、解決しなければならない問題として認識されるからである。

今後、日本全国の自治体の人口はどのように推移していくのであろうか？まんべんなく人口が減っていくのであろうか？それとも、より便利な大都市圏への人口集中が進むのであろうか？市町村レベルでの推計はないが、都道府県レベルの推計は、国立社会保障・人口問題研究所がコーホート要因法¹²を用いて行っている。図 1.4 は、

¹⁰ 「過疎」は、「過密」の対概念をなすものとして 1960 年台から公の場に多く登場するようになった。また、「過疎問題」という言葉は、人口流出そのものだけでなく、むしろそれを契機として生じるさまざまな問題の総称を意味している。例えば、安達は過疎について、「農村人口の流出が大量かつ急激に発生した結果、その地域に残った人々の生産と社会生活の諸機能が麻痺し、地域の生産の縮小とムラ社会自体が崩壊に向かって作用してゆく悪循環過程である」と定義している[安達 1970]。

¹¹ 戦後、地域間格差の縮小(あるいは解消)は、地域政策の目標の一つとして、広く承認されてきたが、その重要度および政策手段については、意見の相違があった。地域間格差縮小のためには政府による積極的な政策的介入が必要である[Myrdal 1959, Hirschman 1963]という意見と、経済成長が進めば、市場メカニズムの作用によって、自動的に地域間格差は縮小する[Borts 1965, Mera 1978]という意見である。現在も、雇用機会や所得など、経済面での格差が、さまざまなメディアを通じて指摘されている。しかし、人々が感じているほど統計的には地域間格差の拡大傾向を明確に認められないという報告もある。例えば、日本総研の報告

http://www.jri.co.jp/press/2006/jri_060927.pdf (2007.10.27 参照)

¹² この方法は、ある年の男女・年齢別人口を基準として、ここに人口動態率や移動率などの仮定値を当てはめて将来人口を計算する方法である。詳しくは付録 1 で説明する。

都道府県人口の推移である。横軸は西暦年、縦軸は人口を表している。1920 年から 2005 年までが実測値で、2010 年から 2035 年までが推計値である。人口の少ない県の推移は狭い範囲に密集しているの、人口の多い 9 県のみ推移との対応を示した（例—：東京）。第 1 回国勢調査が実施された 1920 年に、人口が 200 万人以上を記録したのは、47 都道府県のうち、6 都道府県のみであった。これが 2006 年には 20 になったが、2050 年には 13 に減ると予測されている。そして、将来的には全ての都道府県において、多かれ少なかれ人口が減少する。

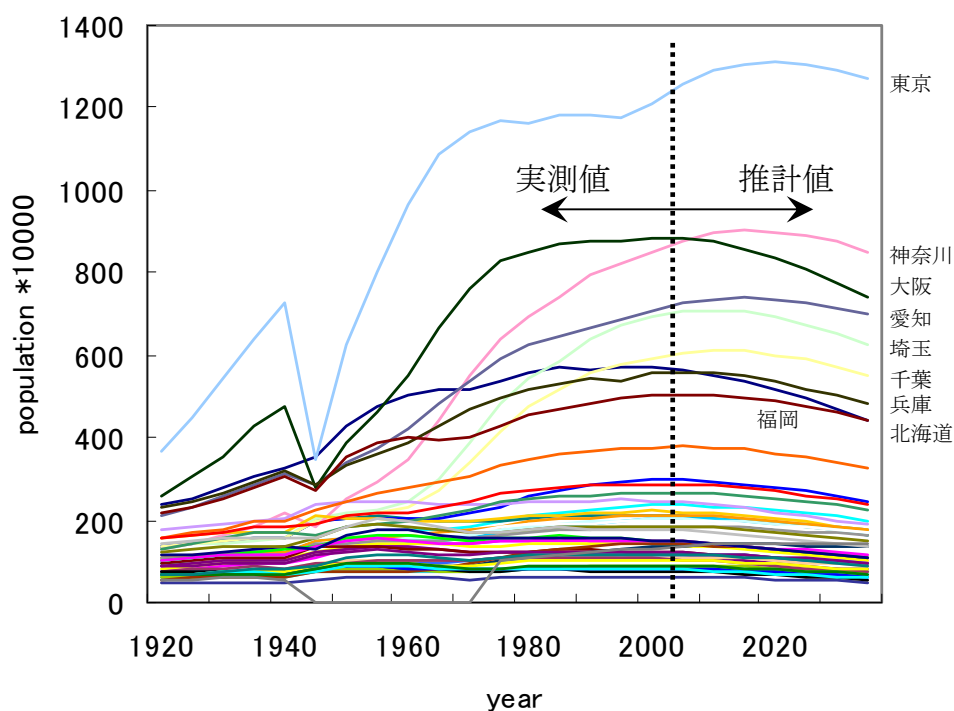


図 1.4 都道府県別人口の推移(1920～2035 年)¹³

実測値のほうに目を向けてみると、3 つのことが読み取れる。

結果 1) 1980 年頃から人口の変動が小さくなっている。

結果 2) 人口の多い都道府県は、人口の少ない県と比べて増減の変動も大き

¹³ 国立社会保障・人口問題研究所の都道府県別将来推計人口をもとに作成。
<http://www.ipss.go.jp/pp-fuken/j/fuken2007/t-page.asp> (2007.10.27 参照)

い.

結果 3) 戦後の経済復興, 高度経済成長を機に, 人口を大幅に増やした県と, それ以外の県に分かれる. 1950~70 年代, 東京, 大阪の人口が急増し, これに引っ張られる形で東京周辺の神奈川・埼玉・千葉, 大阪周辺の兵庫の人口が増加している.

結果 1 は, 都道府県の人口分布がある定常状態に近づいていっていることを予感させる. また, 結果 2, 3 から 2 つの問いを設定することができる.

問 1) なぜ, 人口が多い地域ほど変動も大きいのであろうか?

問 2) なぜ, ある地域が大きく発展し, 別の地域がしないのであろうか?

上述の 2 つの問いは, 発展した地域には, 発展するだけの特別な理由があるのではないかと, 暗に期待して発せられたものである. ある地域が発展し, 別の地域がしない, もしくは衰退する理由を知るためには, 数え切れない歴史的事実, 地理的要因と, 無数の社会的, 経済的影響力を分析しなければならないと考えられてきた. これまで, 人口の集中と分散に関して, そのメカニズムを経済学, 地理学, 社会学といったさまざまな視点から明らかにしようとする研究がなされ, 国際的な比較も行われている[Champion 1989]. このような研究の利点は, その時々, その地域ごとの社会情勢等を踏まえつつ, 具体的な要因について議論できることにある.

しかしながら, 多数の国ごとの個別の事情を勘案すれば, 関連する要因を網羅的に列挙して結論とせざるを得ないのが現状で, 「仮説は多いが検証は少ない[Greenwood 1985]」という状況である. 地域の成長には複雑な力が影響していることや, 人々の移動は個人的な理由で行われることから, 人口移動や地域人口の変動についての, 統一的な説明は難しいと考えられてきた.

一方, 相当の人口規模をもつ国における都市の**サイズ分布**¹⁴には, ある規則性が観察されることが知られている. それは, 都市人口と都市人口順位の関係が, べき乗

¹⁴ サイズ分布とは, 複数のもの(ここでは都市人口)の大きさについて, 「どれくらいの大きさのものが, どれくらいの個数あるか」ということ.

則に従うという事実である。これは、「ジップの法則」、あるいは「ランクサイズルール」ともいい、1913 年にアウエルバッハ[Auerbach 1913]が初めて指摘し、1949 年にジップ[Zipf 1949]によって見出された。都市の進化は明らかに混沌としているにもかかわらず、この規則性は 1950～1990 年のアメリカ[Glaeser 1995]、20 世紀のフランス[Eaton 1997]、19 世紀中盤の中国[Rozman 1990]、1911 年のインド[Zipf 1949]といった歴史も経済体制も異なる多くの国のデータで実証されてきた。

この順位とサイズの関係は、もともと自然界において多く観察されてきたものである。それが人口のサイズ分布に見られるということは、自然界と人間社会に、共通の法則性があるかもしれないことを意味する。人間社会に固有の複雑さの多くは、複雑な人間の心理学とはほとんど関係がなく、比較的単純な原理に従うのかもしれないのである[Buchanan 2005]。人間は、概して気まぐれであったり、強い意志を貫いたりと予測できないものではあるが、そのような人間が多く集まると、時には複雑で細かい部分の多くを無視した基本原理から、さまざまな現象が理解できる可能性がでてきたのである[Watts 2004]。

本研究の目的は、これまで複雑と思われてきた人口現象¹⁵の背後に規則性を見出し、その基本メカニズムを解明することである。これにより、上述の 2 つの問いに対する答えが自ずと導かれる。以下に、その枠組みを示す。

枠組み 1) 人口や人口移動量の**サイズ分布**に着目し、その背後にある規則性を明らかにする。これまで調べられてきたデータは、上位の数パーセントの都市についてのみであり、人口の小さい都市や町は、データの不足などを理由に無視されてきた。さらに、なぜこのような規則性が現れるのか、その本質的なメカニズムの解明については未解決のままである。本研究では、都道府県、市町村という 2 つの枠組みで、日本全体のデータを用いて検証を行う。

枠組み 2) 確率モデルによって、この規則性が再現できることを示す。統計力学的アプローチにより、メカニズムの解明を逆問題として捉え、現実の

¹⁵ 出生・死亡、人口移動などの現象をまとめて人口現象と呼ぶことにする。

規則性を再現できる変数を求める.

枠組み 3) 実データとモデルの推定値の比較から, モデルの妥当性, 社会的意味を明らかにする.

本研究は, 図 1.5 に示す流れで進める. 第 2 章で本研究のアプローチ方法を述べた後, 自然界, 社会に共通に見られる分布の規則性を発見し, この規則性に従う現象の一つである人口の分布を抽出する. そして, この人口の分布を再現できる確率モデルの提案および妥当性の検討を行う.

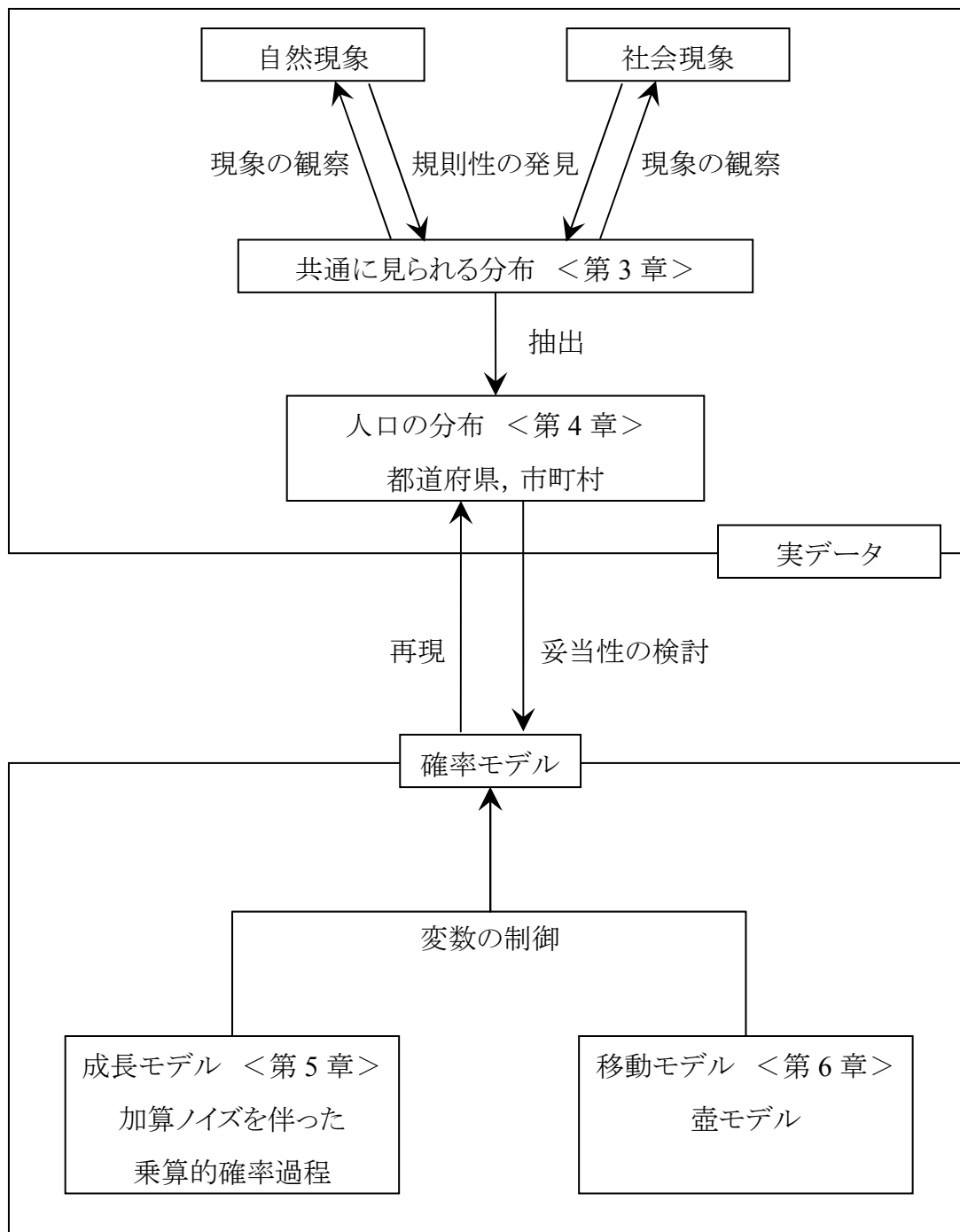


図 1.5 研究の流れ

1.2 本論文の構成

本論文は、序論と結論を含めて全体を 7 章で構成している。以下に、その構成を述べる。

第 2 章では、これまで行われてきた人口に関する研究の流れを俯瞰するとともに、その限界性について触れ、本研究の位置づけを明確にする。

第 3 章では、自然界や社会のいろいろな現象について、その大きさの分布 (サイズ分布) に着目すると、共通の規則性が見られることを示し、なぜこのような共通の規則性が見られるのかについて考察を行う。

第 4 章では、都道府県、市町村という 2 つの枠組みで人口のサイズ分布を観察したときに現れる興味深いパターン、法則性を提示する。

第 5 章では、人口の成長過程を、加算ノイズを伴った乗算的確率過程という確率モデルとして捉え、このモデルにおける変数を操作することによって、実際の人口サイズ分布を再現できることを示す。

第 6 章では、人口の移動を、壺モデルという確率モデルとして捉えることによって、加算ノイズを伴った乗算的確率過程と同じように、実際の人口サイズ分布を再現できることを示す。

第 7 章では、結論として本研究のまとめを述べる。併せて、本論文では扱えなかった、将来の研究課題に関しても述べる。

第 2 章

人口に関する研究

2.1 はじめに

これまで人口に関する研究は、さまざまな視点からのアプローチがなされ、多くの成果が蓄積されてきている。しかし、これらの成果を足し合わせても、全体像の理解が得られないというのが現実である。

本章では、これまで行われてきた人口に関する研究の流れを俯瞰するとともに、本研究の視点、学問的位置づけを明確にする。

2.2 人口に関する研究の俯瞰

2.2.1 人口問題の認識

人口の減少や増加といった変動によって、生活に不都合がもたらされるようなとき、我々は、解決しなければならない問題として、それを認識する。人口問題の認識と議論は古くからあった。古代のギリシャやローマの哲学者たち、プラトン (Platon 427～347 B.C.) やアリストテレス (Aristoteles 384～322 B.C.) も人口問題を論じている。プラトンは、人口に過不足が生じないように、必要に応じて調整することが国家の任務であると考え、結婚を国家当局の監督のもとにおいて、生まれた子供は国家が管理するという過激な人口政策を提唱したという[兼清 2002]。

マルサス (1766～1834) は、人口問題をめぐる議論を学問的なものに高めたという点で、近代人口学の先駆者であったとされている。1798 年にマルサスは、『人口論』を著した[Malthus 1978]。マルサスの人口問題認識は、貧困の根源が人口増加にあるというものであった。人口の増加傾向が、人類の進歩に不可欠なものであることを

認めた上で、それが社会を貧困¹⁶に陥れ、人類の幸福を妨げるという矛盾を問題にしたのである。この問題は、その後の目覚ましい経済発展のために、先進諸国では忘れ去られた感がある。しかし、大量生産・大量消費・大量廃棄をベースとした経済システムの発展により、エネルギー問題や環境問題など、新しい次元での生存資料の問題が生じている。

人口問題は、政治的、経済的、社会的、環境的視点といった、さまざまな方向からの認識が可能であり、また必要である。よって、必然的に人口に関する研究も多様なものとなる。

2.2.2 人口学の研究領域

人口に関する研究は、人口学(demography)と呼ばれる学問分野として発達してきた。この分野は、もともと体系的な学問として出発したのではなく、さまざまな既成の科学が、それぞれに人口現象と関係する側面を取り上げて、それぞれに異なる方法で考察してきた。そのため、関係する科学の数だけ研究方法も細分化され、経済学、地理学、生物学、社会学、政治学など、さまざまな分野で、人口の規模や構造に関する諸問題が研究され、さまざまな研究成果が蓄積されてきたのである。

アメリカの人口学者ハウザー(P. M. Hauser)は、人口学を人口分析(demographic analysis)と人口研究(population studies)の2部門に分けて、「人口分析は、人口の変化量と構成要素の研究に限定される。人口研究は、人口変数だけでなく、人口変動と他の社会的、経済的、政治的、生物的、遺伝的、地理的諸変数との関係にかかわるものである」と定義している[Hauser 1959]。さらに、「人口学は、狭義においては人口分析と同義と考えられ、広義においては人口分析と人口研究を含む」としている。このような人口学の定義に対して、いくつかの批判¹⁷もあるが、とにかく、広義の人口

¹⁶ 生存資料, 主として食料の不足

¹⁷ 南は、このように分けることに対して、「常識的な理解と矛盾する」、「人口研究の内容や性格、相互関係が明らかでない」などと批判している[南 1964]。

学の研究領域は、自然科学と社会科学の両方にまたがる広い領域にわたっており、その研究方法も多様であることに異論はない。人口学の概観は、図 2.1 のようになる。

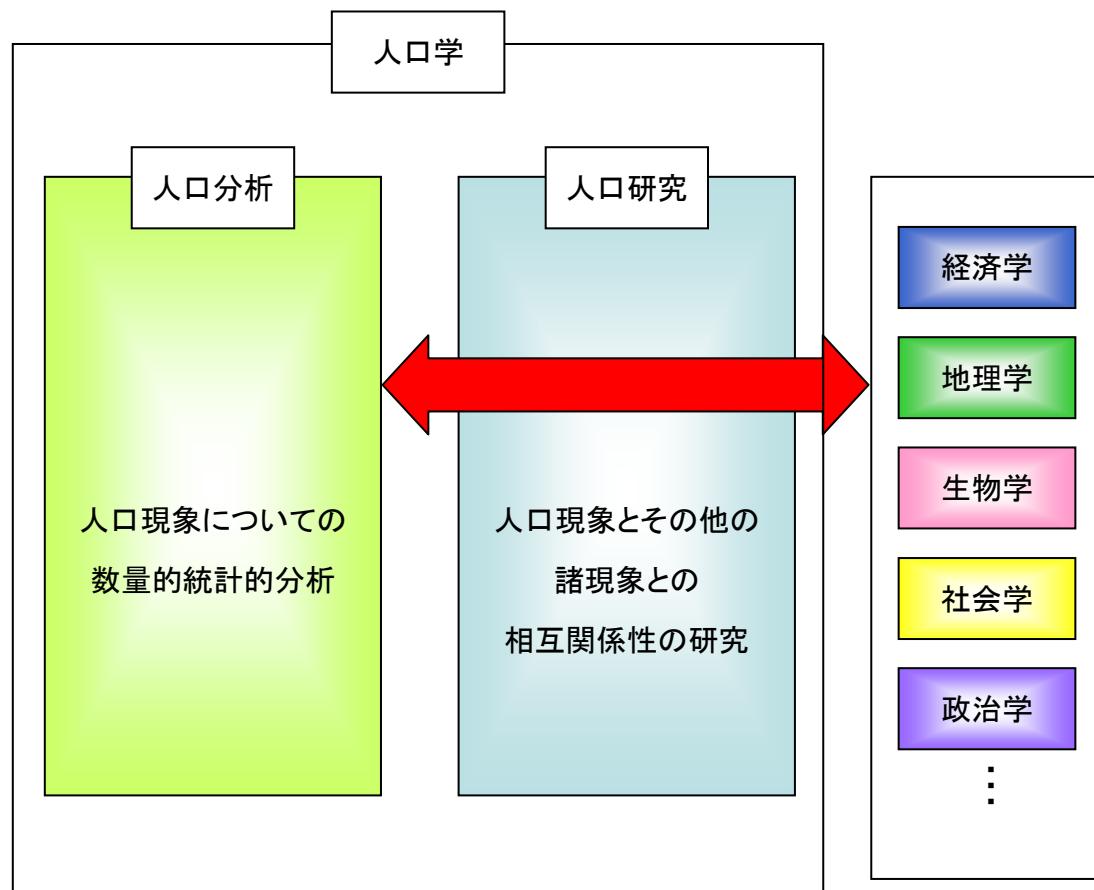


図 2.1 人口学の概観

自然科学は、自然現象についての法則を発見しようとする経験科学である。そのため、実証的な研究方法が用いられ、自然現象の量的な変化の関係が一般法則として把握される。これに対し、社会科学において一般法則を求めることが可能かどうかは問題である。

新カント派の哲学者リカード (Heinrich Rickert 1863～1936) は、自然科学と文化科学 (法学や経済学などの社会科学) の本質的な違いを強調し、文化科学の研究は、普遍的な法則を追及する自然科学的方法論とは異なっていて、歴史的方法によって

研究されなければならないと考えた。この場合、歴史とは、特殊性と個別性を有する一回生起¹⁸という概念である。自然現象と異なり、人間が価値を認める目的に従って行動した結果である社会現象の研究は、個別性を追及する歴史的研究が必要であるという意見はもつともである。しかし、社会現象を自然現象と比較して、一般法則を発見することは可能である。人口現象の研究は、自然科学か社会科学かという二者択一ではなく、自然科学と社会科学の両方を必要としているといえる。

2.2.3 人口移動のモデリング

これまで人口移動を説明するために、さまざまなアプローチでモデル化がなされてきている。伝統的なアプローチは、クロス・セクション法と呼ばれるもので、明示的には時間の次元を立てないで、異なる時点のデータに同一の分析法を適用し、結果の違いからその間の変化を検討するものである。この方法の代表的なものがグラビティーモデル(重力モデルとも呼ばれる)で、ニュートンの万有引力法則という物理法則を理論的背景としている。グラビティーモデルは、1940年代に Stewart によって人口移動の説明に用いられて[Stewart 1947]以来、旅客移動、情報通信量移動等の問題にも適用がなされ、その後の数学モデルによる社会事象のモデル分析あるいは意思決定などの分野における学問的、理論的そして実証的な分析の発展に大きな影響を及ぼしている。ニュートンの万有引力法則は、2つの物体間に働く引力の大きさ f がこれら2つの物体の質量 m_1 , m_2 の積に比例し、物体間の距離 d の2乗に反比例するというもので、次式のように表される。

$$f = G \frac{m_1 m_2}{d^2} \quad (2.1)$$

¹⁸ 人間に関わる事象は、時間的推移の中でただ一回だけ生起するものであり、歴史の中に普遍的な法則を探求しようとするのは間違いであるという考え。

ここで G は比例定数である. これを n 個の地域の集合 $N = \{1, 2, \dots, n\}$ における任意の 2 地点間の人の移動として一般化する. 地域 i から地域 j への移動量を f_{ij} , 地域 i の人口を p_i , 地域 j の人口を q_j , i と j 両区間の距離を d_{ij} とするとき, i から j への移動量 f_{ij} と移動総量 p_i と q_j の間に次の関係が成り立つと仮定する.

$$f_{ij} = K \frac{p_i^\mu q_j^\nu}{d_{ij}^\gamma} \quad i \in N, j \in N \quad (2.2)$$

ここで, K, μ, ν, γ は変数である. μ は出発地の人口の影響, ν は到着地の人口の影響, γ は出発地—到着地間距離の影響を表す. これらの変数の推定については, 式の両辺の対数を取り,

$$\log f_{ij} = \log K + \mu \log p_i + \nu \log q_j - \gamma \log d_{ij} \quad (2.3)$$

に基づいて $\{f_{ij}, p_i, q_j, d_{ij}\}$ に最小二乗法を適用する. モデルの適合度は, 通常は決定係数 R^2 によって示される¹⁹.

グラビティーモデルを実際の人口移動に適用した例として, 熊本県内 98 市町村の人口移動の研究[大山 1998]がある. この例では, 地域 i の事業所数を p_i , 地域 j の事業所数を q_j とおいて, 出発地, 到着地の経済規模・距離と人口移動の関係を調べ, ν の値が大きいことから, 到着地の経済的集積規模の大きいところへの人口移動傾向が強まっていると結論付けている. 1975 年, 1984 年, 1988 年の 3 時点で推定を行っているが, 決定係数 R^2 はどの時点でも 0.8 を越えており, グラビティーモデルの説明力は高いといえる.

では, 一地域で人口移動をうまく説明できるグラビティーモデルを全国規模の人口移動にも適用できるであろうか? 南北に細長い日本全国の人口移動にグラビティー

¹⁹ 決定係数: $R^2 = \frac{\text{回帰によって説明される変動}}{\text{全変動}}$ は, 0.0 から 1.0 の範囲を動き, 全

体として, 推定値が観測値に類似していればいるほど, すなわち適合が良好なほど 1.0 に近づき, 逆に推定値が観測値とずれていればいるほど 0.0 に近づく. 経験的には, 0.7 以上で適当とされている.

モデルを適用して計算した結果を、表 2.1 に示す。この表から、到着地の人口の影響 ν が大きいこと、出発地－到着地間距離の影響 γ は低く、ほとんど変化がないことがわかる。しかし、決定係数自体が低い値となり、説明力が低下している。また、特定の時間断面における回帰分析は、相互作用パターンの変動を、基本的にその時点だけの変数から説明することを目指すもので、変化の説明に直結しないという難点がある。

表 2.1 変数推定結果

	$\log K$	μ	ν	γ	決定係数 R^2
1960	-22.949	0.641	0.960	0.930	0.310
1965	-26.357	0.780	1.085	0.907	0.392
1970	-28.831	0.885	1.174	0.885	0.283
1975	-31.035	0.997	1.180	0.933	0.552
1980	-33.661	0.989	1.174	0.930	0.586
1985	-34.624	0.996	1.210	0.961	0.587
1990	-33.362	1.030	1.278	0.956	0.601
1995	-33.363	1.061	1.202	0.963	0.652
2000	-28.199	0.816	1.141	0.941	0.580

これに対し、ある時点における空間的相互作用を、それ以前の時点における空間的相互作用自体、あるいはそれに関連する社会経済的・人口学的諸変数から説明しようとするモデル、あるいは複数の時点の間の時間経過があつてはじめて得られる空間的相互作用の変化量を記述するモデルを、動態モデリングという。人口移動のモデルは、時間の次元が明示的に取り込まれた動態モデリングへと進展している。

2.3 還元論に基づくアプローチの限界

人口学の専門分化は、ますます進んでいる。そのため、多様で特殊的な発見を一つにまとめるバインダーとして役立つ理論が必要だといわれてきた。しかし、自然科学と社会科学の広い領域にわたる人口研究の成果を総合的にまとめることは容易ではない。

方法の多様性に起因するこの問題に対して、大淵は「対象からの接近」という構想を提示した[大淵 1987]。これは、人口変数を中心に、それと何らかの関連をもつ諸変数を集めて、一つの分析領域を構成し、一定の秩序のもとに体系付けようとする試みである。例えば、出生力を分析の対象として取り上げると、その水準はまず生物学的な妊孕力²⁰によって規定されるが、さらに経済的、社会的、文化的な規制を受けて弱められた結果、現実出生力になるという関係がある。したがって、出生力という一つの対象の分析に、さまざまな既存の科学を導入する必要がある。「方法的には、多数の異なる科学分野を導入するが、出生力を対象とすることによって、各科学の守備範囲はおのずから限定されるであろう。つまり、ある科学の説明力の限界で他の科学がバトンを引き継ぐという仕方で、出生力についてすべてが説明されうる」としている。

対象からの接近という大淵の方法は、特定の研究対象への多科学的接近であり、人口研究に必要なさまざまな科学の協力を重視するものである。例えば、人口と経済との関係を追及して、そこに経済人口学と称される分野が形成されている。しかし、経済変数だけでは説明しきれない部分が残っている。そこで別の分野、例えば社会人口学や地域人口学にバトンが渡され、その力が及ばなくなったところで、また次の研究分野にバトンが渡されるという仕方で、さまざまな専門的人口学が集まって、総合的な理解が得られるというものである。

これは建設的な見解であるが、各専門分野がその限界を認めて、次の分野の研究にバトンタッチするとき、それらの間に溝が残ってしまうという問題が生じている。各々の専門分野における成果や知見の間には矛盾や齟齬も多く、統一的な理解は得ら

²⁰ 妊娠する能力のこと。

れない状況である。これは、「群盲象を撫でる」²¹という言い回しに例えることができる。人口現象のような複雑な系²²について、それを構成する要素に分解し、それらの個別の要素を理解しても、元の複雑な現象全体の性質や振る舞いを理解することができないのである。すなわち要素還元論に基づくアプローチには、限界があることも事実である。

人口移動に関する研究[斎野 1987, Fray 1988, Illeris 1990, 酒井 1991, 矢野 1996, Vining 1998, Ishikawa 2001]でも、全体像の理解よりも個別の現象を理解しようとする細分化が進んでいる。人口移動に関する研究は、人口全体に関する分析から、性別・年齢階級別・職種別などの移動者の属性ごと[Boyle 1999, 早瀬 2002, 阿藤 2004], 進学・就職・結婚・引退などの移動理由ごと[Fagan 1993, Rogers 1997, Streib 2001, 田原 2007]に細分化した分析,あるいは移動経歴を重視した分析に重点が移っている。また、これらの現象についてのモデル化も、2.2.3 節で述べたように単純なものから複雑なものへ、静的なものから動的なものへと、より詳細な記述を行う方向に進んでいる[石川 2001]。これらの研究によって、人口移動に関するより詳しい事実が明らかになっている。その反面、それぞれの現象を足し合わせても全体の理解にはつながっていないのである。

2.4 本研究の戦略：統計力学的アプローチ

現在、日本では29秒に1人が生まれ、死亡している²³。また、時々刻々と人々は移

²¹ 元々の逸話は、以下のようなものである。6人の盲人が象を撫でて調べている。一人は足を触って、「象は柱のようだ」と言い、別の盲人は腹を触って、「壁のようだ」と言い、またある者は尻尾を触って…と、偏った見方しかできない人たちが大勢集まっても、物事の本質を見抜けずに不毛な議論に始終するのである。彼らは、少なくとも特定の部分に関しては的を射た理解を持っているのだが、全体をうまく説明できないのである。日本の人口についての全体像は、付録2, 3に示す。

²² 相互に影響を及ぼしあう要素から構成される、まとまりや仕組みの全体のこと。

²³ <http://www.mhlw.go.jp/toukei/saikin/hw/jinkou/suikai06/index.html> (2007.10.27 参照)

動している。これらの人々の出生、死亡、移動といった人口現象をモニタリングして、データを蓄積することは、原理的には可能であるが、現実的には不可能である。また、このようなデータを蓄積することができたとしても、それだけで人口現象について理解することはできない。これらのデータに対して、統計的な整理を行って、傾向や法則性を見出して、はじめて現象の理解につながるといえる。自治体の人口や転入・転出の全リストよりも、例えば1万から10万人の範囲の人口をもつ自治体はいくつあるか、10万から20万人の自治体はいくつあるか、というような巨視的な記述、すなわち**分布 (distribution)**が人口の集中と分散というダイナミクスをとらえるのに必要な情報といえる。このような考え方は、統計力学のとりアプローチと同じである[久保 2003]。

統計力学はミクロな世界の力学法則に基づいて、マクロな世界を記述する体系である。ミクロな世界とマクロな世界を結ぶのは、極めて非自明で困難な課題であり、今日でも未解決の点が多くある。それでも「平衡状態」と呼ばれる限定された状況については、ミクロな世界の法則がどのようにしてマクロな世界と対応するかについてのほぼ完全な一般論が得られている。それは、人類の科学の中でも最も大きく成功した基礎分野の一つであり、平衡統計力学あるいは単に統計力学と呼ばれている。統計力学は、マクロな系の平衡状態を扱う際の必須の道具であり、物理、工学、化学などの諸分野の基礎の一つになっている。

一般に、粒子の数が増えれば増えるほど、力学の問題を解くのは難しくなる。外力がなく、粒子どうしが相互作用を及ぼし合っている系に限定しても、二体問題は(中心力の一体問題に還元することで)場合によっては一般解が求められるが、三体問題は一般には解くことができない。つまり、粒子の数が三つ以上になれば、力学にもとづいて運動方程式を書くことはできても、その結果として生じる運動を閉じた数式で表現することはできないのである。このため、膨大な数の粒子からなる系の性質を、力学法則に基づいて議論することは、ほぼ不可能といえる。ところが、構成要素の個数が極めて大きくなることで、逆に、ある種の問題の扱いは簡単になるという逆転が起こる。より正確に言えば、マクロな系が平衡状態という特殊な状態にあるときには、力学の問題を完全に解かずにマクロな物理量のふるまいを正確に特徴づけることができるのである。さらに、この際、系の記述には、力学ではなく**確率論**という道具が有用なのである。マクロ系の平衡状態に注目し、これを説明し得る基本メカニズムを抽出すると

というのが、統計力学の基本的な戦略である。

気体にしろ、固体にしろ、磁性体にしろ、統計力学の対象となるマクロな系は、一般には、極めて複雑なミクロな構造をもっている。これらの系の(量子)力学的なミクロな詳細を完全に特徴づけるには、膨大な数のミクロな変数が必要だが、通常、それらの値を正確に知ることなどできない。人間社会もまた、非常に複雑で自由度の大きい物理系である。そして、個人、企業、地域などが、さまざまなレベルで複雑に相互作用しあっている。この意味で、ミクロな自由度を予測することも制御することも、ほぼ不可能といえる。本研究のアプローチを図 2.2 に示す。個々の要素を理解して、それらの成果を積み重ねていくことによって、全体像を理解しようとする要素還元論的アプローチでは、限界がある。これに対して、統計力学的アプローチでは、予測不能なミクロな自由度を一旦**確率的な変数**として扱い、マクロな世界(人口現象)の記述になげる。そして、この記述が可能などきの変数の意味を後から考察する。逆問題としてとらえるこのような試みは、一つの方法論として健全であると考ええる。つまり、本研究が目指すのは、様々な要因を細かく計算することではなく、系のミクロな詳細に依存しない普遍的なふるまい(分布)を探しだし、それらを的確に記述することである。

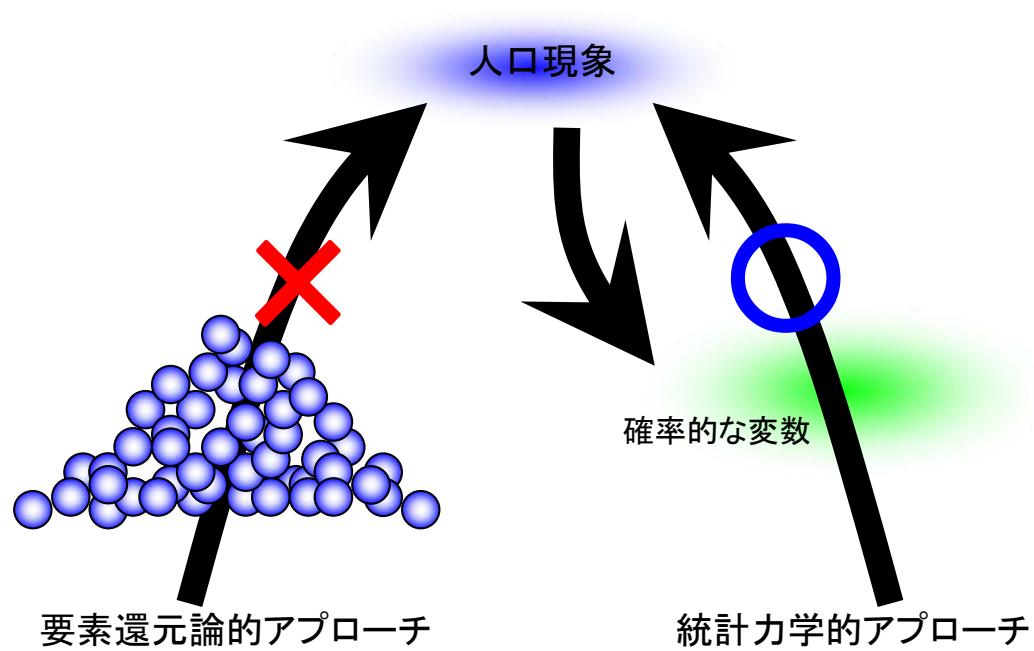


図 2.2 本研究のアプローチ

上記のようなアプローチで人口現象の解明を試みるが、その流れを図 2.3 に示す。現象の観察において、着目するのは人口および人口移動の**サイズ分布**である。このサイズ分布にパターン、規則性を見出す。そして、そのパターン形成の最も本質と思われる物理的メカニズムを抽出し、モデル化およびシミュレーションを行う。最後に、このシミュレーション結果と現実データの比較・検討を行う。

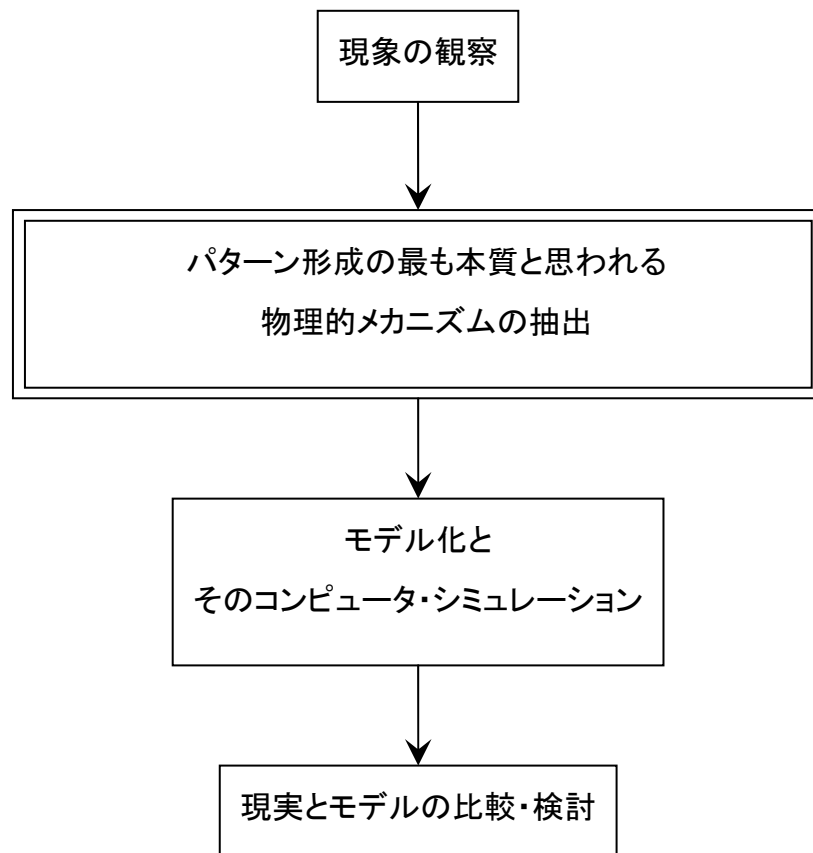


図 2.3 研究の流れ²⁴

2.5 まとめ

本章では、これまで行われてきた人口に関する研究の流れを俯瞰するとともに、本研究の視点、学問的位置づけを明確に述べた。これまで人口に関する研究は、社会的視点、経済学的視点、地理学的視点といったさまざまな視点から行われ、多くの成果が蓄積されてきている。しかし、これらの成果や知見を足し合わせていくことによ

²⁴ 『フラクタルの物理(Ⅰ)』[松下 2002], p95 を参考に作成。

って、全体の理解を深めていくという要素還元的なアプローチには、限界があることも事実である。統一的な理解が深まるどころか、成果や知見の間には矛盾や齟齬も多く、群盲象を撫でるがごとくである。これに対し本研究は、予測不能なミクロな自由度は一旦**確率的な変数**として扱い、マクロな世界の記述につなげ、変数の意味については後から考察する、逆問題として人口現象を捉える。本研究が目指すのは、様々な要因を細かく計算することではなく、系のミクロな詳細に依存しない普遍的なふるまい(分布)を探しだし、それらを的確に記述することである。

次の章で、自然界や社会において、この普遍的なふるまい(分布)が同じように見られることを示す。

第 3 章

自然界と社会に見られる共通の分布

3.1 はじめに

自然界や社会におけるさまざまな現象や事象の特徴を理解しようとする場合、統計的な整理を行った情報、すなわち**分布**が、最も基本的かつ重要な道具である。序論では、都市のサイズ分布がべき乗則に従うことを述べた。このべき乗則は自然現象の中に多く見出されてきたものである。しかし、人間の社会活動に関するデータの蓄積と公開が進むにつれ、社会現象の分布の中にも、同様の規則性が見られることが報告されるようになってきた。本章では、このような分布に見られる規則性についての基本事項を確認し、このような規則性をもつ他の現象の探索と分析を行う。

3.2 サイズ分布とは

3.2.1 確率分布

我々の身の周りの多くのものの大きさの分布は、正規分布になることが知られている。例えば、スーパーで売られているたくさんのリンゴの重さを量って、分布を見ると正規分布になっているだろう。実際に、大きさの分布が正規分布になっている例を示す。図3.1は、12歳男子の身長**の確率分布**²⁵である。この図は、横軸を身長 x 、縦軸を割合とするヒストグラムである。 $p(x)$ は確率密度関数²⁶ (probability density function) と呼ばれる。確率分布は、度数分布すなわちヒストグラムを割合にしたものである。この場合、割合をすべての区間について足すと1になる。この分布は、150cmあたりを中心として、その両側で急激に減少する「**釣鐘型**」の分布になっている。

²⁵文部科学省 平成17年度学校保健統計調査 身長の年齢別分布を基に作成。

http://www.mext.go.jp/b_menu/toukei/001/h17.htm (2007.8.20 参照)

²⁶ 確率密度, 分布関数, 確率分布と同義。

10cm や 300cm といった、極端に小さい、もしくは大きい身長の子は存在しない。身長差が 2 倍や 3 倍より大きくなることもほとんどない。このように、ある大きさを中心にして、その両側で急激に減少するような分布である正規分布は、釣鐘型分布の代表的なものである。正規分布の確率密度関数 $p(x)$ は、次のように書ける。

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \quad (3.1)$$

ここで、 μ は平均、 σ^2 は分散を表す。

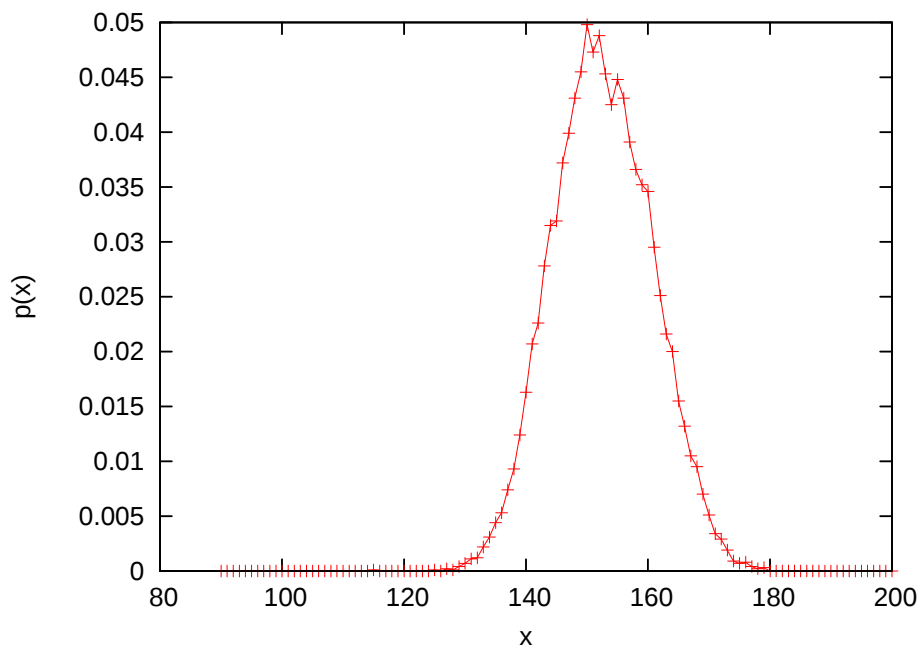


図 3.1 12 歳男子の身長確率分布

19 世紀頃は、すべての事象が正規分布で説明できると考えられていた。しかし、20 世紀以降、そういった考え方に修正が加えられている。生物集団の現象や社会現象において、右の裾野が長い分布が多く存在するという報告が相次いでいる。このような報告の詳細については 3.3 節で触れる。右の裾野が長いとは、身長の例でいうと、300cm、3000cm といった規格外の巨人が存在するような、左右非対称な分布のことである。その一例として、世界各国の人口について見てみる。人口は整数なので、離

散的に扱う必要がある。しかし、ここではサイズが十分大きく、実数と見なしてもよいと考え、連続的に扱う²⁷。図 3.2 は、2005 年における各国の人口を、1000 万人ごとに階級分けした確率密度関数をプロットしたものである。横軸は人口サイズ x ，縦軸は確率分布 $p(x)$ を表す。

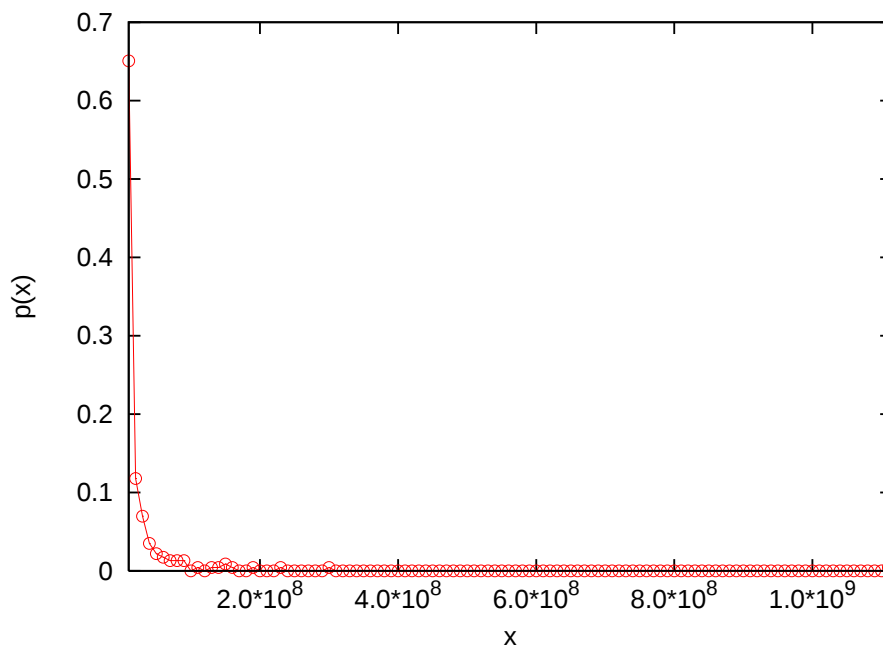


図 3.2 2005 年における各国人口の確率分布²⁸

2000 万人以下の国が数多くある一方で、中国やインドのように 10 億を超える国も存在する。世界各国の人口サイズ分布は、右に歪んだ (right-skewed) 分布になっている。このように右の長い裾野は、ファット・テール (fat-tail) もしくはヘビー・テール (heavy-tail) とも呼ばれるが、この図には、あるパターンが潜んでいる。

図 3.3 は、図 3.2 の横軸と縦軸それぞれの対数をとったものである。この図から、人

²⁷ 統計学では、母集団があり、それから抽出したものが標本データである。本研究では、母集団の分布関数を連続の形で求める。この確率分布から、整数だけを抽出するという条件のもとに得られた結果が、実際の現象として現れていると解釈する。

²⁸ 国連の人口データを基に作成。 [United Nations \(2006\) Population, Resources, Environment and Development: The 2005 Revision](#)

口が大きくなるにつれて、直線的に出現頻度が低くなることが見て取れる。これは、確率密度関数が、

$$\log p(x) = c - \alpha \log x \quad (3.2)$$

と、近似できることを示している。ここで α 、 c は、それぞれ傾き、切片の係数を表す。(3.2)式を書き換えると、

$$p(x) = e^c x^{-\alpha} \quad (3.3)$$

となる。このような形の分布をべき乗分布という。

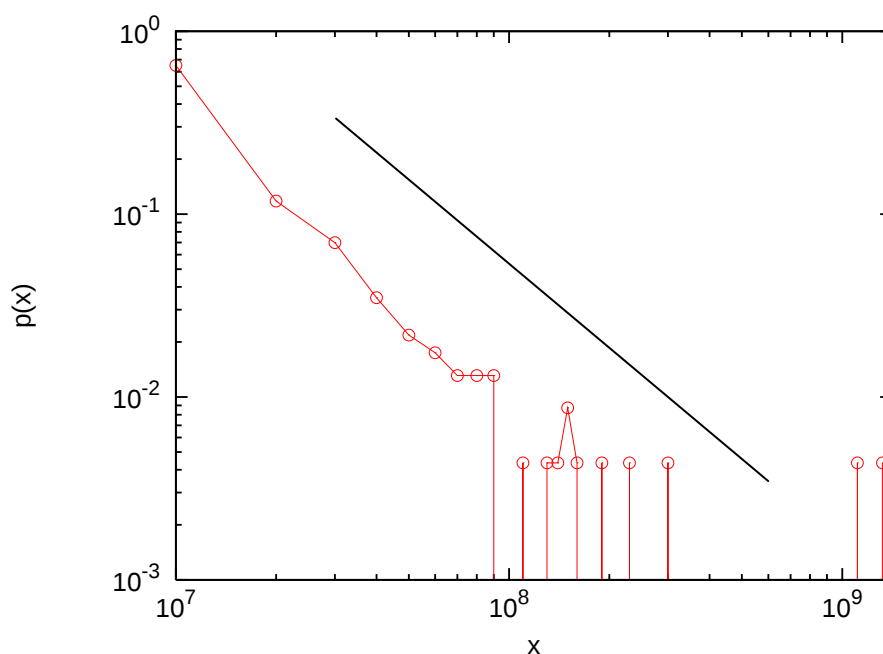


図 3.3 2005 年における各国人口の確率分布の両対数プロット

3.2.2 累積分布

ところで、図 3.3 は、 x の大きい区間 (10^8 以上) において確率分布が上下に振動していることが分かる。これは、 x の大きな領域に対しては、区間の大きさが相対的に小さくなるので、度数が 0 となる区間が次第に増えていくためである。そこで、こうした上下の振動を平滑化する目的で、確率密度関数の代わりに、**累積分布関数**²⁹ (cumulative distribution function) が用いられることが多い。本研究では、累積分布関数を $P_{>}(x) = [\text{変数の大きさが } x \text{ 以上となる対象の数が全体に占める割合}]$ と定義する。これは、 $P_{>}(x) = [x' > x \text{ となる } x' \text{ について } p(x') \text{ をすべて足したもの}]$ 、つまり、 $P_{>}(x)$ は $p(x)$ を積分したものである。このことから、

$$P_{>}(x) = \int_x^{\infty} p(x') dx' \quad (3.4)$$

となる。その確率密度関数は、

$$p(x) = -\frac{dP_{>}(x)}{dx} \quad (3.5)$$

のように、累積分布関数を微分することによって求まる。このような、確率密度関数と累積分布関数の関係から、

$$P_{>}(x) \propto x^{-\beta} \Leftrightarrow p(x) \propto x^{-\beta-1} \quad (3.6)$$

となる。(3.3)と(3.6)式から、 $\beta = \alpha - 1$ という関係になっている。 α 、 β はべき指数 (exponent of the power-law) と呼ばれるが、確率密度関数として表した場合と累積分布関数として表した場合では、値が 1 異なる。ネットワーク科学において、次数分布について議論する場合は確率密度関数のべき指数 α 、経済物理学において、企業サイズなどの分布について議論する場合は累積分布関数のべき指数 β が一般的に用いられる。本研究では、累積分布関数のべき指数 β を用いる。

²⁹ 累積分布、累積確率と同義

3.2.3 累積分布とランクサイズプロット

累積分布関数を実際のデータから求めるには、**ランクサイズプロット** (rank-size plot)を用いる。データが n 個ある場合、それらを大きなもの順に並び替えて $x_1, \dots, x_n$ とする。そうすると、順位が i 番目のデータは x_i となるが、 $x \geq x_i$ となるデータはちょうど i 個になる。よって、横軸にサイズ、縦軸に順位をデータ数 n で割った量をプロットすると、それがちょうど累積分布関数になる。同じことであるが、縦軸に順位そのものをプロットすると、累積分布関数を n 倍したもの $n \times P_{\geq}(x)$ が得られる。このように、累積分布のプロットには、分布関数を求めるのに必要な**区間の選択**や**情報の欠落**(区間内に度数としてまとめてしまう)がない。これがランクサイズプロットの利点である。サイズ分布についてのまとめを、図 3.4 に示す。

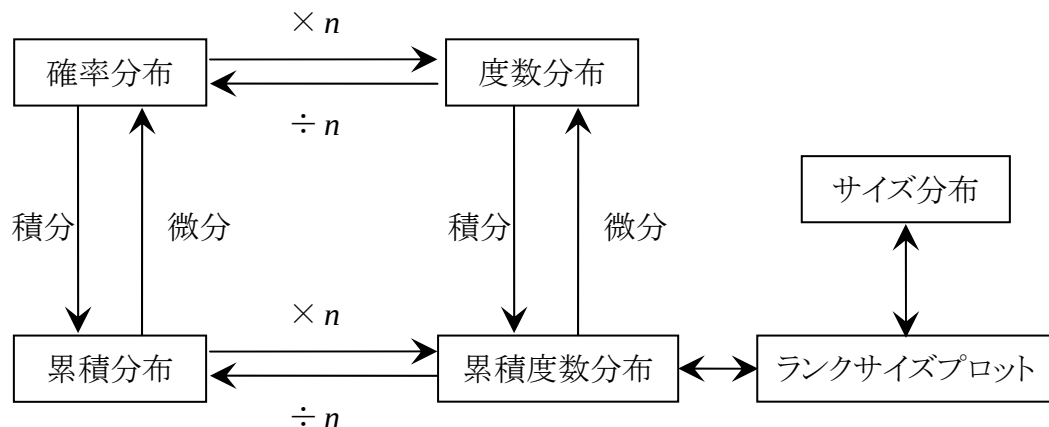


図 3.4 分布の関係図

本研究では、実際のデータの個数が分かるほうが、対象とする現象をイメージし易いという理由で、 $n \times P_{\geq}(x)$ すなわち順位を縦軸にもつランクサイズプロットを用いることにする。

順位を横軸、サイズを縦軸に表しているような研究もあるが、両軸が入れ替わっているだけである。また、データを小さなもの順に並べて、より小さなものが裾野の部分にくるようにプロットする場合もあるが、 $P_{\leq}(x) = [\text{変数の大きさが } x \text{ 以下となる対象の数}]$

が全体に占める割合]と定義すれば,

$$P_{<}(x) = 1 - P_{>}(x) \quad (3.7)$$

となるので, 同じものを別の見方で見ていることになる.

図 3.5 は, 2006 年における各国人口のランクサイズプロットである. 横軸は人口 (サイズ), 縦軸は順位を表している. この図から, サイズの大きい上位 (10^7 以上) の国の分布は, べき乗則にしたがっていることが分かる. これは, これまで数多く実証されてきたジップの法則, あるいはランクサイズルールに相当する.

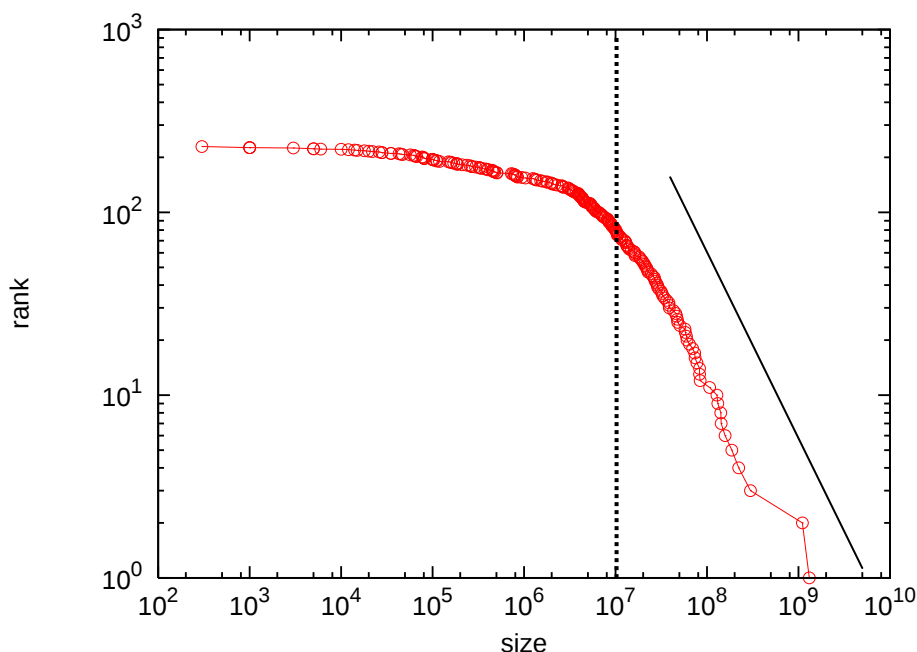


図 3.5 2005 年における各国人口のランクサイズプロット

ここで, べき乗分布は, 変数 x の大きい領域での漸近的な振る舞いが決まっているだけである, という点に注意しなければならない. 例えば, 指数関数的なカットオフ (減衰する因子) を含むようなガンマ分布や, 引き伸ばされた指数型分布 (stretched exponential distribution) など, 変数 x のある領域ではべき乗分布に従う. また, 長い裾野を持つ分布は, べき乗分布だけとは限らない. 変数 x の大きな領域では落ち込みがあるものの, 対数正規分布も裾野の長い分布の一つである. 対数正規分布とは,

変数 x の対数をとったものが正規分布に従うような分布のことをいう. x は 0 から ∞ の範囲で定義されていて, 確率密度関数 $p(x)$ は,

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left[-\frac{(\log(x/T))^2}{2\sigma^2}\right] \quad (3.8)$$

と与えられる. ここで, T は平均, σ は分散を表す. これは, 正規分布との関係から以下のように導くことができる.

確率変数 X が対数正規分布に従い, $X = e^Y$ とする. ここで Y は正規分布に従う. すると, 確率 $\Pr(X < k)$ は,

$$\begin{aligned} \Pr(X < k) &= \Pr(e^Y < k) \\ &= \Pr(Y < \log k) \\ &= \int_{-\infty}^{\log k} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(y-\mu)^2}{2\sigma^2}\right] dy \end{aligned} \quad (3.9)$$

ここで, $y = \log x$, $dy = \frac{1}{x} dx$ と変数変換を行い, さらに μ の代わりに, $\mu = \log T$ を使って,

$$\Pr(X < k) = \int_{-\infty}^k \frac{1}{\sqrt{2\pi}\sigma x} \exp\left[-\frac{(\log(x/T))^2}{2\sigma^2}\right] dx \quad (3.10)$$

となる. よって, 確率密度関数は(3.8)式のように与えられる.

累積分布関数 $P_{\geq}(x)$ は, (3.8)式を x から ∞ の範囲で積分して,

$$P_{\geq}(x) = \frac{1}{2} \left[1 - \operatorname{erf}\left(\frac{\log(x/T)}{\sqrt{2}\sigma}\right) \right] \quad (3.11)$$

と表せる. ここで $\text{erf}(x)$ は, 誤差関数で,

$$\text{erf}(x) \equiv (2/\sqrt{\pi}) \int_0^x \exp(-t^2) dt \quad (3.12)$$

と定義される. さらに順位 $N(x)$ は, データの総数 N_T をかけて,

$$N(x) = \frac{N_T}{2} \left[1 - \text{erf} \left(\frac{\log(x/T)}{\sqrt{2}\sigma} \right) \right] \quad (3.13)$$

となる.

3.3 これまでに報告された事例

これまで, べき乗分布であると報告されてきた分布の中には, 対数正規分布に従うとされるものもある. 収集データの質と量の違いによって, べき乗分布と対数正規分布は混同されてきたのである. いずれにしても, 裾野の長い分布は, 自然現象, 社会経済現象共に広く存在することが, 数多く報告されている. これらの一部を列挙すると, 表 3.1 のようになる.

表 3.1 裾野の長い分布をもつ現象

裾野の長い分布	自然現象	酵母菌のたんぱく質ネットワークにおけるたんぱく質の次数分布[Ito 2000] バクテリア (<i>E. coli.</i>) の代謝ネットワークにおける代謝体の次数分布 [Huss 2006] 最近の4半世紀における哺乳類の種の数[Smith 2003] 鳥の種別目撃頻度 http://www.mbr-pwrc.usgs.gov/bbs/ 魚の群れの大きさ[Bonabeau 1995] 血管の直径分布[Takayasu 1986] カリフォルニアの森林火災規模[Newman 2005] 太陽フレアの強度[Newman 2005] エアロゾル, 小惑星の大きさ[向井 2007] 地震の強度[Newman 2005]
	社会現象	小説『Moby Dick』における単語の出現頻度[Newman 2005] インターネット上での自律システムの次数分布[Holme 2007] 長距離通信会社 AT&T 社の顧客の受信数[Abello 1998] テロによる死者数[Clauset 2007] HTTP 通信のバイトサイズ[Willinger 1998] 停電によって影響を受けた客数[Newman 2005] 本の売上高[Hackett 1967] アメリカの都市人口[US Census 2000] 大学の email アドレス録[Newman 2002] 宗教の信者数 http://www.adherents.com/ アメリカにおける姓の頻度[US Census 1990] アメリカの富裕層の財産[Newman 2005] 論文の引用数[Redner 1998] 学者による論文発表数 http://www.ams.org/mathscinet/ ウェブサイトのハイパーリンク数[Broder 2000]

以上の多くの観測例で、べき指数 β はだいたい1から3の範囲の値をとる。 β がより小さくなると、分布がより広い範囲に広がることになる。あるいは、より大きなサイズを持つものが増えることになる。その意味で、べき乗則に従っている領域での不平等性を表しているといえる。これらの現象において、裾野の長い分布が生じる明確な理由については、まだ研究が進行中であるといえるが、多くのものが相互作用するような複雑系において、べき乗分布が頻繁に見られることは確かである。

3.4 データの再発掘

インターネット時代といわれて久しい。社会(特にネット上)には、あらゆる分野のデータがデジタル化され散在している。そこで、3.3 節で列挙した事例以外に、裾野の広い分布が見出せないか、インターネット等からデータの再発掘を行う。

データの再発掘に際しては、**人間の社会活動の総体**として得られるデータに着目する。実際に、入手したデータは、世界の上位112都市人口、世界各国の人口密度、世界の軍事費、戦争の死者数、戦争の期間、日本の姓のランキング、マクドナルドの店舗数、トヨタの販売台数、世界各国の電力消費量、世界各国の殺人事件の数、世界各国の囚人の数、世界各国の輸出入量である。これらのデータのサイズ分布を作成し、分布形に何らかの共通点があるのか、ないのか、あるとすればどのような点かを調べる。それぞれのデータのランクサイズプロットと、最小二乗法³⁰によって求めた近似式を図3.6および3.7に示す。また、どの分布関数で近似可能かによって、タイトルの横に、1本のべき乗分布の場合:○, べき乗分布と対数正規分布の組み合わせの場合[Mitzenmaher 2003]:△, 1本の対数正規分布の場合:□, べき乗分布2本の組み合わせの場合:*という記号をつけた。

³⁰ 測定で得られた数値の組を、適当なモデルから想定される関数を用いて近似するときに、想定する関数が測定値に対してよい近似となるように、残差の二乗和を最小とするような係数を決定する方法。

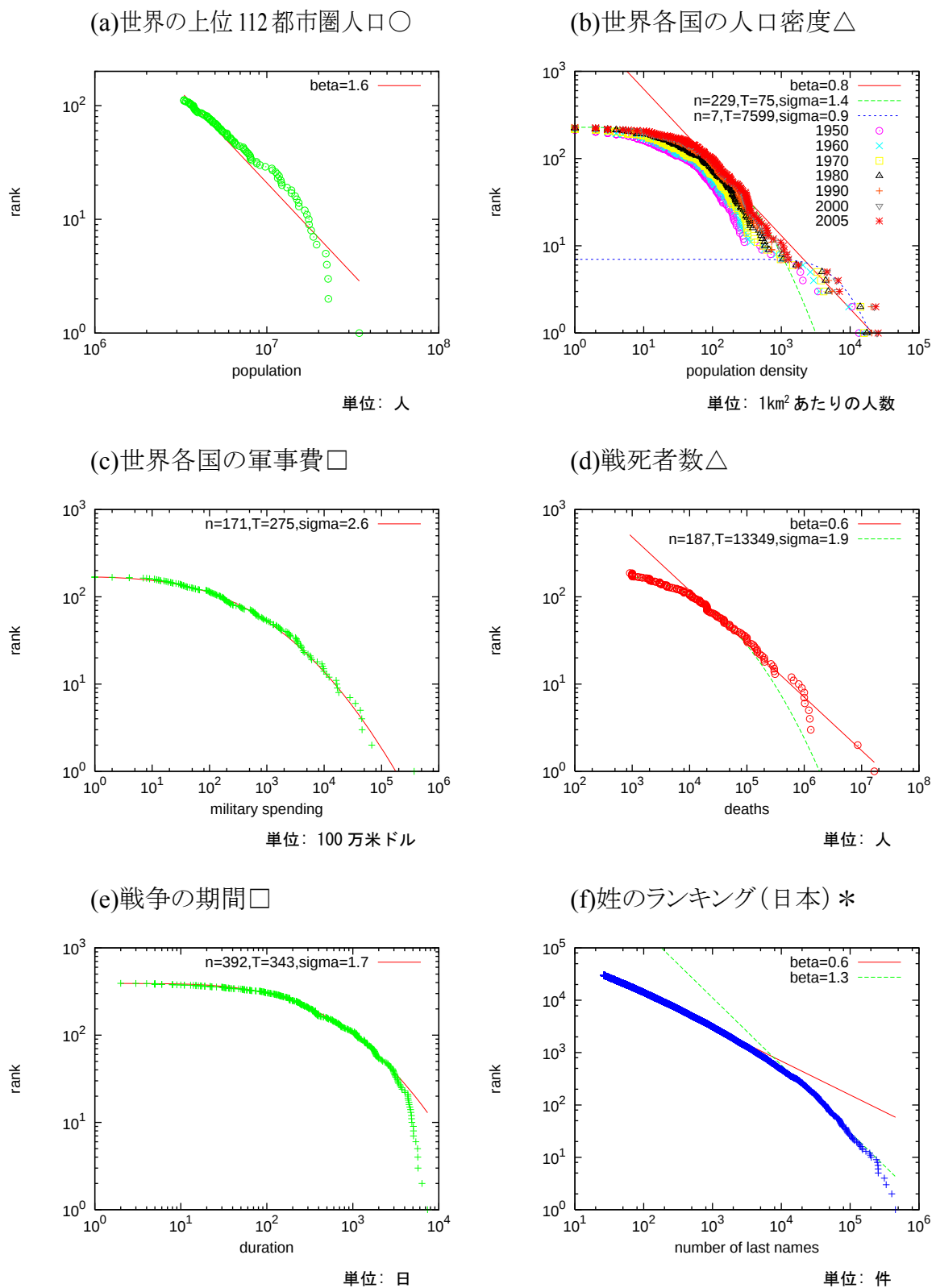
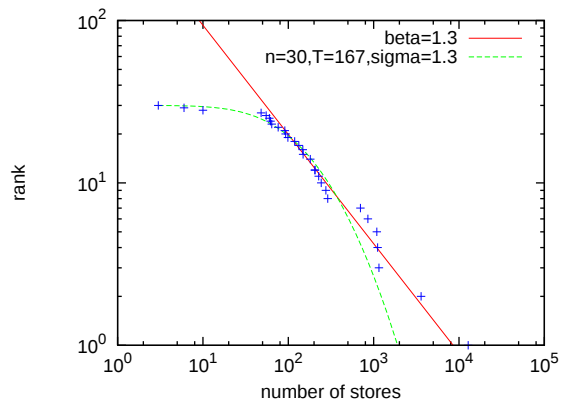


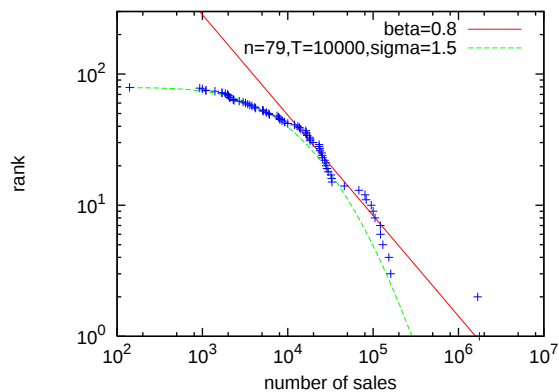
図 3.6 いろいろな現象のランクサイズプロット(その 1)

(g) マクドナルドの店舗数△



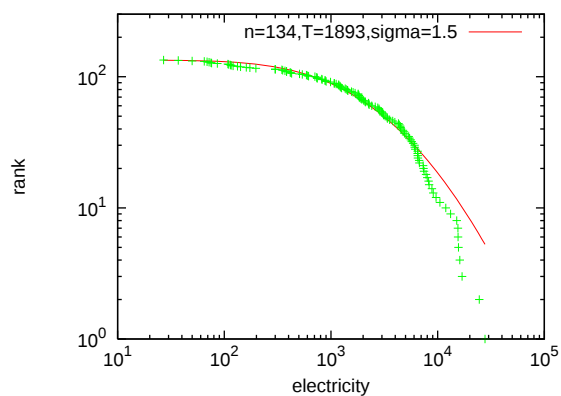
単位：店

(h) トヨタの販売台数△



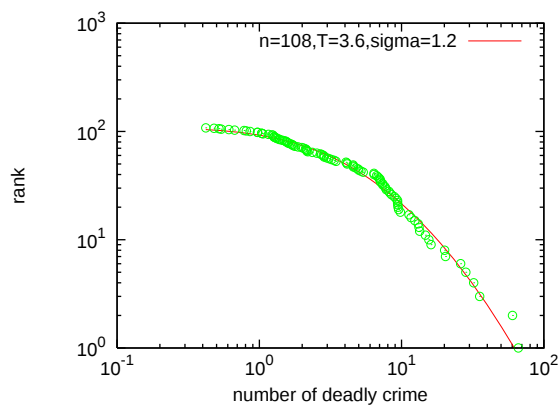
単位：台

(i) 世界各国の電力消費量□



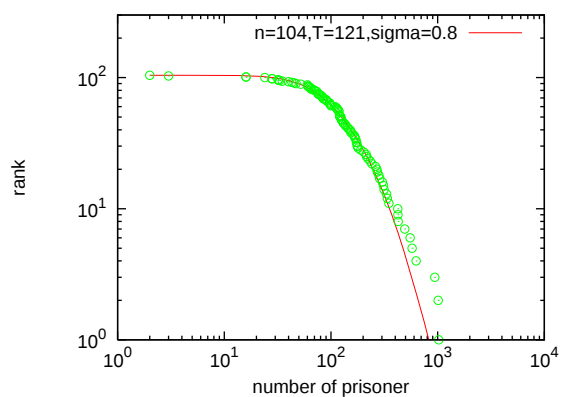
単位：kWh

(j) 世界各国の殺人事件の数□



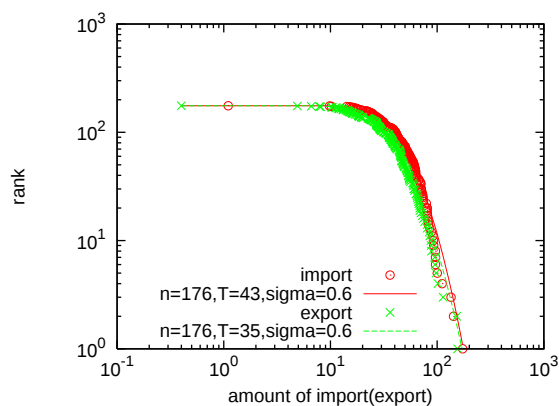
単位：人口 10 万人あたりの件数

(k) 世界各国の囚人の数□



単位：人口 10 万人あたりの人数（1990-2000 年の平均）

(l) 世界各国の輸出入量□



単位：GDP に占める割合

図 3.7 いろいろな現象のランクサイズプロット(その 2)

図 3.6 と 3.7 にプロットしたそれぞれのデータについて、以下に説明する。

(a) 世界の上位 112 都市圏人口³¹: ○

世界の都市圏人口とは、世界の主要都市と主要都市地域近くの郊外都市も含んだ都市圏の人口である。都市圏の定義は様々あり、それぞれの国の都市行政事情も異なるが、さまざまなウィキペディア言語版に掲載されている情報の中から最大の数値を用いた。データは 2005 年のものである。横軸は人口規模、縦軸は順位を表している。最大値付近でカットオフが見られるものの、 $\beta = 1.6$ のべき乗分布で、ほぼ近似できる。

(b) 世界各国の人口密度³²: △

世界各国の人口密度は、229 カ国各々の人口を面積で割って得られた値である。1950 年から 2005 年までの数値を用いた。横軸は人口密度、縦軸は順位を表している。上位区間は $\beta = 1.6$ のべき乗分布、下位区間は $T = 26$, $\sigma = 1.8$ の対数正規分布で近似できる。また、 10^2 から 10^3 の区間はべき乗分布でも対数正規分布でもよくフィットしている。

さらに別の見方として、2 本の対数正規分布の組み合わせと見ることもできる。この場合、 10^3 以下は $T = 26$, $\sigma = 1.8$ の対数正規分布、 10^3 以上は $T = 7599$, $\sigma = 1.9$ の対数正規分布で近似できる。

(c) 世界各国の軍事費³³: □

世界各国の軍事費は、国防費、防衛費、防衛関係費とも呼ばれ、軍隊(時に

³¹ 出典: フリー百科事典『ウィキペディア (Wikipedia)』

<http://ja.wikipedia.org/wiki/%E4%B8%96%E7%95%8C%E3%81%AE%E9%83%BD%E5%B8%82%E5%9C%8F%E4%BA%BA%E5%8F%A3%E3%81%AE%E9%A0%86%E4%BD%8D>

³² 出典: 国連のサイト <http://unstats.un.org/pop/dVariables/DRetrieval.aspx>
台湾のデータは、

<http://eng.dgbas.gov.tw/lp.asp?CtNode=2351&CtUnit=1072&BaseDSD=36>

³³ 出典: CIA, The World Factbook
<https://www.cia.gov/library/publications/download/>

は準軍事組織を含む)の維持運用の経費や軍事施設の生活環境の整備, 安全保障に関する会議の運営費などを含む軍事政策や軍事作戦に関する費用である. データは 2004 年のものである. 横軸は軍事費の規模, 縦軸は順位を表している. $T = 275$, $\sigma = 2.6$ の対数正規分布で近似できる.

(d) 戦死者数³⁴: $\triangle$

戦死者数とは, 戦争における作戦・戦闘において死亡した軍や市民防衛隊あるいはゲリラなどの構成員の数である. データは, 1816 年から 1997 年までに起こった戦争のうち, 死者数が公表されている 187 回分の戦争を対象としている. 横軸は戦死者数, 縦軸は順位を表している. 上位区間は $\beta = 0.6$ のべき乗分布, 下位区間は $T = 13349$, $\sigma = 1.9$ の対数正規分布で近似できる. また, 10^4 から 10^5 の区間はべき乗分布でも対数正規分布でもよくフィットしている.

(e) 戦争の期間³⁵: $\square$

戦争の期間とは, 戦闘が開始されてから終結するまでの日数である. データは, 1816 年から 1997 年までに起こった戦争のうち, 期間が公表されている 392 回分を対象としている. 横軸は戦争期間の長さ, 縦軸は順位を表している. $T = 343$, $\sigma = 1.7$ の対数正規分布で近似できる.

(f) 姓のランキング³⁶: $*$

姓は, 現在の日本においては, すべての日本国民が有する, 先祖から受け継がれてきた家の名称や夫婦を中心にした家族の名称を指す. 姓のランキングとは, 同じ姓 (例えば佐藤) が何件あるかをカウントし, 順位づけをしたものである. 横軸は苗字の件数, 縦軸は順位を表している. 10^4 以下は $\beta = 0.6$, 10^4 以上は $\beta = 1.3$ のべき乗分布で近似できる.

³⁴ 出典: CORRELATES OF WAR <http://www.correlatesofwar.org/>

³⁵ 出典: 脚注 9 と同じ

³⁶ 出典: 苗字館 <http://park14.wakwak.com/~myj/>

(g) マクドナルドの店舗数³⁷: Δ

マクドナルドは多国籍企業で、数多くの国にファーストフード店をチェーン展開している。このうちの 30 カ国について、店舗数をカウントし、順位づけをした。横軸は店舗数、縦軸は順位を表している。上位区間は $\beta = 1.3$ のべき乗分布、下位区間は $T = 167$, $\sigma = 1.3$ の対数正規分布で近似できる。また、 10^2 から 10^3 の区間はべき乗分布でも対数正規分布でもよくフィットしている。

(h) トヨタの販売台数³⁸: Δ

トヨタは、日本の自動車メーカー最大手で、全業種の中でも最大規模の企業である。世界生産台数及び販売台数が世界第 1 位で、多くの国で販売されている。このうち、79 カ国について、販売台数をカウントし、順位づけをした。データは 2003 年のものである。横軸は販売台数、縦軸は順位を表している。上位区間は $\beta = 0.8$ のべき乗分布、下位区間は $T = 10000$, $\sigma = 1.5$ の対数正規分布で近似できる。また、 10^4 から 10^5 の区間はべき乗分布でも対数正規分布でもよくフィットしている。

(i) 世界各国の電力消費量³⁹: $\square$

134 カ国の電力消費量について、順位づけをした。データは 2004 年のものである。横軸は 1 時間あたりの電力消費量、縦軸は順位を表している。 $T = 1893$, $\sigma = 1.5$ の対数正規分布で近似できる。

(j) 世界各国の殺人事件の数⁴⁰: $\square$

108 カ国の殺人事件の数について、順位づけをした。データは 2002 年、2003 年のものである。横軸は人口 10 万人あたりの殺人事件数、縦軸は順位を

³⁷ 出典: マクドナルドのサイト <http://www.mcdonalds.com/>

³⁸ 出典: トヨタのサイト http://www.toyota.co.jp/jp/about_toyota/gaikyo/index.html

³⁹ 出典: IEA (2004) Key World Energy Statistics 2004

⁴⁰ 出典: UNODC, United Nations Survey on Crime Trends and the Operations of Criminal Justice Systems (Fifth, Sixth and Seventh) (2002 年 12 月現在), Republic of China (2003) Statistical Yearbook of the Republic of China 2003.

表している. $T = 3.6$, $\sigma = 1.2$ の対数正規分布で近似できる.

(k) 世界各国の囚人の数¹⁵: □

104 カ国の囚人の数について, 順位づけをした. データは 1999 年から 2000 年までの平均をとったものである. 横軸は人口 10 万人あたりの囚人の数, 縦軸は順位を表している. $T = 121$, $\sigma = 0.8$ の対数正規分布で近似できる.

(l) 世界各国の輸出入量⁴¹: □

176 カ国の輸出入について, 順位づけをした. データは 2003 年のものである. 対内・外直接投資⁴²を含むデータなので GDP に占める割合が 100%を超える国もある. 横軸は輸出入量, 縦軸は順位を表している. 輸出は $T = 43$, $\sigma = 0.6$, 輸入は, $T = 35$, $\sigma = 0.6$ の対数正規分布で近似できる.

⁴¹ 出典: World Bank (2003) World Development Indicators 2003, Republic of China (2003) Statistical Yearbook of the Republic of China 2003.

⁴² 直接投資とは, 外国の企業に対して, 永続的な権益を取得する(経営を支配する)ことを目的に行われる投資である. 配当や金利といったインカム・ゲイン, 売却益といったキャピタル・ゲインを得ることを目的とした投資(間接投資)に対する概念である.

日本では, 日本企業による海外の企業に対する直接投資を対外直接投資, 海外の企業による日本企業に対する直接投資を対内直接投資という.

これらのデータについて、分布形によって整理したのが表 3.2 である。

表 3.2 データと分布形の関係

記号	データ名	分布形
(a)	世界の上位 112 都市圏人口	○
(b)	世界各国の人口密度	△
(c)	世界各国の軍事費	□
(d)	戦死者数	△
(e)	戦争の期間	□
(f)	姓のランキング(日本)	*
(g)	マクドナルドの店舗数	△
(h)	トヨタの販売台数	△
(i)	世界各国の電力消費量	□
(j)	世界各国の殺人事件の数	□
(k)	世界各国の囚人の数	□
(l)	世界各国の輸出入量	□

○：1 本のべき乗分布の場合，

△：べき乗分布と対数正規分布の組み合わせ，

□：1 本の対数正規分布，

*：べき乗分布 2 本の組み合わせ．

3.5 まとめ

人間の社会活動に関するデータは、対数正規分布とべき乗分布が混ざり合ったものになるのではないかと考えられる。順位でいうと、下位の大部分は対数正規分布に従い、上位の少数がべき乗分布に従う事象が多いようである。これは、図 3.8 のような

2つの解釈が可能である。

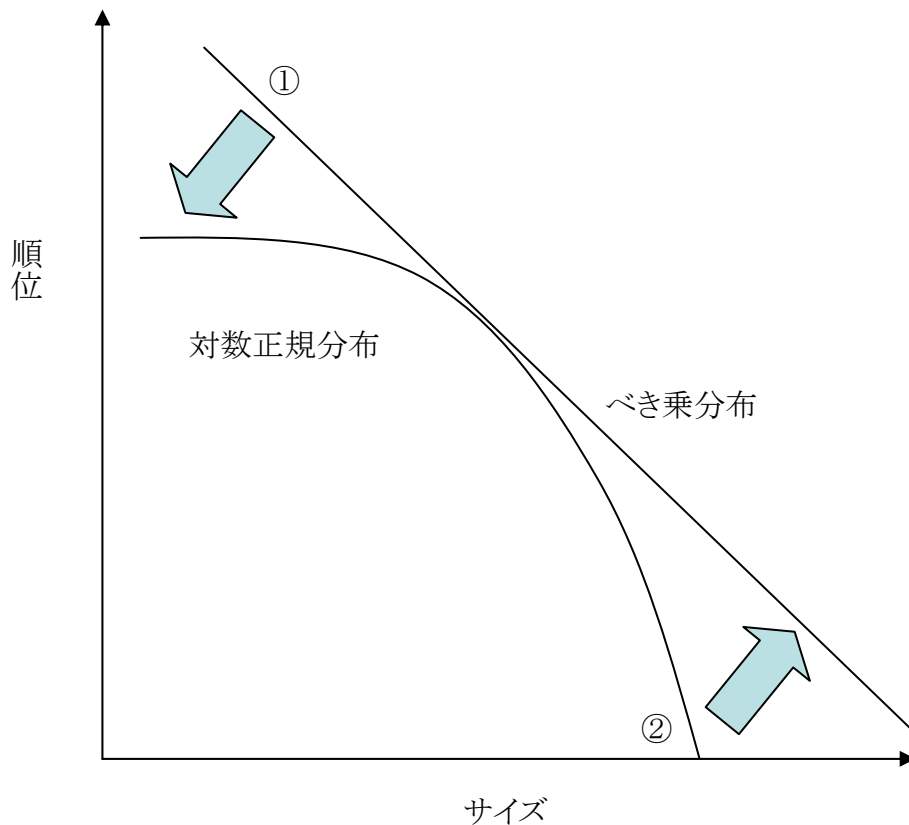


図 3.8 社会データについての 2 つの解釈

1つ目は、べき乗分布が基本にあり、下位の部分が①の矢印のように押さえつけられているという解釈である。これは、極端にサイズの小さいものがないように、不平等が拡大しすぎるのを抑える社会的圧力によるものと考えられる。2つ目は、対数正規分布が基本にあり、②の矢印のように上位の部分が持ち上げられているという解釈である。このような分布は、上位のものがより大きくなる開放的な場合に生じると考えられる。下位の部分で格差を小さくする平等主義、上位の部分で自由主義の力が働いているのではないかと考えられる。

人間の社会活動の総体は、釣鐘型の分布ではなく「The rich get richer」というように、格差が開くように進行する傾向があり、政策や世論は、これを抑制したり、促進したりする程度の役割を果たしているにすぎないという解釈ができる。

第 4 章

人口データの分析

4.1 はじめに

本章では、都道府県、市町村という 2 つの枠組みで、人口を観察したときに現れる興味深いパターン、規則性を提示する。これは、個々の自治体に注目するのではなく、全体をひとまとめにして整理することによってはじめて浮かび上がってくるパターンである。このパターン発見の道具は、第 3 章で導入した**サイズ分布**である。

具体的には、自治体のサイズ分布が、

見方 1) 上位 20%はべき乗分布、下位 80%は対数正規分布に従う。(都道府県、市町村)

見方 2) 1 本のべき乗分布に従う。(都道府県)

というように、2 つの見方ができることを示す。

4.2 人口統計データ

人口という標識によって作られた統計集団、すなわち人口統計集団の数量的記述が人口統計(population statistics)である。一般に、人口統計は人口の大きさ、構造、およびそれらの変動要因を把握することを目的として作成される。人口統計は、大別するとデータが一時点によって示される人口静態統計と、データが一定期間によって示される人口動態統計とに二分される。人口センサス⁴³のデータは前者に属し、人口の大きさや構造を表す。日本における人口センサスは、政府の実施する「国勢調査」である。国勢調査は 1920 年の第 1 回調査以来、戦後の混乱期を除いて 5 年ごとに

⁴³ 人口センサスは、一国あるいは一定の地域のすべての住民を対象とした統計調査で、直接的には特定の時点における人口の規模や構造を表す人口統計の作成を目的とするものである。

行われている。それに対し、出生、死亡、移動などのデータは後者に属し、人口の大きさや構造の変化を引き起こす人口現象を表す。いずれにしても、人口統計はそれらの人口静態や人口動態を、時間的、場所的な限定を付したうえで量的に表現したものである。

このような統計情報は、官庁のホームページ(総務省統計局⁴⁴や国立社会保障・人口問題研究所⁴⁵など)で、広く公開されている。本研究で使用する都道府県、市町村の人口は総務省統計局のホームページから、都道府県間人口移動量は住民基本台帳人口移動報告 総合報告書から入手した。

4.3 自治体のサイズ分布

現在、我々が“地域”として利用している枠組みは、数多く存在する。そうした枠組みを行政域の大きさに従って並べると、地方・州・都道府県⁴⁶・市町村⁴⁷となる。本研究では、都道府県と市町村という2つの地域レベルでの人口を扱う。これら2つは、レベルこそ異なるが、日本全体の人口をカバーしているという点は共通している。これまでいろいろな国において、都市の人口分布がジップの法則に従うという主張がされてきたが、人口の大きい都市だけではなく、一国全体での議論はほとんどない。アメリカの25359の町や村についての議論[Gabaix 1999, Eechout 2004]と、日本の市町村についての議論[Sasaki 2007]くらいである。この2つの議論では、サイズ分布がべき乗分布ではなく、むしろ対数正規分布に近い分布になる、もしくは対数正規分布とべき乗分布の組み合わせであるということが指摘されている。以下に、都道府県と市町村

⁴⁴ 総務省統計局：国勢調査結果を過去にさかのぼって見ることができる。

<http://www.stat.go.jp/>

⁴⁵ 国立社会保障・人口問題研究所：人口統計資料集、将来推計人口、少子化統計情報など、専門的な調査結果が公開されている。 <http://www.ipss.go.jp/>

⁴⁶ 都道府県は、日本における行政区分の一つで、「市町村を包括する広域の地方公共団体」である。

⁴⁷ 市町村は、地方公共団体である市、町、村の総称。包括的(広域的)地方公共団体である都道府県に対比される。

の人口サイズ分布について見ていく。

4.3.1 都道府県の人口サイズ分布

日本には 47 の都道府県がある。この 47 という数字は道州制⁴⁸が導入されるかどうかによって、今後どのように変わっていくかは分からないが、国勢調査開始以来、この枠組みでのデータが充実している。図 4.1 は 1920 年からの 2000 年まで 10 年おきの都道府県人口のランクサイズプロットである。横軸は人口 (サイズ)、縦軸は順位を表している。

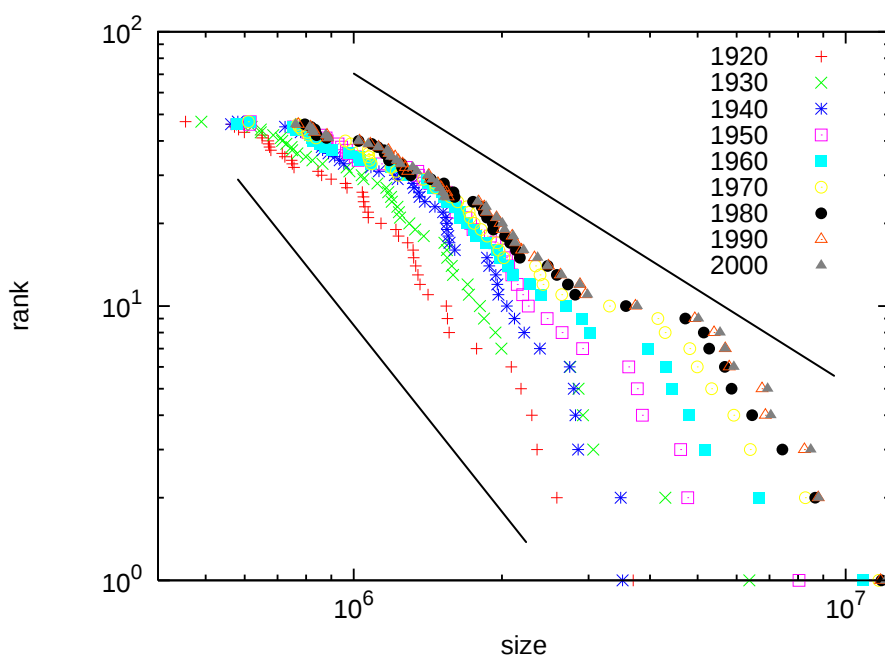


図 4.1 都道府県人口のランクサイズプロット (1920 年～2000 年)

1 本のべき乗分布で近似した場合、その傾きは、年を追うごとに緩やかになっていることが分かる (べき指数は、だんだん小さくなっている)。これは、大きい県と小さい県の格差が広がっていることに相当する。しかし、1980 年から 2000 年の分布は、ほぼ

⁴⁸ 都府県の内、幾つかを分割しその上で、都府県の広域連合の地方公共団体として道州を設置するという案など、様々な案がある。

重なっている。つまり、この間の人口分布に変化がほとんどないということである。そこで、1980年から2006年までの分布を詳しく見るために、5年おきのデータをプロットしたのが図4.2である。この図から1980年から2006年までは、ほぼ同じ分布とみなしてもよいほど一致していることが分かる。

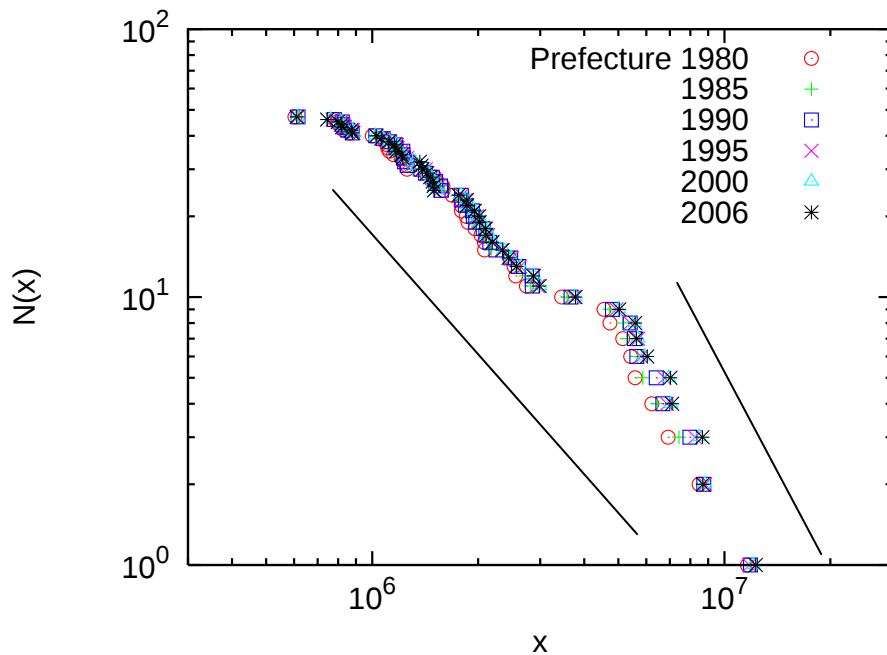


図 4.2 都道府県人口のランクサイズプロット(1980 年～2006 年)

1920 年から 2006 年までの人口のランクサイズプロットに対して、1 本のべき乗分布で近似した時のべき指数の推移を図 4.3 に示す。このべき指数は、全体的に見て右肩下がり傾向を示している。第 2 次世界大戦の時に、一旦大きくなる(断絶する)が、それ以降急激に小さくなり、本来の右肩下がり軌道に戻ったようにも見える。1980 年以降、1.2 あたりで横ばいになっていることから、このあたりの指数において、安定的な定常状態になっているのではないかと考えられる。ここでいう定常状態とは、図 1.4 のように順位の変動はあるが、図 4.2 のようにサイズ分布自体は安定している状態のことを指す。都道府県のランクサイズプロットは、1 本のべき乗分布以外に、2 本の対数正規分布で近似できる[Kobayashi 2006]という主張もあるが、縦軸の上限が 47 ということもあり、いかようにも解釈できるのが現実である。本研究では、次に示す市町村のランクサイズプロットと歩調を合わせるという意味も込めて、都道府県人口は、

見方 1) 上位 20%はべき乗分布, 下位 80%は対数正規分布に従う.

見方 2) 1 本のべき乗分布に従う.

という, 2 つの見方ができると解釈する.

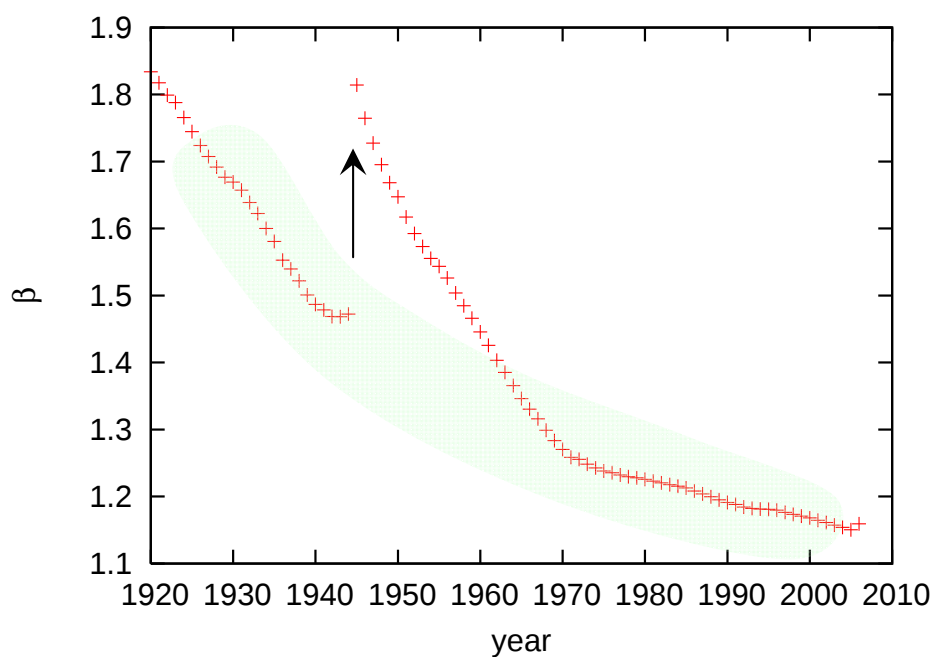


図 4.3 べき指数の推移

4.3.2 市町村の人口サイズ分布

表 4.1 は, 1980 年から 2006 年までの間の市町村数と, その減少数を示している. 1980 年から 2000 年までは, ほとんど数が変わらないが, 2000 年から 2006 年までの間に 1407 の市町村がなくなっている. これは, 「平成の大合併」によるものである.

表 4.1 市町村数の推移⁴⁹ (1980 年~2006 年)

Year	1980	1985	1990	1995	2000	2006
市町村数	3257	3253	3253	3234	3229	1822
減少数	-4	0	-19	-5	-1407	

年度によって数が異なるため、すべての市町村人口を時系列データで付き合わせることは難しい。しかし、サイズ分布で見るとは可能である。

図 4.4 は、市町村人口のランクサイズプロット(1980 年~2006 年)である。横軸は人口(サイズ)、縦軸は順位を表している。1980 年から 2000 年までは、市町村数も変わらず、ほとんど分布が重なっている。一方 2006 年は、市町村数が激減しているが、それにもかかわらず、傾きは他の年とほとんど変わらない。市町村のサイズ分布も、都道府県同様、定常分布になっていると考えられる。市町村のプロットの場合、データ数が多いので、べき上分布でうまくフィッティングできる区間と、対数正規分布でフィッティングできる区間が明確にわかる。市町村は合併のため年によって数が異なるため、各年毎にべき乗分布と対数正規分布でフィッティングを行った。この結果は付録 5 に付す。

⁴⁹ 『平成の市町村合併 早分かりマップ』より作成。

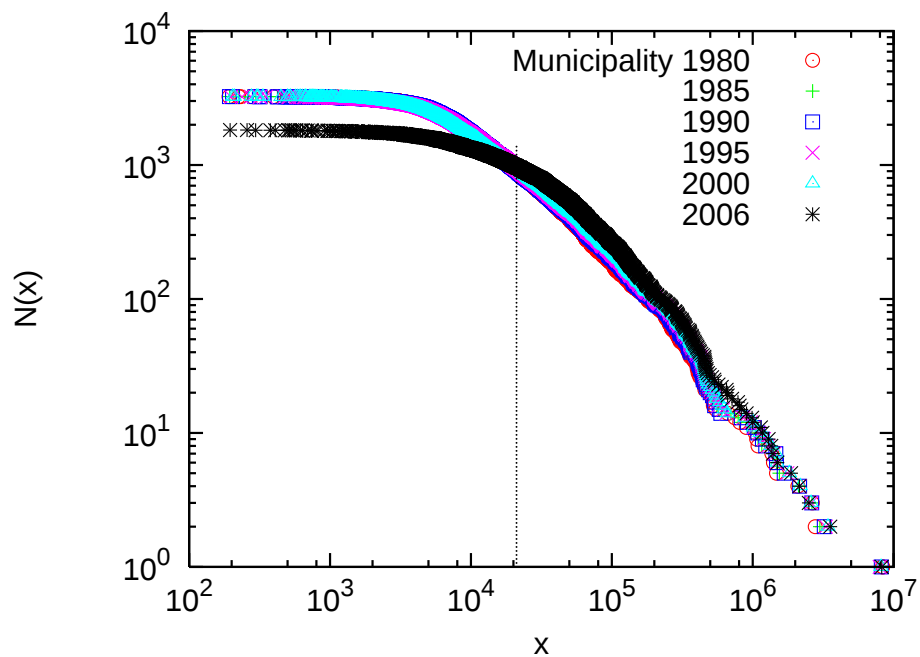


図 4.4 市町村人口のランクサイズプロット(1980 年～2006 年)

4.4 移動量のサイズ分布

図 4.5, 4.6 は, それぞれ 1965 年から 2003 年までの人口の転入・転出量についてのサイズ分布である.

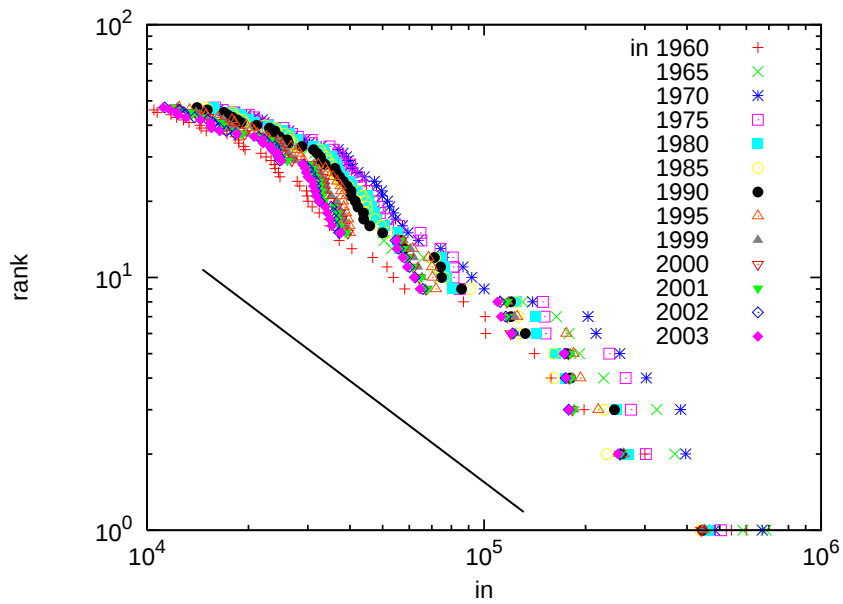


図 4.5 転入量のランクサイズプロット

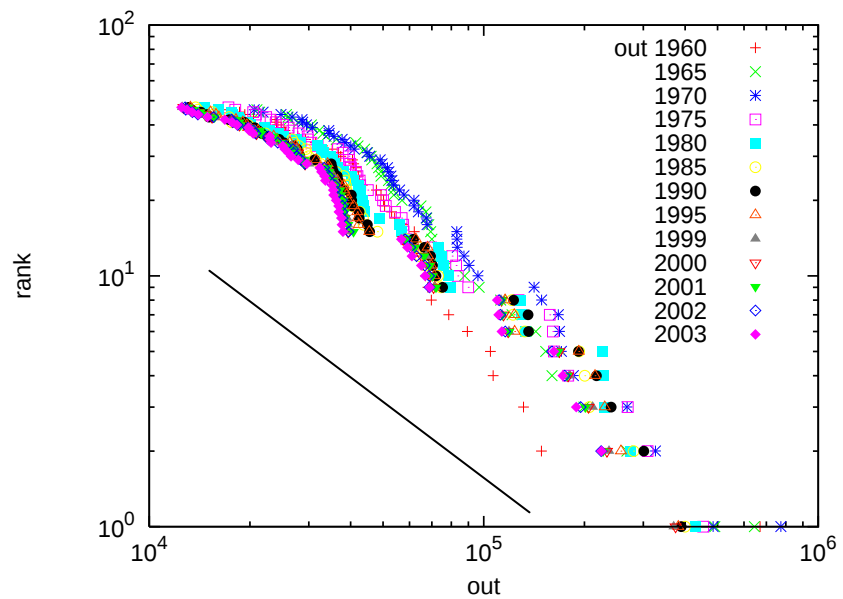


図 4.6 転出量のランクサイズプロット

移動量のサイズ分布についても、べき乗則が成り立つといってよい。また、この転入・転出量は、人口のサイズと関係がある。図 4.7 は、2005 年の転入、転出量を人口と移動量の関係を表している。人口の多い県ほど転入量、転出量共に大きくなる傾向が見て取れる。

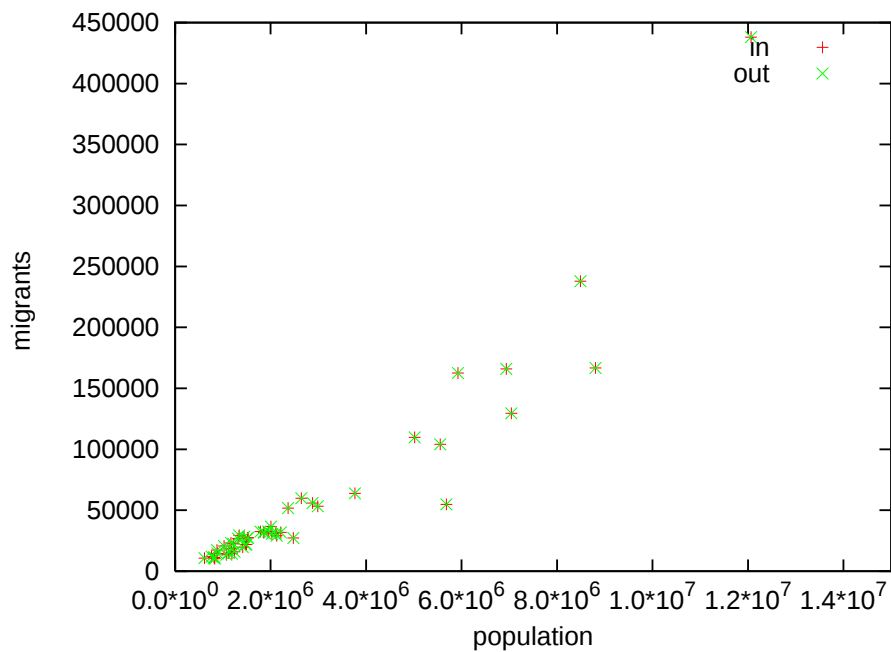


図 4.7 人口と転入・転出量の関係(2005 年)

4.5 まとめ

本章では、都道府県、市町村という 2 つの視点から人口のサイズ分布を分析した。特に、1980 年以降の都道府県、市町村の分布は、

予想 1) 上位 20%はべき乗分布、下位 80%は対数正規分布で近似できる(都道府県、市町村)。

予想 2) 1 本のべき乗分布で近似できる(都道府県)。

というように、都道府県人口、市町村人口ともに定常分布になっているという見方ができる。サイズ分布が定常状態になっているとは、ある自治体が発展し階層構造の中を上位に移動すると、他の自治体が衰退し、下位の階層に移動することを示している。

個々の自治体としては大きな変動があっても、全体としての分布は変わらないということである。また、移動量についても、べき乗則が成り立つ。これらの人口データの規則性を再現できる確率モデルを、次の第 5 章と第 6 章で提案する。

第 5 章

成長モデル

5.1 はじめに

第4章では、1980年以降の都道府県、市町村の人口サイズ分布について、

見方 1) 上位 20%はべき乗分布、下位 80%は対数正規分布に従う。(都道府県、市町村)

見方 2) 1本のべき乗分布に従う。(都道府県)

のように、2つの見方ができることを示した。このような分布は、第3章で示したさまざまな現象の分布と同じである。個々の自治体の人口サイズは時間とともに変化するが、自治体を集団としてみて、ある時点でのスナップショットである分布を見ると、規則性を見出すことができたのである。

分布を生み出しているのは、そのフローとストック量の時間的な変化、すなわちダイナミクスなので、その中身がどうなっているのかが分かれば、べき乗分布や対数正規分布の起源について理解が進む。そこで、本章では、自治体のストック量の時間的な変化について分析する。

まず、時系列の都道府県人口データを用いて、その変化にジブラ則というパターンがあることを示す。そして、自治体のサイズ分布を見方 1)のように捉えた場合、各自治体の成長メカニズムに簡単なノイズを加えるだけで、実データと極めて似通った分布を生成できることを示す。

5.2 成長率とジブラ則

5.2.1 自治体の成長

自治体人口サイズの成長を測るには、異なる時点でのサイズを比較すればよい。ある時点、 t_1 期と t_2 期における自治体サイズを、それぞれ x_1 , x_2 とする。そして、成長率 (growth rate) を

$$R = \frac{x_2}{x_1} \quad (5.1)$$

と定義する。 R は人口が何倍になったかを示す変数である。また、対数成長率 (logarithmic growth rate) を

$$r = \log_{10} R = \log_{10} x_2 - \log_{10} x_1 \quad (5.2)$$

とする。ここで $\log_{10}$ は、10 を底とする対数である。変化がなければ $R=1$ なので、対数成長率 r は 0、人口が増加すれば $R>1, r>0$ 、減少すれば $R<1, r<0$ となる。 $r=0.2, 0.4$ は成長率 R がそれぞれ約 1.6 倍、2.5 倍であり、 $r=-0.2, -0.4$ は R が約 0.6 倍、0.4 倍であることに対応する。図 5.1 は、都道府県人口の対数成長率の時系列推移である。縦軸は対数成長率、横軸は西暦年を表している。対数成長率は、戦前・戦中戦後の復興期と 1980 年あたりからの安定期という 2 つの時期に分けられる。

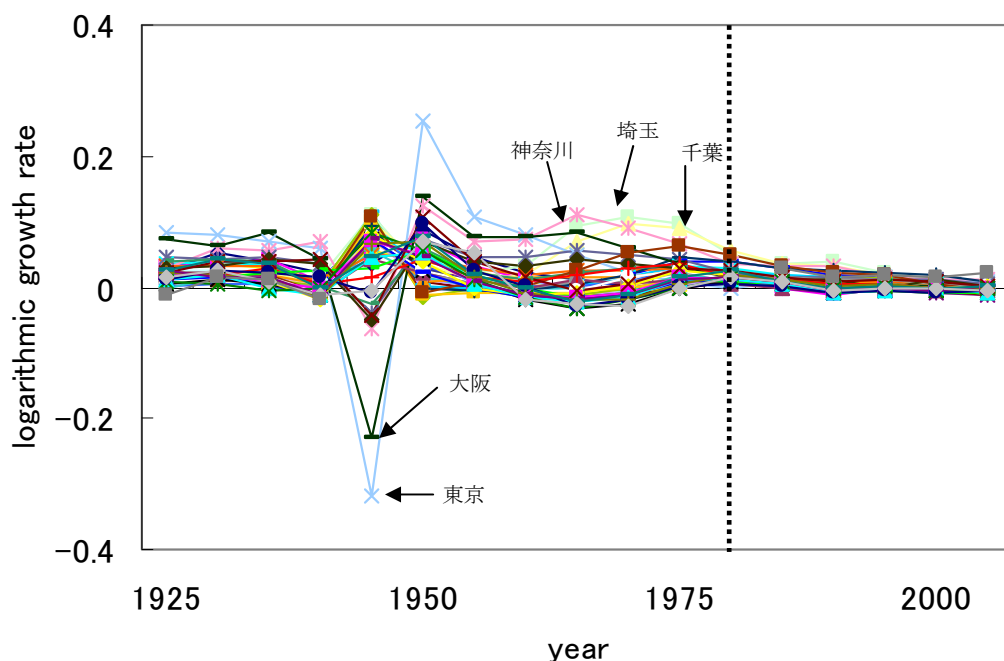


図 5.1 都道府県人口の対数成長率の推移

図 5.1 について、横軸を人口サイズに変更してプロットしたのが図 5.2 である。これらの点について、最小二乗法でフィッティングすると、傾きはほぼ 0 の直線で近似できる。これは一般的な直感とは違う結果といえる。というのは、サイズの異なる県 A と B では、何らかの成長率の分布が異なっていて、それゆえに A と B ではサイズが異なるという見方のほうが自然だと思われるからである。この図から、大きい県が大きい成長率、小さい県は小さい成長率を持っているとはいえないことが分かる。このことから、都道府県の人口は、そのサイズには統計的に依存しない成長率で変動している、という見方が成り立つ。この性質は、ジブラ則 (Gibrat's law) と呼ばれている[Gibrat 1931]。同様のことがアメリカの市町村[Echout 2004]、日本の市町村[Sasaki 2007]においても指摘されている。

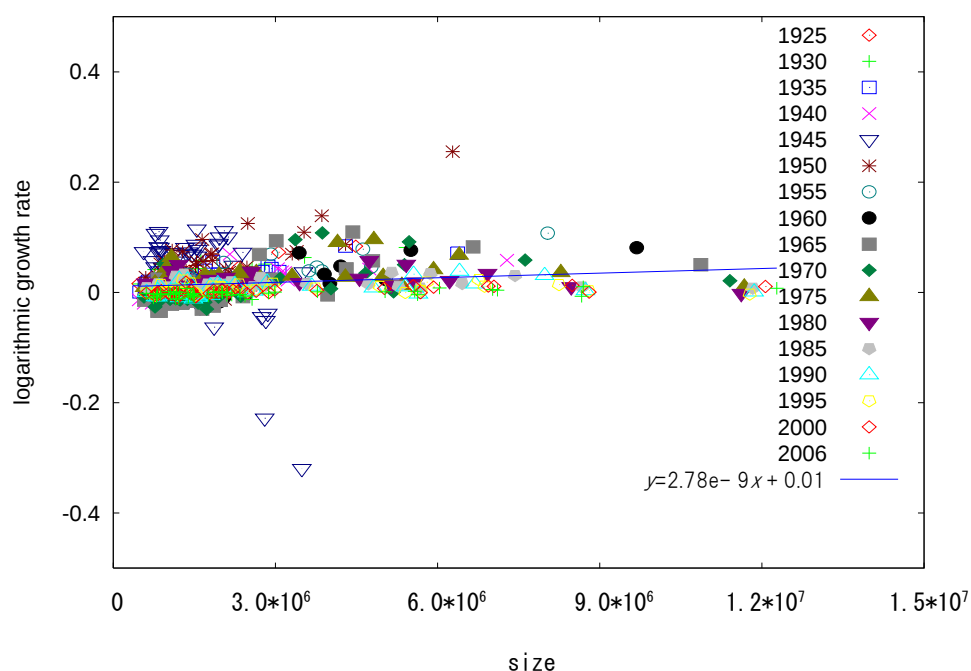


図 5.2 人口サイズと対数成長率の関係

フランスの経済学者ジブラ(R. Gibrat)は、その著書『経済における不平等性(Les inégalités économiques)』の中で、企業サイズのダイナミクスとしてある確率過程を1931年に提案した。これは、ある時点 t での企業サイズ x_t の成長率 R_t が、サイズとは無関係で、かつ時間的にも独立な確率変数であるような過程である。この過程は、比例効果の法則(law of proportionate effect)といい、企業サイズの変化の大きさが、ある確率変数とサイズのかげ算で決まるということを意味している。この単純な乗算的確率過程をジブラ過程(Gibrat's process)という。ここで、第1章で挙げた2つの問いの1に対する答えが導かれる。成長率は、サイズとは無関係であるが、たまたま大きなサイズの県にたまたま大きな成長率がかけられたら、そのサイズの変化も大きくなるのである。

5.2.2 ジブラ過程

5.2.1 節で、都道府県の人口サイズの成長率について、ジブラ則が成り立っていることを示した。ある年 t の自治体サイズを $x(t)$ とし、次の年のサイズを $x(t+1)$ とする。ジブラ則は、 $x(t+1)$ とは独立な成長率 $b(t)$ を用いて、

$$\frac{x(t+1)}{x(t)} = b(t) \quad (5.3)$$

と書ける。この両辺に $x(t)$ をかけると、

$$x(t+1) = b(t)x(t) \quad (5.4)$$

となる。つまり、 t 期のサイズ $x(t)$ に、それとは独立な $b(t)$ をかけたものが、 $t+1$ 期のサイズになるということである。このようなプロセスは、ジブラ過程と呼ばれている。このように、確率変数のかけ算として発展する過程は、一般に乗算的確率過程 (multiplicative stochastic process) と呼ばれている。ジブラ過程は、乗算的確率過程の中でも最もシンプルなものであるため、純乗算的確率過程と呼ばれる。このプロセスを 1 期さかのぼると、

$$x(t) = b(t-1)x(t-1) \quad (5.5)$$

と書ける。よって、

$$x(t+1) = b(t)x(t) = b(t)b(t-1)x(t-1) \quad (5.6)$$

となる。よって、このプロセスをどんどん溯ってみると、

$$x(t+1) = b(t)b(t-1)...b(1)b(0)x(0) \quad (5.7)$$

が得られる. ここで, 両辺の対数をとると,

$$\log x(t+1) = \log b(t) + \log b(t-1) + \dots + \log b(1) + \log b(0) + \log x(0) \quad (5.8)$$

となる. つまり, 右辺は足し算に変わる. ここで, $b(t)$ は $x(t)$ と無関係で, 異なる時刻の $b(t)$ 同士の間に関係がなければ, 中心極限定理 (central limit theorem) によって $\log x(t+1)$ は時間とともに正規分布に漸近する. よって, $x(t+1)$ は対数正規分布に漸近する. 対数正規分布はジブラにちなんでジブラ分布 (Gibrat's distribution) と呼ばれることもある.

以上のように, ジブラ過程によって, 対数正規分布が得られる. しかし, 本章では, 自治体の分布を, 下位 80% が対数正規分布で, 上位 20% がべき乗分布という 2 つの分布の組み合わせとして再現することを目的としている. では, どうすれば, 上位の部分にべき乗分布が得られるであろうか. この問題を解決する方法は, 物理学者によって多くの提案がなされ, 数値的にべき乗分布となることが示されている[青山 2007]. 例えば, リセットイベントを伴った乗算的確率過程 [Zanette 1997, Manrubia 1999, 相馬 2006] や, 下限を設けてそれより小さくなったものをその下限で反射させるというような, 反射壁を伴った乗算的確率過程 [Levy 1996], 加算ノイズを伴った乗算的確率過程 [Champernowne 1953, Kesten 1973, Sornette 1997, Takayasu 1997, Sato 2004] などがある.

リセットイベントを伴った乗算的確率過程は, ジブラ過程に対して, 確率 q で振り出しに戻り, 確率 $q-1$ で従うというルールを追加することによって, べき乗分布を再現できるようにしたものである.

$$x(t+1) = \begin{cases} x_0 & \text{確率 } q \\ b(t)x(t) & \text{確率 } 1-q \end{cases} \quad (5.9)$$

反射壁を伴った乗算的確率過程は、ジブラ過程に対して、 x_t がある値より小さくならないように反射壁を設けることによって、べき乗分布を再現できるようにしたものである。これは、(5.4) 式に対して、

$$\begin{cases} \langle \log b(t) \rangle < 0 \\ x_{\min} \leq x(t) \end{cases} \quad (5.10)$$

という制限を加えたものである。ここで $x_{\min}$ は反射壁の値である。

加算ノイズを伴った乗算的確率過程は、ジブラ過程に対して、成長率とは独立なノイズを加えることによって、べき乗分布を再現できるようにしたものである。これらの改良は、説明する現象によってさまざまな解釈が可能なので、適用範囲は広い[林 2007]。

5.3 モデル:加算ノイズを伴った乗算的確率過程

5.2 節で述べたように、べき乗分布を生成するための様々な乗算的確率過程が提案されているが、本研究では、人口現象との関係性を考慮して、加算ノイズを伴った乗算的確率過程に着目する。

ある地域の人口は出生によって増え、死亡によって減少する。これらの差し引きによってもたらされる人口の増減を自然動態という。また、人口は、転入によって増え、転出によって減少する。これらの差し引きを社会動態という。つまり、

$$\begin{aligned} \text{人口動態} &= (\text{出生数} - \text{死亡数}) + (\text{転入数} - \text{転出数}) \\ &= \text{自然動態} + \text{社会動態} \end{aligned}$$

が成り立つ。これは人口学的方程式と呼ばれ、地域の単位を国、県あるいは市町村といったスケールまで、いかなる大きさに設定しても成り立つ。ただし、どのような地域

を考えるかで、自然動態と社会動態それぞれの寄与には大きな違いがでてくる。日本の場合、島国であること、移民を積極的に受け入れてこなかったこと等から、最近まで人口増加はもっぱら自然動態によってもたらされてきた。これに対して、都道府県や市町村の場合、自然動態も重要であるが、他都道府県(他市町村)との間の転入、転出すなわち社会動態が大きな意味を持ってくる。この人口学的方程式は、加算ノイズを伴った乗算的確率過程と同じ式の形をしているのである。

加算ノイズを伴った乗算的確率過程は、(5.4)式に**加算ノイズ** $f(t)$ を加えて、以下のように書ける。

$$x_i(t+1) = b_i(t)x_i(t) + f_i(t) \quad i = 1, 2, \dots, N \quad (5.11)$$

ここで、添字は、各々の自治体を区別するために用いている。乗算ノイズ $b_i(t)$ は、正の値をとるノイズで、 $b_i(t) > 1$ のとき $x_i(t+1)$ は増加し、 $b_i(t) < 1$ のとき減少する。高安は、乗算ノイズ b と加算ノイズ f の分布形を、それぞれポアソン分布と左右対称なガウス分布と設定した[Takayasu 1997]が、最終的に得られるべき乗分布は、これらノイズの分布形に依存しないことが確認されている[Manrubia 1999]。

人口学的方程式との対応関係でみると、 $b_i(t)x_i(t)$ は自然動態、 $f_i(t)$ は社会動態に相当する。すなわち、乗算ノイズ b は成長率(正確には自然動態の成長率)、加算ノイズ f は転入・転出量による増減と解釈することができる。 b は割合であり次元をもたず、 f は人数であり次元をもつ変数である。実際問題として、 b と f は様々な地理的、社会的、歴史的、経済的要因を含んでおり、厳密な分布形を特定することは難しい。しかし、どのような分布形を設定しても、裾野の部分にべき乗分布が得られることが分かっているので、ここでは簡単のため、 b と f を「時間から独立した、ある範囲の一樣分布⁵⁰」と仮定する。このように、ある範囲の一樣分布と仮定する理由は 2 つある。1 つは、これら 2 つの変数が、無限ではなくある範囲に存在するであろうと予想されるため。もう 1 つは、一樣分布は、ある特定の分布形を仮定しないので、他の様々なケースにも適用しうると考えられるためである。

⁵⁰ 擬似乱数生成器にはメルセンヌ・ツイスタを使用。

5.3.1 初期状態と成長過程

初期状態として、各サイト⁵¹の人口を 100 人⁵²とした。この初期人口から乗算ノイズ $b_i(t)$ と加算ノイズ $f_i(t)$ の影響を受けながら、各々のサイトが成長していくのである。図 5.3 は、47 サイトの成長過程の 1 例である。横軸は時間、縦軸は人口サイズ x を表している。 2×10^6 時間あたりまでは、どんぐりの背比べのような感じで、大きいサイトと小さいサイトに大きな差はなく順位の変動も激しいが、 4×10^6 , 6×10^6 と時間が経つにつれて大きくなるものと小さいままのものとの差が、明確に現れてくる。一度大きくなる軌道に乗ったサイトは、その他の多数のサイトの数十倍にも成長する。しかし、長い時間軸で見ると、このように一度大きくなったサイトも、やがては衰退してその他大勢の中に紛れ込み、別のサイトが大きくなっていく。まさに栄枯盛衰である。このように、どのサイトが成長し衰退するかは予測できないが、全体のサイズ分布は、定常分布に収束していくのである。

⁵¹ サイトとは、自治体のことを指す。以後、都道府県と市町村を区別なく呼ぶときに用いる。

⁵² 100 人という人数に、特に意味はない。すべてのサイトの初期値を同じにすることで、平等なスタートラインから格差が生まれていく様子を観察することに意味がある。

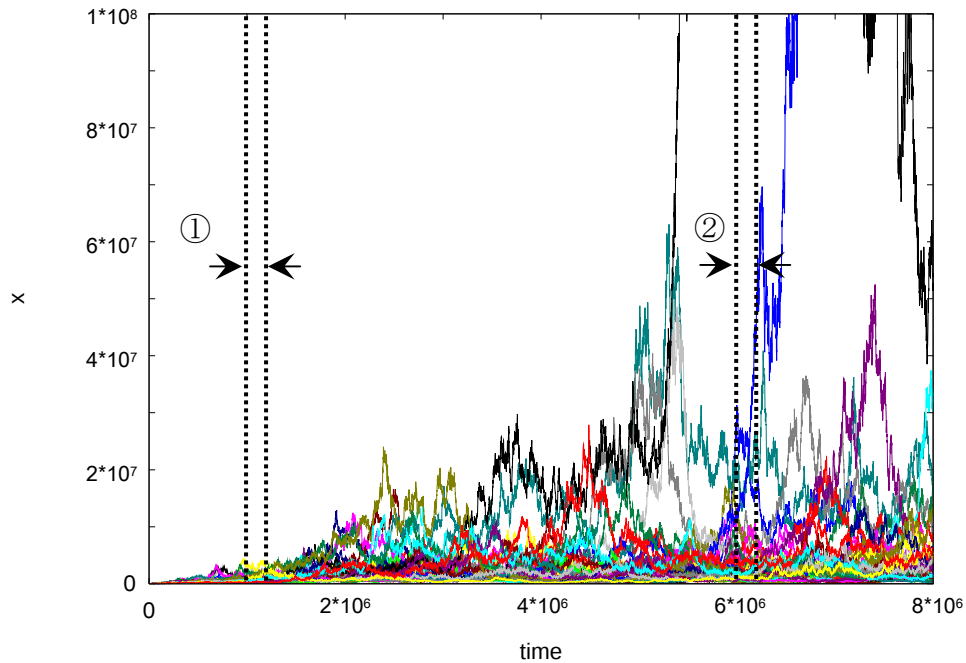


図 5.3 各サイトの人口成長過程

図 5.4 は、成長の初期段階での順位変動である。図 5.3 の①の時間範囲 (10^6 から 100 時間) での順位変動を可視化したものである。横軸は時間、縦軸は順位 (上から 1 位, 2 位, $\dots$ 47 位) を表している。この図から、時々刻々と順位変動がおきており、しかも変動規模が大きいことが見てとれる。これに対し、図 5.5 は、定常状態以後の順位変動である。図 5.4 と同様に、図 5.3 の②の時間範囲 (6×10^6 から 100 時間) での順位変動を可視化したものである。順位変動は少なく、その変動も順位が 1 つ入れ替わるくらいで、小さい変動である。さらに、上位の数サイトと下位の数サイトは、100 時間という範囲では順位の変動がほとんどない。このことから、定常状態に達してからの順位変動は、初期段階に比べて緩やかであるといえる。

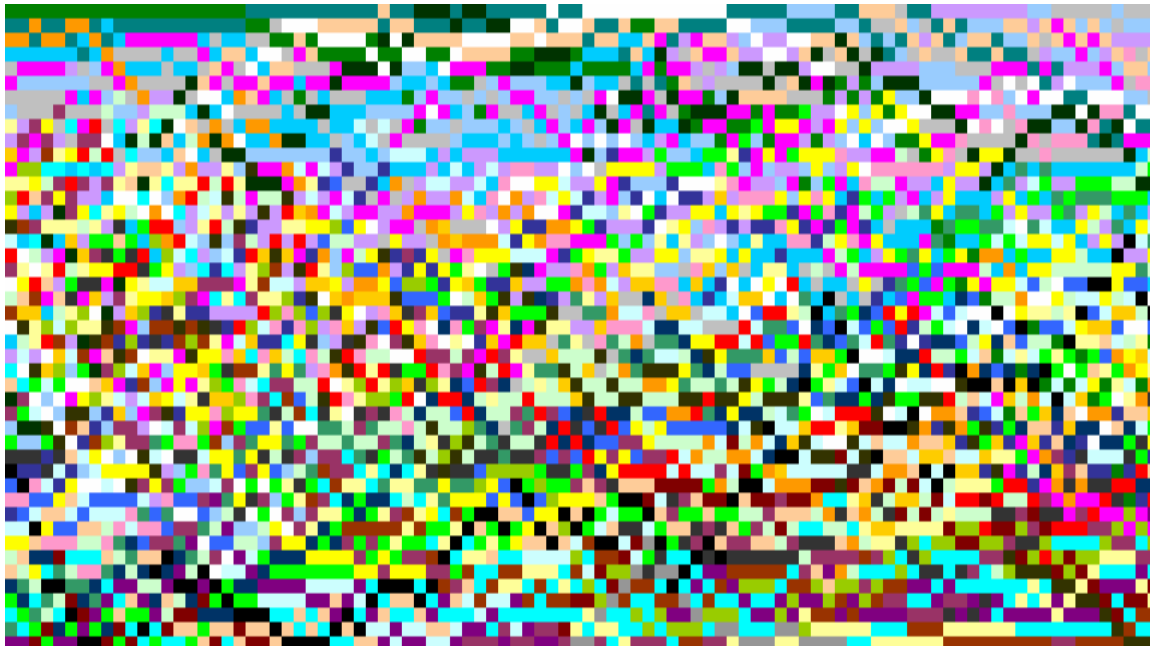


図 5.4 成長の初期段階での順位変動

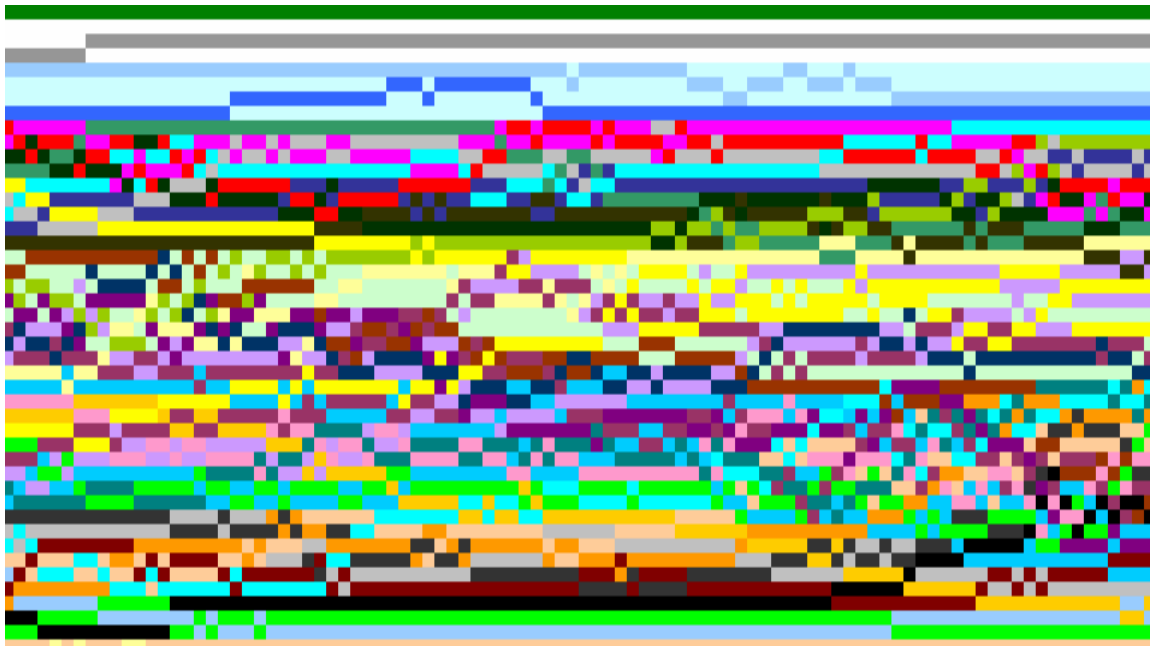


図 5.5 定常状態以後の順位変動

5.3.2 定常状態

加算ノイズを伴った乗算的確率過程では、指定したべき指数 β をもつべき乗分布に収束する. このことから、シミュレーションにおいて、ある時点での分布が定常分布になっているかどうかを判断することが可能となる. あらかじめ指定したべき指数 β と、ある時点での分布のべき指数を比較すればよい.

5.4 変数の制御性

加算ノイズを伴った乗算的確率過程には、べき指数 β 、乗算ノイズ $b(t)$ 、加算ノイズ $f(t)$ 、サイト数 N といった 4 つの変数がある. これら 4 つの変数のどれを変更しても、分布形を変えることができる. モデルと実データのフィッティングを行う前に、これらの変数の制御性について、既存研究によって指摘されている事実とシミュレーションを通じて新たに得た知見を、以下に述べる.

5.4.1 べき指数 β の操作による分布形の制御

5.3.2 節でも触れたが、加算ノイズを伴った乗算的確率過程では、指定したべき指数 β をもつべき乗分布を生成できることが数値的に示されている[Takayasu 1997]. 言い換えれば、べき指数 β を変更することによって、分布の形そのものを制御することができるのである. このような制御は、べき指数 β と乗算ノイズの範囲を決定するためのパラメータ b の関係が、期待値、

$$\langle b^\beta \rangle = 1 \tag{5.12}$$

となるように設定することによって可能となる. ここで, 確率変数 $B = b^\beta$ とする. b の確率分布を P とすると B の期待値は,

$$\langle B \rangle = \int_{-\infty}^{\infty} b^\beta dP(b) = \int_{-\infty}^{\infty} b^\beta p(b) db \quad (5.13)$$

で与えられる. ただし, $p(\cdot)$ は b の確率密度関数とする. $p(b)$ は, さまざまな分布が考えられるが, ここでは簡単のため $[b_{\min}, b_{\max}]$ の一様分布を仮定する. 確率の規格化により,

$$\int p(b) db = p(b) \int_{b_{\min}}^{b_{\max}} db = p(b)(b_{\max} - b_{\min}) = 1 \quad (5.14)$$

よって,

$$p(b) = \frac{1}{b_{\max} - b_{\min}} \quad (5.15)$$

となる. このとき(5.13)式は,

$$\begin{aligned} \langle B \rangle &= \int_{b_{\min}}^{b_{\max}} b^\beta p(b) db = p(b) \int_{b_{\min}}^{b_{\max}} b^\beta db \\ &= \frac{1}{b_{\max} - b_{\min}} \int_{b_{\min}}^{b_{\max}} b^\beta db = 1 \end{aligned} \quad (5.16)$$

よって,

$$\int_{b_{\min}}^{b_{\max}} b^{\beta} db = b_{\max} - b_{\min} \quad (5.17)$$

となる. 図 5.7 は, b の範囲を決める際の方法を示したものである. (5.17)式から, ①の範囲と②の範囲の面積は等しい. つまり, 求めたい分布の傾きであるべき指数 β と b の上限 $b_{\max}$ を決めれば $b_{\min}$ が求まり, b の範囲も決まるのである. 例えば, $\beta=2.0$, $b_{\max}=1.6$ と設定したとき, (5.17)式の計算により, $b_{\min}=0.24$ と求まる. よって b の平均 b_{average} は 0.92 となる. ここで, $b_{\max}$ を小さく設定すると, b の範囲は狭まり, b_{average} は大きくなる. 逆に, $b_{\max}$ を大きく設定すると, b の範囲は広くなり, b_{average} は小さくなる. このような関係に着目し, 以後 b について述べるときには b_{average} の値を用いる.

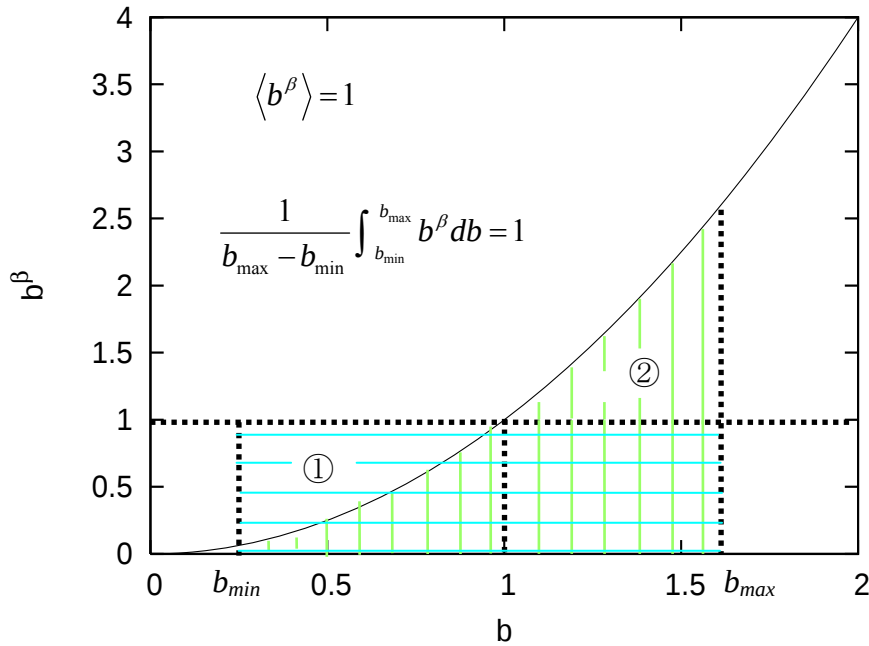


図 5.7 b の範囲を決める方法

5.4.2 ノイズ操作による左右への分布シフト

さらに、乗算ノイズ b もしくは加算ノイズ f の平均や範囲を変更することによって、分布を左右にシフトできることを発見した。図 5.8 は、 $\beta=2.5$, $b_{average}=0.99$, $f_{range}=1$ と設定して、 $f_{average}$ の値を変化させたときのランクサイズプロットである。横軸はサイズ、縦軸は順位を表している。それぞれ、 $f_{average}=4000$ は○, $f_{average}=20000$ は×, $f_{average}=40000$ は□に対応している。この図から、 $f_{average}$ を大きくすると分布は右へ、小さくすると左へ平行移動することが見てとれる。

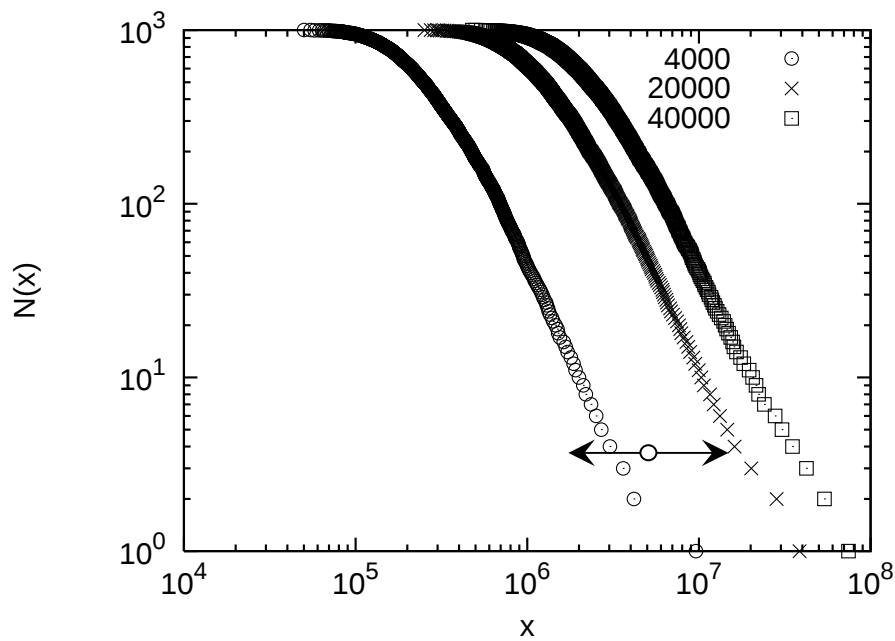


図 5.8 加算ノイズ $f_{average}$ の大きさと分布シフトの関係 $f_{average}$ はそれぞれ(○)=4000, (×)=20000, (□)=40000 を表す。

同様に、 $b_{average}$ を大きくすると分布が右に平行移動し、小さくすると左に移動する。ノイズ操作によって分布が左右にシフトするという特徴は、実データとのフィッティングにおいて明確な方針を与えてくれる。例えば、適当な変数設定で得られた分布が、実データにより作成した分布よりも左側にあったとする。この場合、 b もしくは f を大きくすることによって、分布が右に平行移動し、実データの分布に近づくのである。

ノイズと分布シフトの関係を表 5.1 にまとめた。 $b_{average}$ を固定して、分布を左にシフトさせたい場合、 $f_{average}$ を大きくすればよい。逆に分布を右にシフトさせたい場合、

$f_{average}$ を小さくすればよい. b と f の立場を入れ替えて, $f_{average}$ を固定して, 分布を左にシフトさせたい場合, $b_{average}$ を大きくすればよい. 逆に分布を右にシフトさせたい場合, $b_{average}$ を小さくすればよい. このように, b と f の大小関係には, トレードオフがある. このことについては, 5.5.3 節で詳しく述べる. ここで, f_{range} は分布に影響を与えないことを確認している. つまり, 加算ノイズのばらつきは, 大きくても小さくても関係ないということである.

表 5.1 ノイズと分布シフトの関係

Noise		Distribution	
		Shift to left	Shift to right
b	Average	Decrease	Increase
	Range	Wide	Narrow
f	Average	Decrease	Increase
	Range	No Influence	

5.4.3 サイト数変更による分布幅の拡大縮小

サイト数を変更することによって, 分布幅を変更することができる. 図 5.9 はサイト数を変化させたときのランクサイズプロットである. 横軸はサイズ, 縦軸は順位を表している. それぞれ, $N=100$ は○, $N=1000$ は×, $N=10000$ は□に対応している. この図から, サイト数を増やすと, 分布の形はそのままに, 分布幅を拡大できることが分かる. これは, サイト数が異なる対象(例えば都道府県なら 47, 市町村なら 1600 以上)に対して, 同じようにモデルを適用できることを表している.

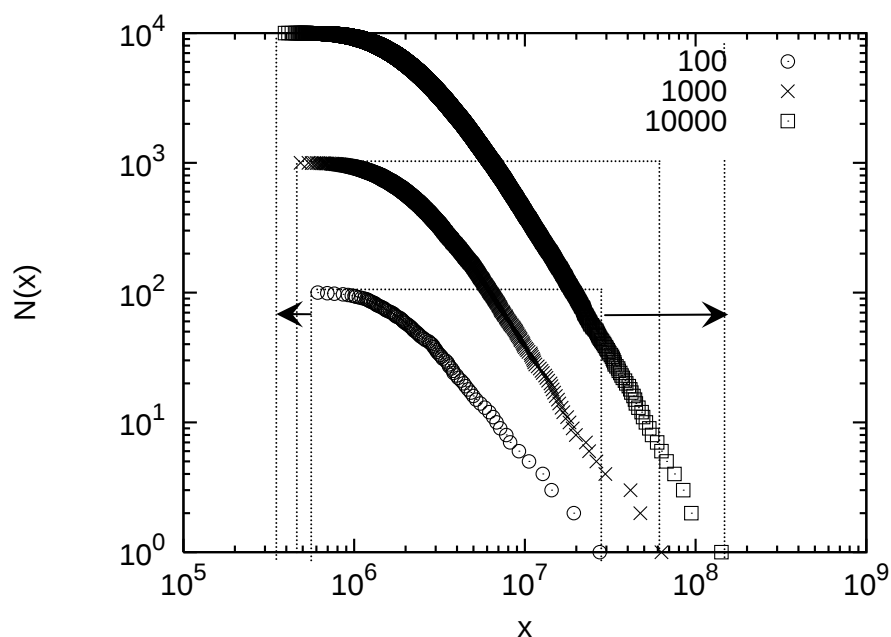


図 5.9 サイト数と分布幅の関係 N はそれぞれ (○)=100, (×)=1000, (□)=10000 を表す.

5.5 データフィッティング

都道府県と市町村, 2 つのレベルの実データに対して, モデルのシミュレーション結果のフィッティングを行う. フィッティングによって, 実データを最もよく再現できるシミュレーションの変数が分かる. この変数の妥当性についても, 5.5.4 節で検討する.

5.5.1 フィッティングの手続き

モデルと実データのフィッティングを行うに際して, 2 分探索に似た方法を用いて, 最適な変数を探索する. 指定した分布は, 乗算ノイズ b , 加算ノイズ f どちらを変更し

ていっても見つけることができるが, ここでは β , $b_{average}$, N を固定して最適な $f_{average}$ を探索する. 図 5.10 は $f_{average}$ を探索する手順を示している.

$f_{average}$ を探索する手順は, 以下の通りである.

Step1. 初期値として $f_{average}$ に小さい値①を最小値として設定し, シミュレーション結果の裾野部分の近似線を計算する. 実データとシミュレーション結果の平均二乗誤差を計算する. 平均二乗誤差の和は, 以下のような式で与えられる.

$$S_e = \sum (y_{real} - y_{sim})^2 \quad (5.18)$$

ここで, y_{real} と y_{sim} は, ランクサイズプロットにおける実データとシミュレーション結果の近似線上の順位に相当する.

Step2. $f_{average}$ に大きい値②を最大値として設定し, Step1 と同様に平均二乗誤差を計算する.

Step3. ①と②の中間の値を③として選択し, Step1 と同様の計算をする.

Step4. ①と②の誤差を比較し, もし, ①のほうが小さければ, ①と③の中間点を④として選択し, Step1 と同様の計算をする.

Step5. ①と③の誤差を比較し, もし, ③の方が小さければ, ③と④の中間点を⑤として選択し, Step1 と同様の計算をする. . . .

Step6. $f_{average}$ の移動量が 1 より小さくなったとき, このプロセスを終了する. このときの $f_{average}$ が, 設定した $b_{average}$ に対する最適な値である.

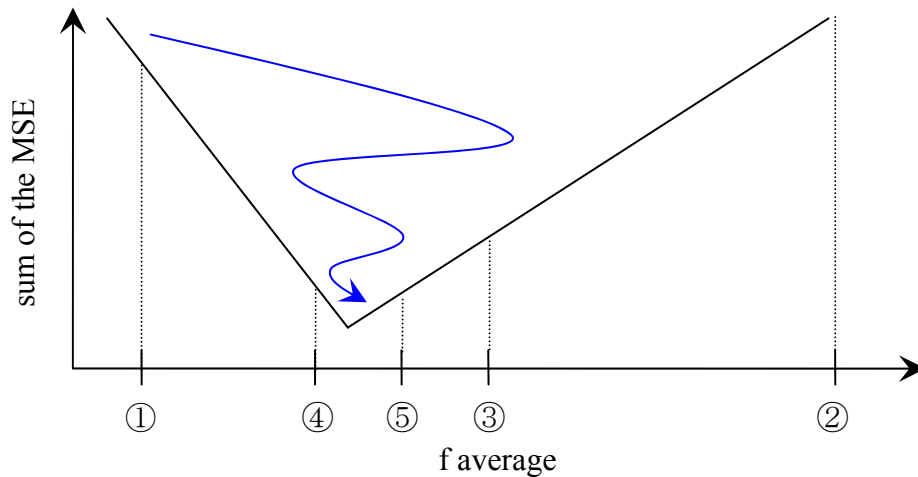


図 5.10 最適な $f_{average}$ の探索手順 $f_{average}$ が最適な値に近づくにつれて誤差は小さくなっていく.

5.5.2 シミュレーション結果と実データのフィッティング

まず, 1980 年から 2006 年までの都道府県の実データの平均値とシミュレーション結果のフィッティングを行った. この際, $\beta=2.49$, $b_{average}=0.99$, $N=47$ と設定したときの最適な $f_{average}$ は 34763.9 で, 平均二乗誤差は 0.17 であった.

次に, 市町村の実データとそれぞれの年の β , N に合わせた設定値の元でフィッティングを行った. それぞれの年で設定値を変えた理由は, 合併によって年ごとに市町村数が増えているため, 同じ土俵で評価できないからである. 1980 年のデータに対しては, $\beta=1.16$, $b_{average}=0.99$, $N=3257$ と設定し, $f_{average}$ を探索した. このときの最適な $f_{average}$ は 67.62 で, 平均二乗誤差は 1.49 であった. 2006 年のデータに対しては, $\beta=1.31$, $b_{average}=0.99$, $N=1822$ と設定し, $f_{average}$ を探索した. このときの最適な $f_{average}$ は 215.7 で, 平均二乗誤差は 1.15 であった.

図 5.11, 5.12, 5.13 は, それぞれ都道府県と市町村(1980 年, 2006 年)の実データ

に対するモデルのフィッティング結果である。これらの図から、市町村については、べき乗分布で近似した上位 20%の部分だけでなく、分布全域にわたってフィッティングできていることが分かる。つまり、数多くの自治体が平等に成長と転入・転出を繰り返していくと、裾野の広い分布になるのである。逆説的ではあるが、全ての自治体が平等であるが故に格差が生まれるといえる。一方、都道府県については、下位 80%の部分で合っていない。この理由としては、自治体数が少ないために生じた統計データ不足によるとも考えられるが、この点に関しては更なる検討が必要である。いずれにしても、ランダムな成長と転入・転出というシンプルな仮定で、実際の分布が再現できたといえる。

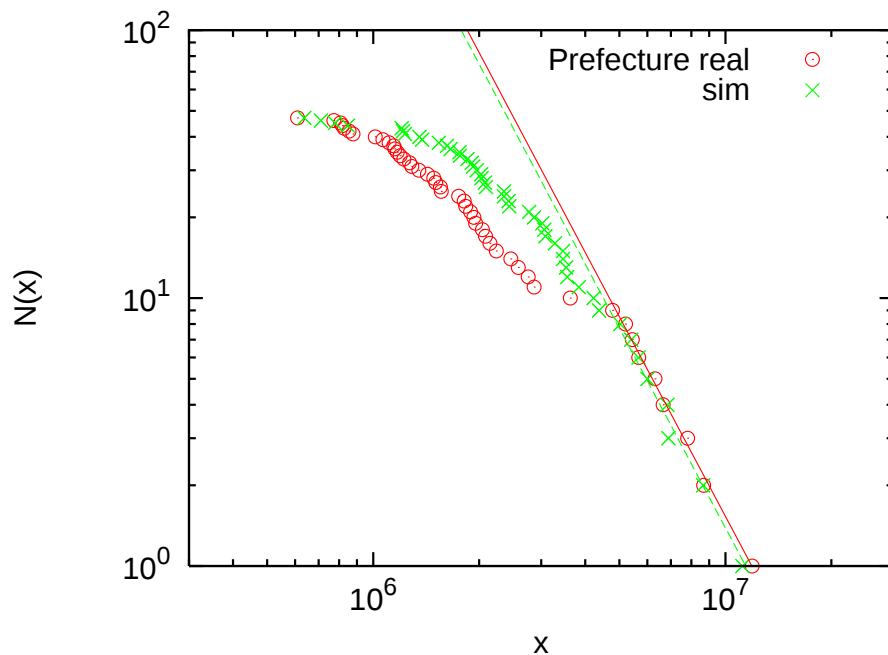


図 5.11 実データとモデルのフィッティング結果(都道府県 1980～2006 年の平均)

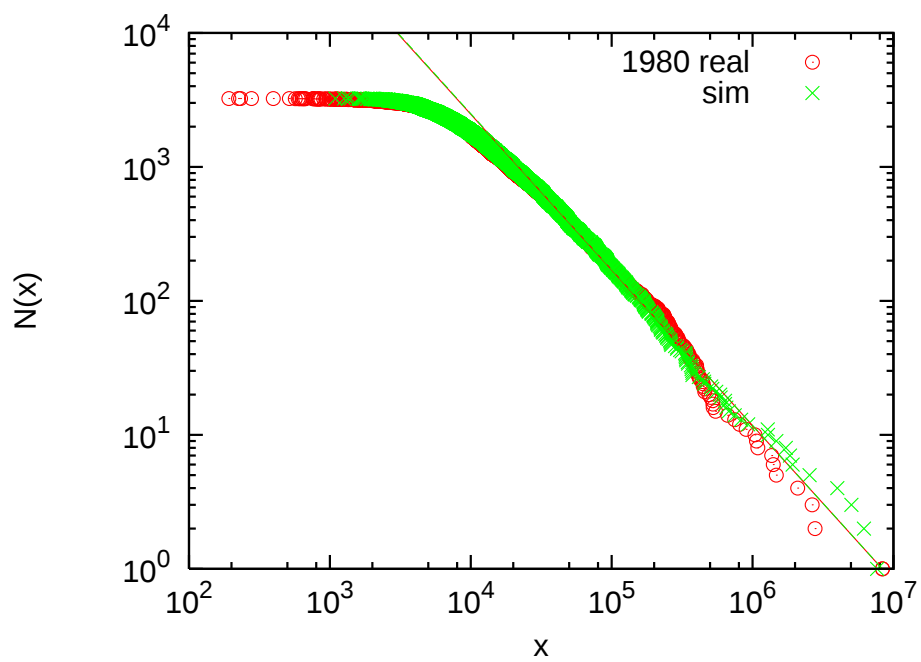


図 5.12 実データとモデルのフィッティング結果(市町村 1980 年)

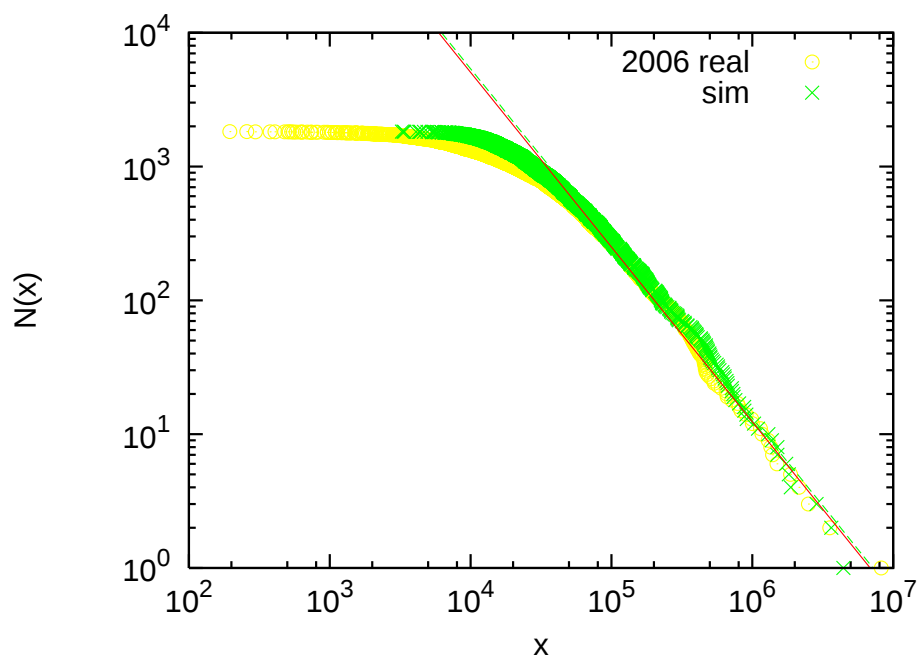


図 5.13 実データとモデルのフィッティング結果(市町村 2006 年)

5.5.3 変数間のトレードオフ関係

前節では、 β , $b_{average}$, N を固定して最適な $f_{average}$ の探索を行ったが、本質的には $f_{average}$ を固定して $b_{average}$ を探索しても同じことである。次に、実データとフィットするときの $b_{average}$ と $f_{average}$ の関係について調べた。図 5.14, 5.15 は、都道府県と市町村について、実データとフィットするときの乗算ノイズと加算の関係プロットしたものである。これらの図から、 $b_{average}$ が大きければ、 $f_{average}$ は小さくてよく、逆に、 $f_{average}$ が大きければ $b_{average}$ は小さくてもよい、というトレードオフ関係にあることが分かる。つまり、実データを再現するノイズの組み合わせは、無数にあるといえる。これらの組み合わせの一つが、現実世界の自然動態と社会動態に対応していると考えられる。そこで次に、どのノイズの組み合わせが実際の成長率と社会動態と一致するのかについて議論する。

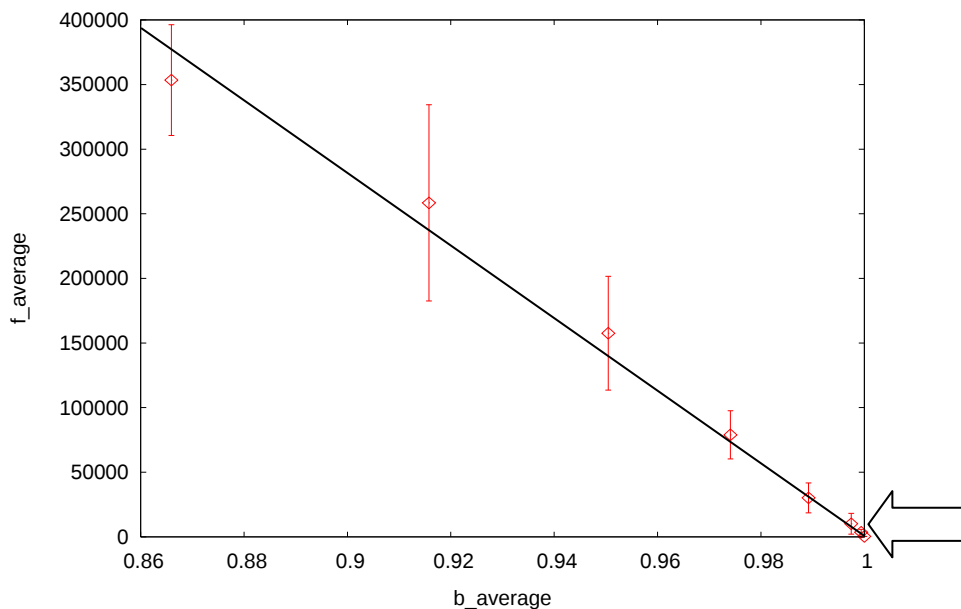


図 5.14 乗算ノイズと加算ノイズのトレードオフ関係(都道府県)

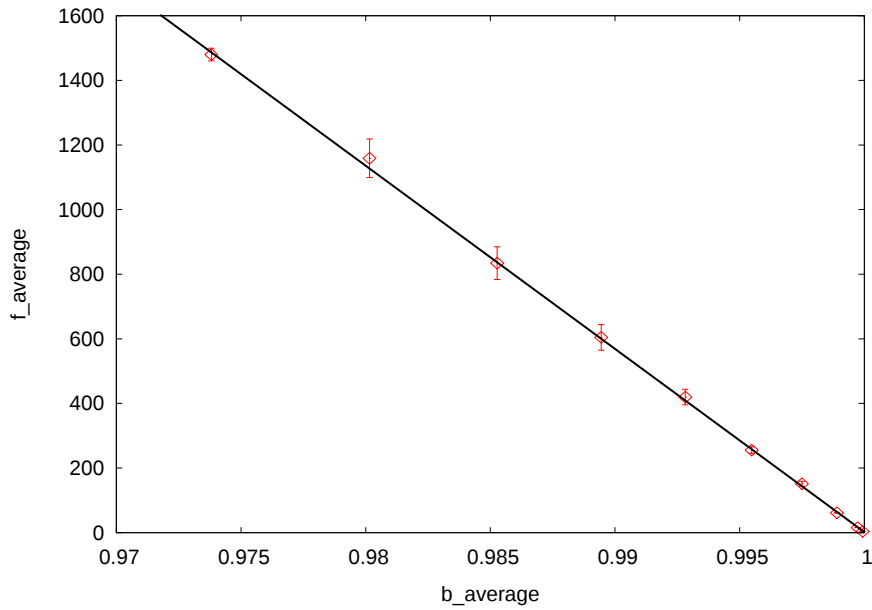


図 5.15 乗算ノイズと加算ノイズのトレードオフ関係(市町村 1980 年)

5.5.4 ノイズの組み合わせと実データの比較

ここで、どのノイズの組み合わせが現実と一致するのかについて考察する。本来、都道府県、市町村両方のデータについて考察を行うべきであるが、市町村についてはその数が増えているために、全ての市町村の成長率の変化を時系列で追うことができない。よってここでは、都道府県についてのみ考察する。乗算ノイズ $b_i(t)$ は、(5.11)式を変形して、

$$b_i(t) = \frac{x_i(t+1) - f_i(t)}{x_i(t)} \quad (5.19)$$

と書ける。この式に基づいて、1920 年から 2000 年までの $b_i(t)$ を求める。図 5.16 は、 $b_i(t)$ すなわち自然動態の成長率の推移を表している。横軸が西暦年、縦軸が成長率である。

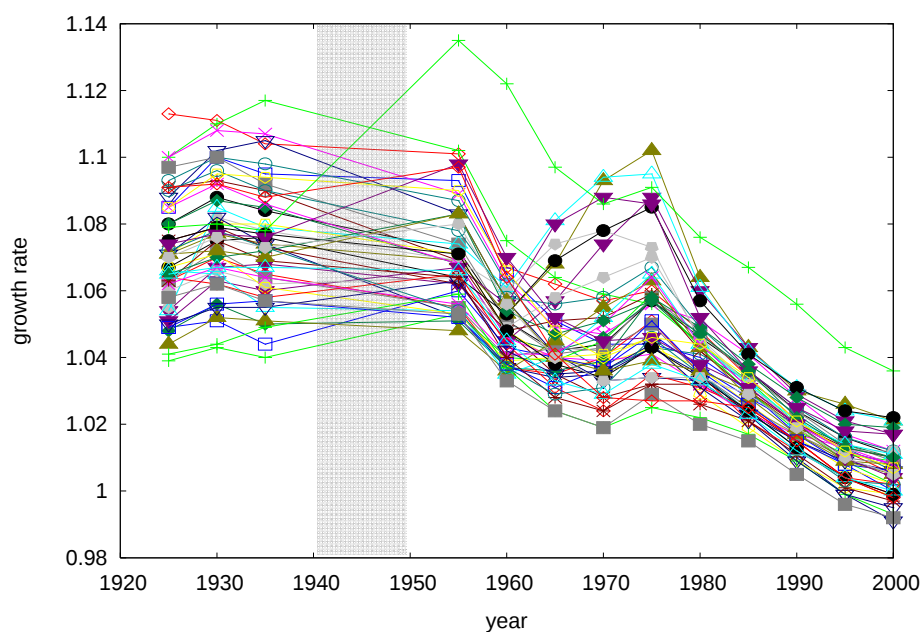


図 5.16 自然動態の成長率の推移⁵³

図 5.17 の上下限と平均を調べて、モデルの b の上下限, 平均と並べたのが図 5.18 である. 左のグラフの背景にも $b_{average}=0.995$ の範囲: ■, 0.997 の範囲: ■, 0.999 の範囲: ■を示した. モデルの b の範囲はおよそ 1 を平均として上下ほぼ対称であるのに対し, 実際の成長率は狭く, 平均も 1 より上のほうに位置する. これは, 1920 年から 2000 年という, 人口増加の時代のデータだからと考えられる. 今後, 実データの範囲が下がるとすると, 長期的に見たときにはモデルの b の範囲と実データの成長率の範囲の適合度は, さらに高まると予想される. いずれにしても, 実際のデータに近いモデルの b は, 図 5.14 の右下のあたりに対応することが分かる.

⁵³ 1940 年から 1950 年まではデータなし.

実際の成長率の平均値と上下限

モデルの b の範囲

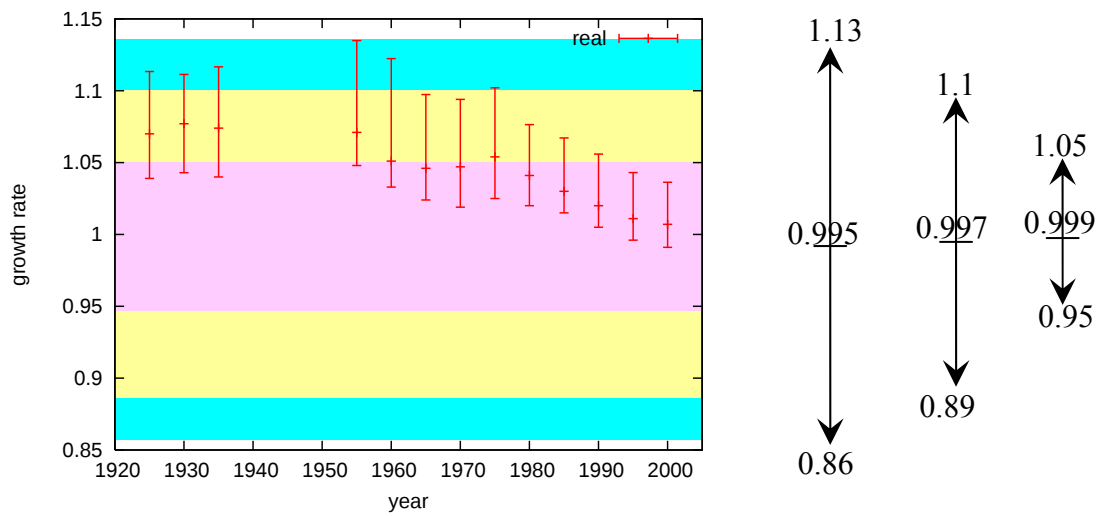


図 5.17 実データの成長率とモデルの b 範囲の比較

$b_{average}$ が 0.995, 0.997, 0.999 のときの最適な $f_{average}$ は, それぞれ 29917, 10103, 3515 である. 図 5.18 は, 社会動態の推移を表している. 横軸が西暦年, 縦軸が社会動態である. 図 5.18 における社会動態の平均を調べて, モデルにおける最適な $f_{average}$ と共にプロットしたのが図 5.19 である. この図から, 実際の社会動態の平均の変動とうまく一致しているのは, $f_{average}=3515$ であることがわかる.

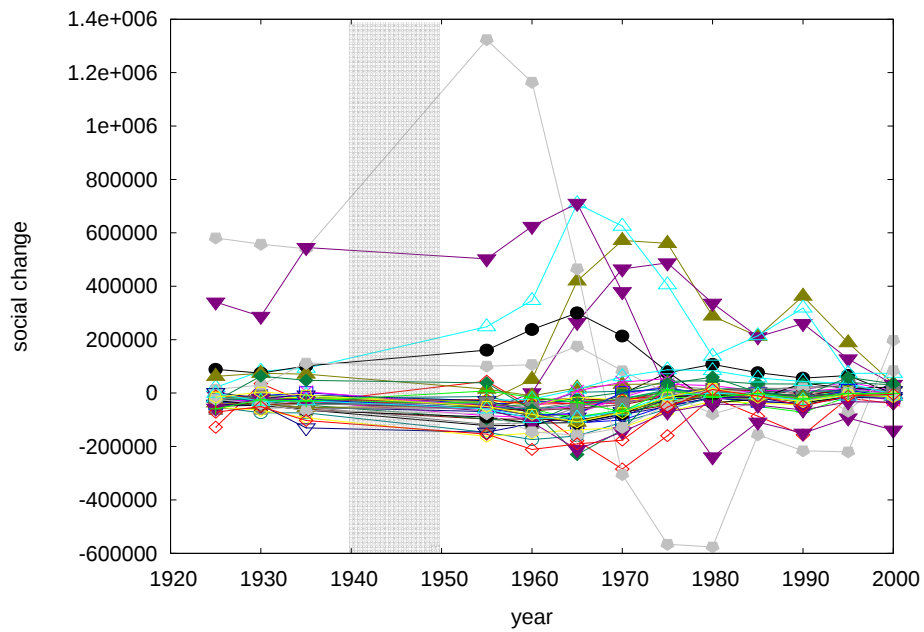


図 5.18 社会動態の推移

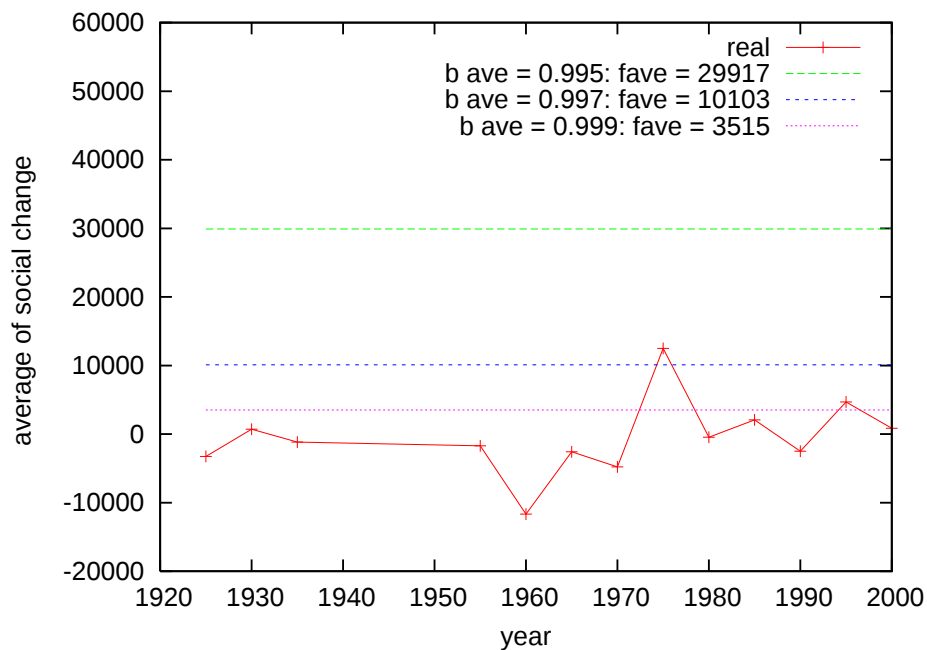


図 5.19 実データの社会動態とモデルの $f_{average}$ の比較

以上のことから、モデルによって得られたノイズの組み合わせ $b_{average}=0.999$, $f_{average}=3515$ が、実際の成長率、社会動態とうまく一致することが分かる。これらの値は、図 5.14 の矢印の部分に対応する。

5.6 まとめ

本章では、自治体の成長モデルとして、加算ノイズを伴った乗算的確率過程を提案し、実データとの比較検討を行った。まず、自治体の成長は、そのサイズに依存しないというジブラ則が成り立つことを確認し、このジブラ則に基づいた乗算的確率過程は対数正規分布を生成することを示した。そして、このジブラ過程にノイズを加えることによって、下位 80%は対数正規分布、上位 20%はべき乗分布で近似できるサイズ分布を生成できることを示した。さらに、この乗算的確率過程の変数の制御性、

性質 1) べき指数 β の変更による分布形の制御(これについては既知)

性質 2) ノイズ操作による左右への分布シフト

性質 3) サイト数変更による分布幅の拡大縮小

を発見し、この制御性を用いて実データへのフィッティングを行った。フィッティングの性能はよく、加算ノイズを伴った乗算的確率過程が、現実をうまく表現できるモデルであることを確認した。

また、実データとのフィッティング作業から、乗算ノイズと加算の間にトレードオフ関係があることを発見し、実際のデータとノイズの対応関係を明らかにした。このように、加算ノイズを伴った乗算的確率過程は、自治体の成長モデルとして、単純ながら現実の本質的なメカニズムを捉えているといえる。

しかし、このモデルは、自治体間の相互作用を想定していないので、加算ノイズの社会的意味合いである社会動態(転入と転出)が名目だけのものになってしまっている。そこで、次の6章では、自治体間での相互作用である人口移動を考慮したモデルを提案する。

第 6 章

移動モデル

6.1 はじめに

第1章でも述べたように、日本の人口は2005年をピークとして現在横ばい、そして減少という道をたどることは確実である。つまり、今後、人口の成長は見込めない社会状況である。

都道府県間の人口移動をネットワークとして捉えた場合、すべての県同士がつながっている完全グラフとみなすことができる。このようにすべての県同士がつながったネットワークにおいて、人々が自由に移動するという状況を考えた場合、どのようなメカニズムによって、現在のような人口分布に落ち着くのであろうか？それぞれの県が、人を引きつける力を持っているとすると、この力はどれくらい違うのであろうか？

上記のような問いに答えるため、壺モデルという確率モデルを考える。5章で扱った加算ノイズを伴った乗算的確率過程は、人口の成長過程を示すものだったが、壺モデルでは、総人口は変わらない。ここでは都道府県間の相互作用、すなわち人口の移動のみに着目する。人口の移動モデルとしては、2.2.3節でグラビティーモデルについて触れた。グラビティーモデルでは、時間の次元を立てずに f_{ij} という個々の地域間移動量の分析を目的とする静的なモデルであった。これに対し、壺モデルは時間の次元を明示的に取り込んだ動的なモデルである。また、壺モデルでは、個々の f_{ij} を説明するというよりは、それを集計した転入・転出量の説明を目指す。

6.2 モデル: 壺モデル

壺モデルは、比較的簡単な動作に基づく確率モデルであることから、統計物理や確率論の分野において広く研究されてきた[Ritort 1995, Bialas 1997]。本研究で用いるのは、大久保によって提案されたモデル[Ohkubo 2005]で、ネットワークの各ノード上に壺を配置して、隣接ノード同士でボールを交換するプロセスである。図 6.1 はそのプロセスの概略である。

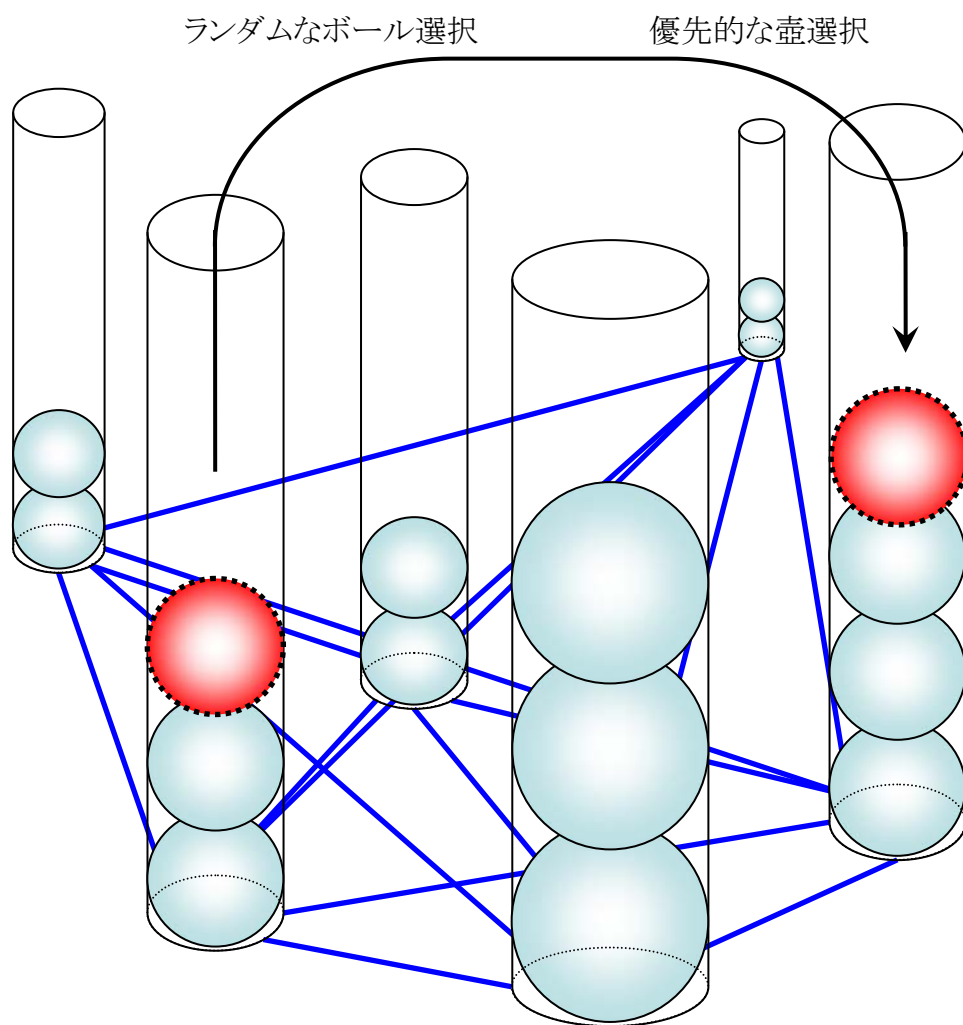


図 6.1 壺モデルのプロセス

このモデルでは、ネットワーク構造と変数がある状況を満たすとき⁵⁴、ベキ乗分布が生成できるとされている。そのダイナミクスは、以下のように定義される。

- Step1. N 個の壺を用意し、 M 個のボールをランダムに分布させる。また、各壺に誘引度 ω を、ある区間の一様分布として割り当てる。この誘引度は、各壺固有のボールを引きつける能力のようなものである。
- Step2. ランダムにボールを 1 つ選ぶ。ここで、選んだボールが入っている壺を i と書くことにする。
- Step3. 以下の確率で壺 i に隣接する壺から壺 j を選び、壺 i から壺 j へとボールを移動させる。

$$W_{i \rightarrow j} = \frac{(n_j + 1)^{\omega_j}}{\sum_{k \in \partial i} (n_k + 1)^{\omega_k}} \quad (6.1)$$

ここで、 n_j は壺 j に入っているボールの数、 ∂i は壺 i に隣接する壺の集合である。

- Step4. Step2 と Step3 の操作を繰り返す。

このモデルは、以下のように考えると、都道府県間の人口移動という解釈が可能となる。各壺は都道府県、ボールは人に対応すると考える。すると、壺同士でのボールの交換プロセスは、つながりあっている県同士で、人が異動し合うプロセスに対応する。また、上のダイナミクスは、第 4 章の図 4.7 と対応して、以下のような状況を指す。

- ・ 人口の多い県は、転出する人も多い(これはボールをランダムに選びだすこと

⁵⁴ 下の表において、ベキ乗分布が生成できるようなネットワーク構造と変数(ボールをひきつける力 ω) の組み合わせに○をつけた。

	完全グラフ	正則グラフ	ランダムグラフ
ω が異なる	○	○	○
ω が同じ			○

に対応している). しかし, 人口の多い県は, 転入してくる人も多い((6.1) 式は the rich-get-richer という優先的選択のメカニズムが含まれている)⁵⁵.

- それぞれの県は, 人を引きつけるための異なった能力を持っている(誘引度 ω が, 分布 $\phi(\omega)$ からランダムに選び出されることに対応している).

こうした人口移動のモデルにおいて, ネットワークの構造と誘引度の分布 $\phi(\omega)$ の選び方にまだ任意性がある. 現在の日本では, 交通手段の発達により, すべての県間で人口移動があるので, ネットワーク構造は完全グラフと考えてよい. 問題となるのは, 誘引度 $\phi(\omega)$ の区間設定である. この変数がある区間の一様乱数と仮定した場合, どの程度の区間幅のとき, 実データとよくフィットするのかを明らかにする.

6.2.1 初期状態

初期状態として, 壺の数 N は県の数に対応付けて 47 とした. また, ボールの数 M について⁵⁶は, 1980 年から 2000 年までの人口の平均を 1000 で割った値 112506 と 10000 で割った値 1250 の 2 パターンを設定した. つまり, 47 個の壺に, 112506 もしくは 1250 のボールがランダムに割り振られているという状況から, シミュレーションはスタートする. そして, 1~4 のプロセスを定常分布になるまで繰り返す.

6.2.2 定常状態

⁵⁵ ただし, ω が 0 の場合は, (6.1) 式の分子が 1 となり, 次の移動先として選ばれる確率が現在の人口とは無関係になってしまう.

⁵⁶ ボール 1 個は人間 1 人に対応させたかったが, コンピュータのメモリの都合で, 計算量が少なくて済む 1000 人と 10000 人の場合について検討した.

壺モデルのシミュレーションにおいて、ある時点での分布が定常分布になっているかどうかを判断することは難しい(移動を繰り返すことによって、まだまだ分布の形が変わっていく途中なのか、移動は起こっていても全体の分布は変わらない状態になっているのかが明確には分からない)。そこでまず、どのくらいの時間シミュレーションをまわせば、定常分布に落ち着くのかを調べた。その方法は以下のような単純なものである。

Step1. ある時間シミュレーションをまわし、分布を作成する。

Step2. 1 よりもさらに長い時間シミュレーションをまわし、分布を作成する。

Step3. Step1 と Step2 の分布を比較し、誤差が大きければ、さらに長い時間を設定し、1 と 2 の操作を繰り返す。

上記のプロセスを繰り返した結果、 $M \times 10$ 時間以上まわせば、定常分布になっていることを確認した。よって、これ以降のシミュレーションでは、 $M \times 10$ 時間後の分布を用いる。

6.2.3 変数の制御性

壺モデルには、壺の数 N 、ボールの数 M 、誘引度 ω と 3 つの変数があるが、誘引度 ω の区間をどのように設定するかによって、定常分布の形も変わってくる。まず、 ω の設定範囲と分布幅の関係について調べた。この調査の方針は、 ω の上限を 1 と決め、下限 d_{min} を変更することによって、分布がどのように変わるのか調べるというものである。図 6.2 は、誘引度 ω の範囲と分布幅の関係を表している。この図から、 ω の範囲が広いほど分布の範囲も広く、 ω の範囲が狭いほど分布の範囲も狭くなっていることが分かる。つまり、 ω の範囲を広くすることによって傾きを緩やかに(べき指数を小さく)することができ、 ω の範囲を狭くすることによって傾きを急に(べき指数を大きく)することができる。この制御性を利用して、実データとのフィッティングを行う。

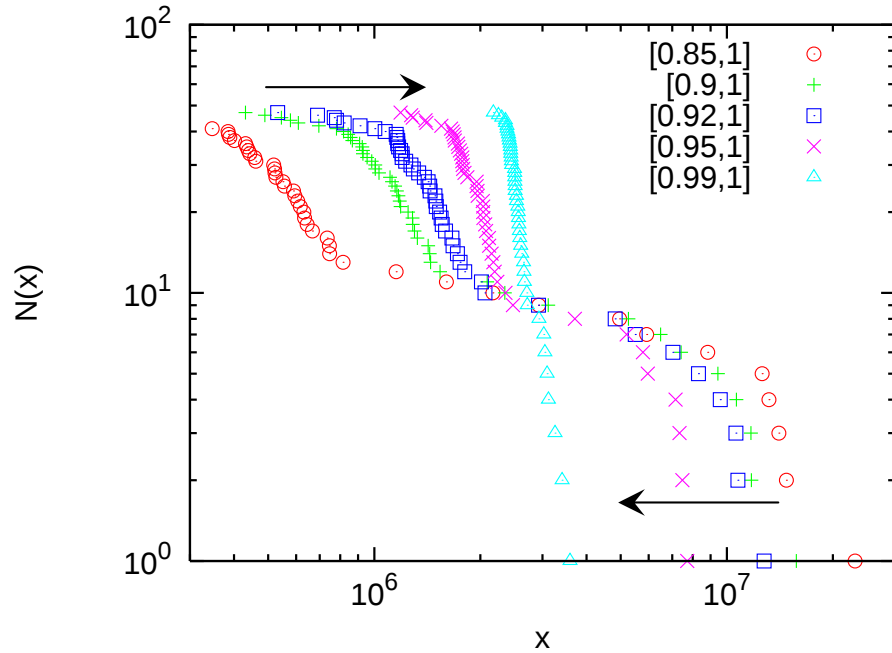


図 6.2 誘引度 ω の範囲と分布幅の関係 ω の範囲についてそれぞれ (○) $= [0.85, 1]$, (+) $= [0.9, 1]$, (□) $= [0.92, 1]$, (×) $= [0.95, 1]$, (△) $= [0.99, 1]$ を表している.

6.3 データフィッティング

モデルのシミュレーション結果のフィッティングを行う。フィッティングによって、実データとシミュレーション結果が最も接近するときの変数が分かる。この変数の妥当性についても、後で検討する。

6.3.1 フィッティングの手続き

モデルと実データのフィッティングを行うに際して、第 5 章と同様に、2 分探索に似た方法を用いて、誘引度 ω の最適な下限 d_{min} を探索する。図 6.3 は d_{min} を探索する

手順を示している.

d_{min} を探索する手順は, 以下の通りである.

Step1. 初期値として d_{min} に小さい値①を最小値として設定し, シミュレーション結果の裾野部分の近似線を計算する. 実データとシミュレーション結果の平均二乗誤差を計算する. 平均二乗誤差の和は, 以下のような式で与えられる.

$$S_e = \sum (y_{real} - y_{sim})^2 \quad (6.2)$$

ここで, y_{real} と y_{sim} は, 実データとシミュレーション結果の近似線上の順位に相当する.

Step2. d_{min} に大きい値②を最大値として設定し, Step1 と同様に平均二乗誤差を計算する.

Step3. ①と②の中間の値を③として選択し, Step1 と同様の計算をする.

Step4. ①と②の誤差を比較し, もし, ①のほうが小さければ, ①と③の中間点を④として選択し, Step1 と同様の計算をする.

Step5. ①と③の誤差を比較し, もし, ③の方が小さければ, ③と④の中間点を⑤として選択し, Step1 と同様の計算をする. ...

Step6. d_{min} の移動量が 0.01 より小さくなったとき, このプロセスを終了する. このときの d_{min} が, 誘引度 ω の最適な下限値である.

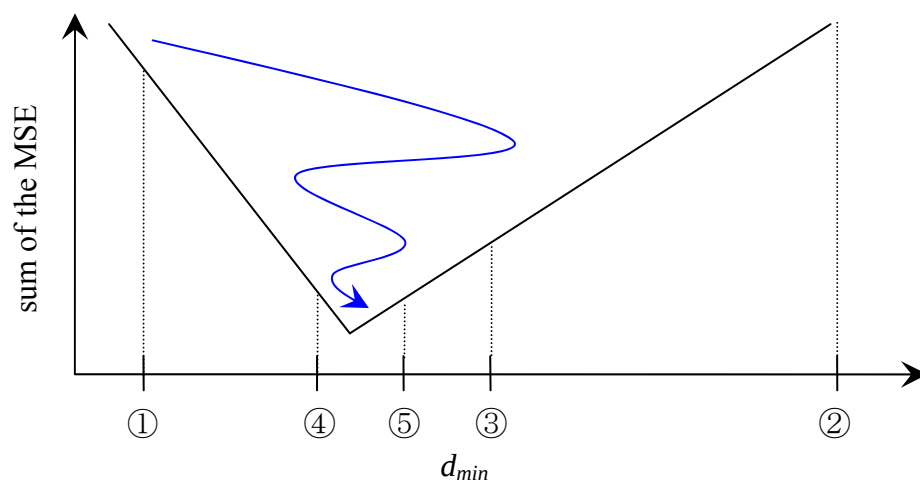


図 6.3 最適な誘引度の下限 d_{min} の探索手順 d_{min} が最適な値に近づくにつれて誤差は小さくなっていく.

6.3.2 シミュレーション結果と実データのフィッティング

ボール 1 個が 10000 人に対応する場合と 1000 人に対応する場合の, 2 パターンについてフィッティングを行った. 実データと比較するときには, ボールと人口の比率によってそれぞれの結果に対して 1000 倍, 10000 倍したものを用いた. 図 6.4, 6.5 は実データとシミュレーション結果の人口のサイズ分布である.

ボール 1 個が 10000 人に対応するとき, 誘引度の範囲が $[0.9, 1]$ で, 実データと最もよいフィッティングになる. また, ボール 1 個が 1000 人に対応するとき, 誘引度の範囲が $[0.92, 1]$ で, 実データと最もよいフィッティングになる. ボール 1 個が 1000 人に対応する場合のほうが精度は高いが, どちらの場合でも, 誘引度の範囲が 0.1 程度である. これを実際の県に当てはめてみると, 人をひきつける力について, 最も大きい東京が最も小さい鳥取の 1.1 倍であるということに対応する. この数値の差は, 大きいのか小さいのか一概に決めつけることはできないが, たくさんの人を引きつける東京の誘引

度が鳥取のわずか 1.1 倍というのは、直感からは大きく外れる結果であるといえる。

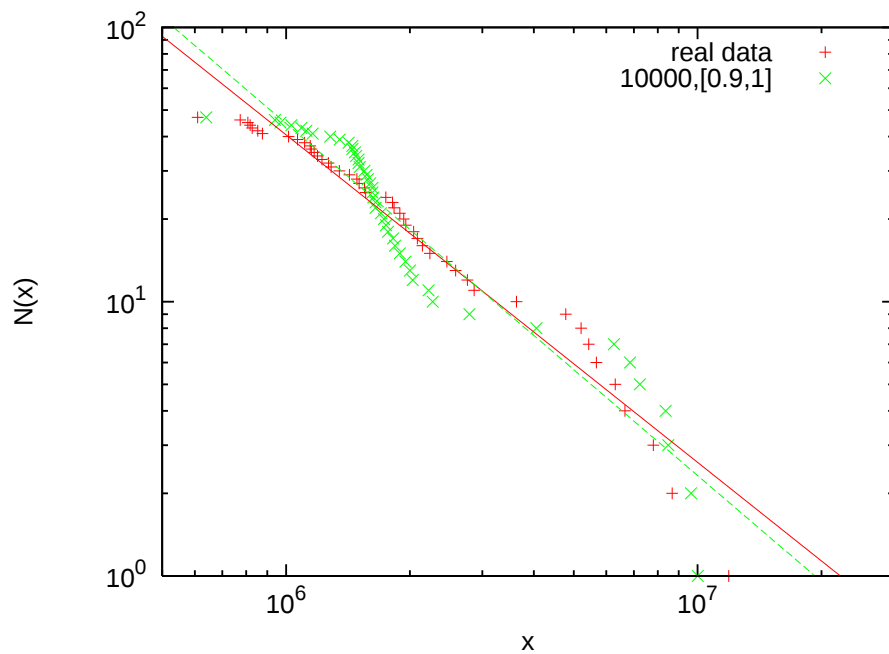


図 6.4 実データとモデルのフィッティング結果(ボール 1 個=10000 人)

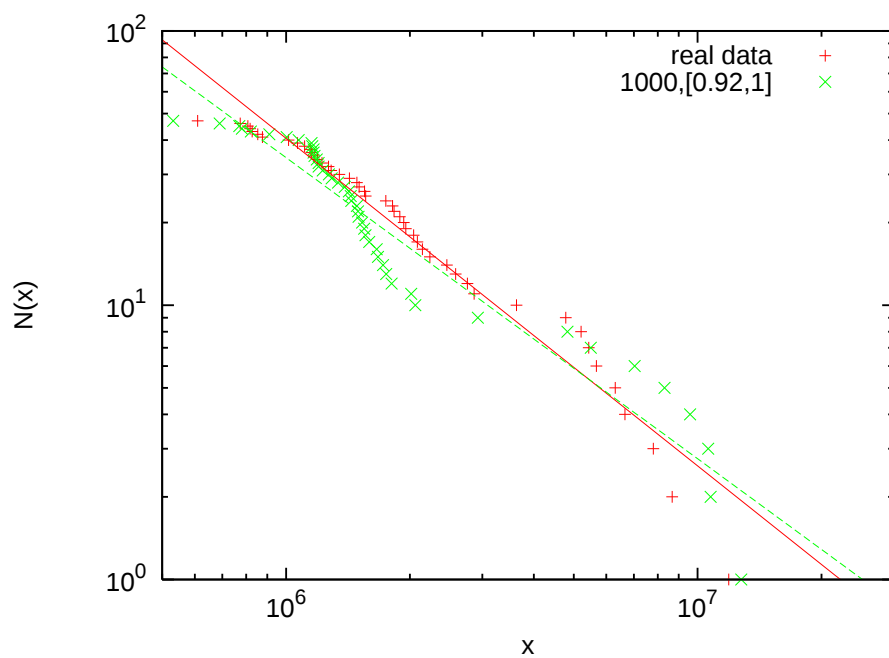


図 6.5 実データとモデルのフィッティング結果(ボール 1 個=1000 人)

つまり、すべての県同士がつながっていて、

- ・自由に人の移動ができる状況が整い、
- ・優先的な選択(人の多い県に引きつけられやすい)という移動メカニズムが働く

場合、人を引きつける力にわずかな差があるだけで、格差がどんどん広がっていき、現在のようなサイズ分布が再現できるものと考えられる。

6.4 定常状態に収束した後の移動量の比較

第4章では、1980年から2006年までの移動量のサイズ分布が、べき乗則に従うと述べたが、この壺モデルにおいても、定常状態に収束した後のボールの移動量のサイズ分布が、べき乗則に従うのか、実データとシミュレーション結果を比較する。

1980年から2000年間の1年間の人口移動者数の平均は310万人、人口移動率⁵⁷でいうと2.5%である。この人口移動率は図6.6に示すように、2.2%~2.9%と安定した推移を示している。壺モデルにおいて、定常状態に収束した後のある時点から、2.5%のボールが移動した時点をも、シミュレーション上での1年とする。この1年間の移動量のサイズ分布について、実データとシミュレーション結果を比較するのである。比較する実データは、1980年から2000年までのサイズ分布の平均とする。図6.7は壺モデルにおけるシミュレーションの推移を表している。収束後の1年間は、この図における赤い区間を指す。

⁵⁷ 総人口に占める移動者数の割合

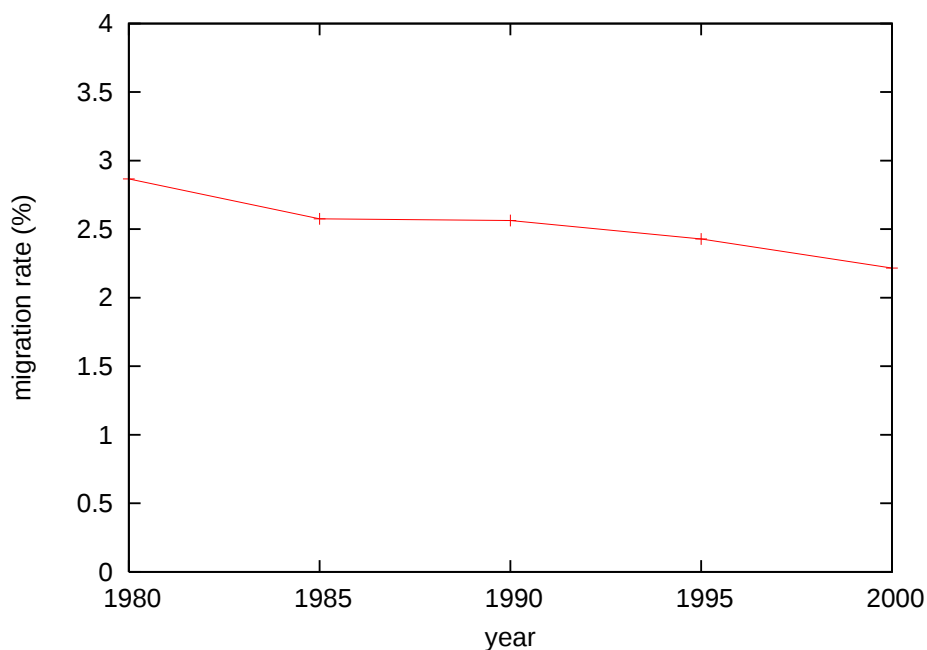


図 6.6 人口移動率の推移(1980 年～2000 年)

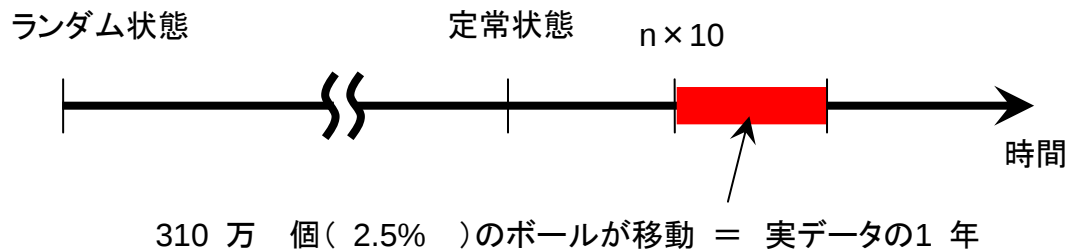


図 6.7 壺モデルの推移

図 6.8～6.11 は、それぞれ転入・出量のサイズ分布について実データとシミュレーション結果をプロットしたものである。全ての比較結果において、実データとシミュレーション結果がよく一致していることがわかる(ここでもボール 1 個=10000 人よりも 1000 人のほうが精度もよい)。このことから、壺モデルの優先的選択というメカニズムと誘引度の設定に妥当性があるといえる。

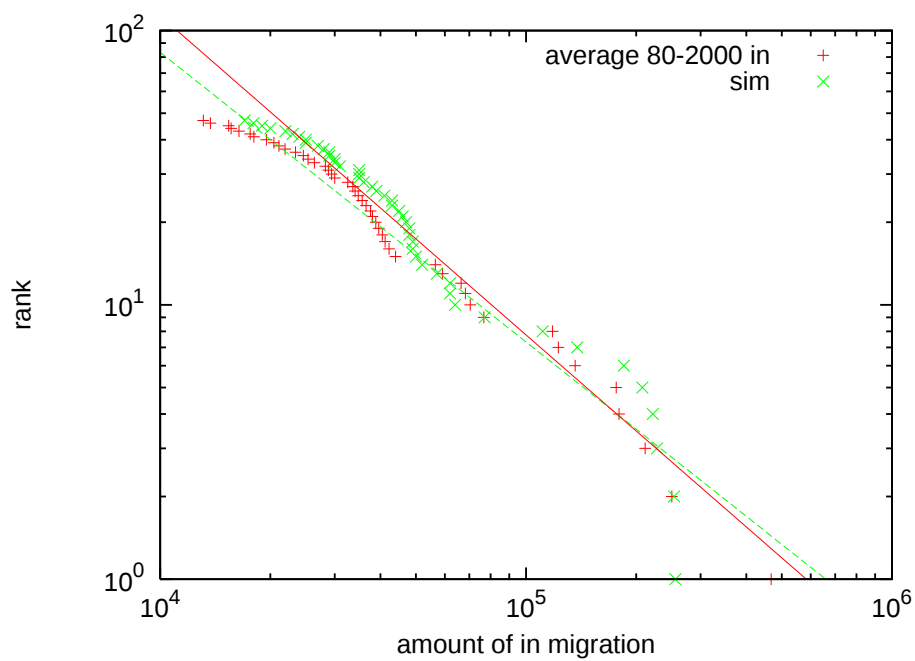


図 6.8 転入量のサイズ分布 (ボール 1 個=10000 人)

(+): 実データ, (x): シミュレーション結果

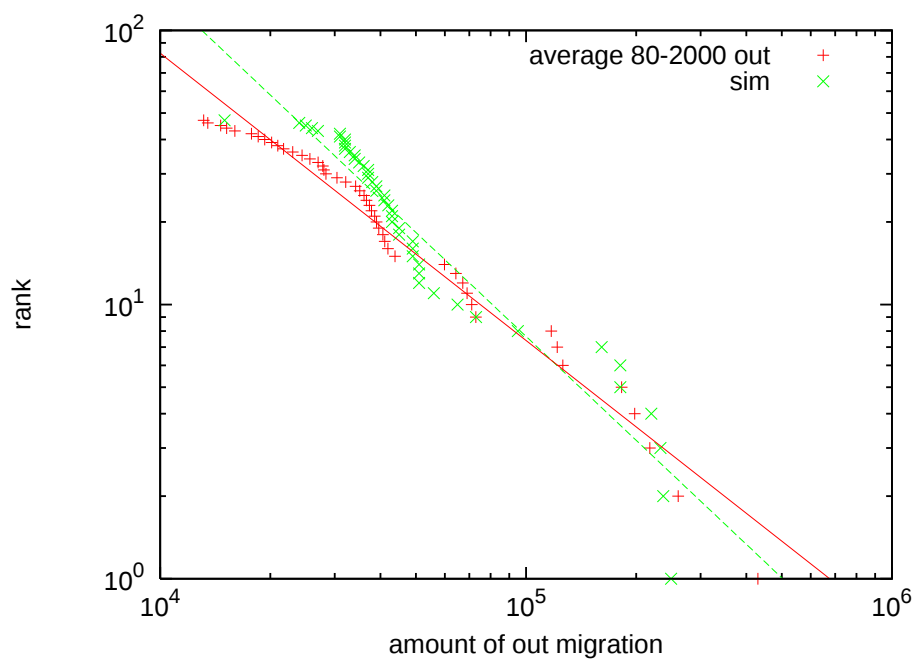


図 6.9 転出量のサイズ分布 (ボール 1 個=10000 人)

(+): 実データ, (x): シミュレーション結果

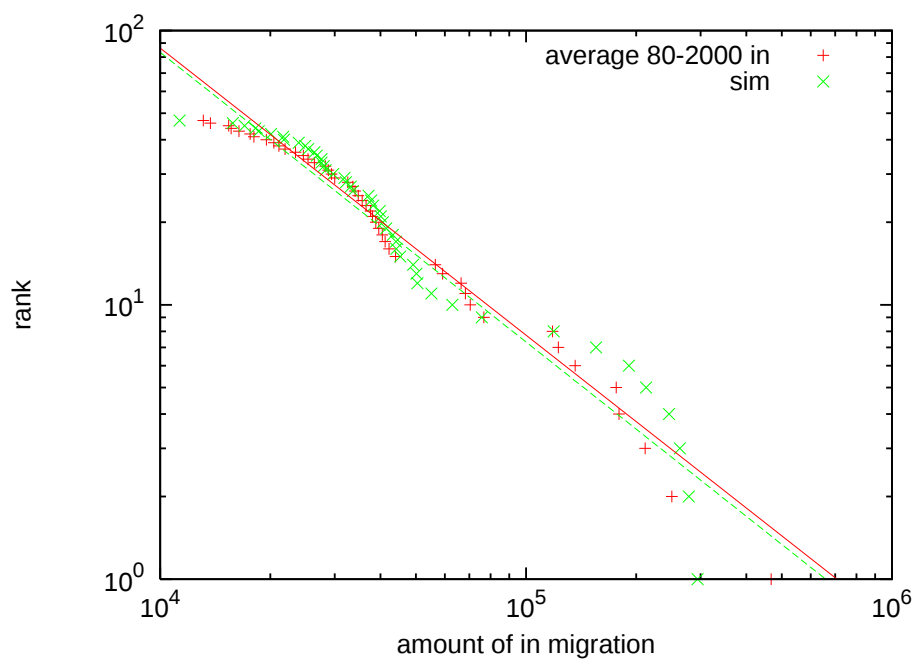


図 6.10 転入量のサイズ分布 (ボール 1 個=1000 人)

(+): 実データ, (x): シミュレーション結果

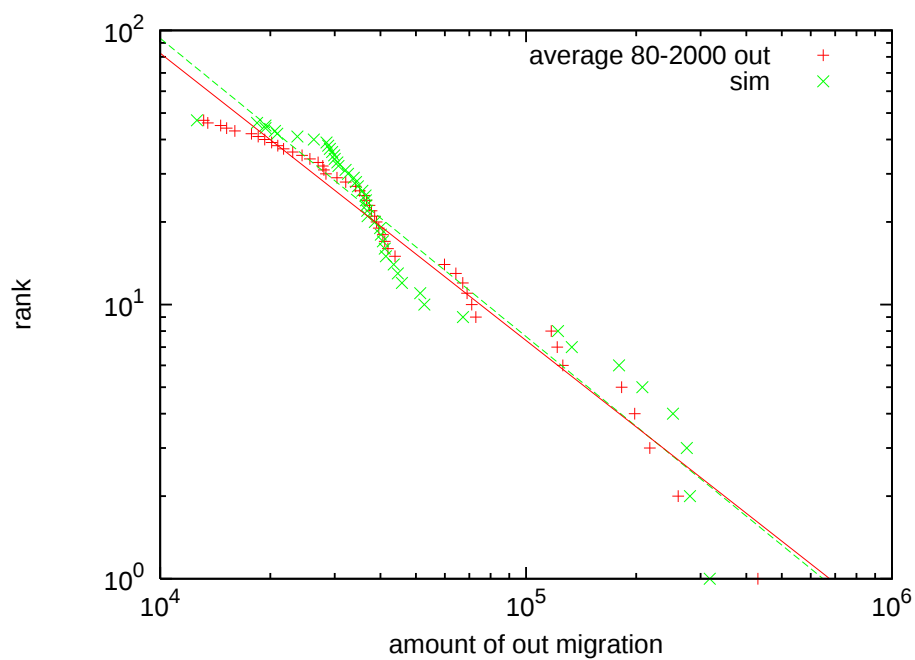


図 6.11 転出量のサイズ分布 (ボール 1 個=10000 人)

(+): 実データ, (x): シミュレーション結果

6.5 まとめ

本章では、都道府県を壺に、人をボールに対応付けて、完全グラフとした壺間のネットワーク内をボールが移動するという人口移動モデルを提案した。移動先が選ばれる際のルールは、ボールの分布がべき乗則に従うことを利用して、多くのボールが入っている壺ほど、次の移動先として選ばれやすいという優先的選択メカニズムになっている。

特に、各壺がボールを引きつける力(誘引度)をある区間の一様乱数と設定し、この区間を操作することによって、実データへのフィッティングを行った。ここでもフィッティングの性能はよく、壺モデルが、現実をうまく表現できるモデルであることを確認した。

さらに、定常状態後のボールの移動量のサイズ分布が、実際の移動量のサイズ分布とよくフィットすることを確認した。

このように、壺モデルもまた、自治体の人口移動モデルとして、単純であるが現実の本質的なメカニズムを捉えているといえる。

第 7 章

結論

7.1 本研究のまとめ

本論文は、大きさの分布という視点から、人口現象という一見複雑に見える現象の背後にある基本メカニズムを明らかにすることを目的に、これまで行ってきた研究成果をまとめたものである。以下にその成果を要約する。

第 1 章では、現在の世界の人口から日本の人口へと視点を絞り、2 つの問いを設定した。

問 1) なぜ、人口が多いほど変動も大きいのであろうか？

問 2) なぜ、ある地域が大きく発展し、別の地域がしないのであろうか？

第2章では、人口に関する研究が、より細分化して精緻化していることの反動として、逆に全体的な理解が得られないという要素還元主義のジレンマについて触れ、統計力学的アプローチの意義を述べた。これは、予測不能なミクロな自由度は一旦**確率的な変数**として扱い、マクロな世界の記述につなげ、変数の意味については後から考察するという、逆問題として人口現象を捉えるアプローチである。様々な要因を細かく計算するのではなく、系のミクロな詳細に依存しない普遍的なふるまい(分布)を探しだし、それらを的確に記述しようとするこのアプローチは、人口に関する研究はもちろん、その他の社会科学にも応用できるものと考えられる。

第3章では、複雑な現象を分布として理解することの意義と、その考え方の基本事項を整理した。また、裾野の長い分布になるようなデータを、インターネット等から探索し、それらの分布に2つの解釈を与えた。1つ目は、べき乗分布が多くの場合に見られることから、それを基本として、下位の部分が押さえつけられて対数正規分布になるという解釈である。これは、不平等が拡大しすぎるのを抑える社会的圧力によるものと考えられる。2つ目は、対数正規分布が基本にあり、上位の部分が持ち上げられてべき乗分布になるという解釈である。これは、上位のものがより大きくなる開放的な場合に生じると考えられる。どちらの解釈がより妥当であるかについては、今後さらな

る検討が必要であろう。

第 4 章では、1980 年以降の都道府県、市町村の人口サイズ分布が、上位 20% はべき乗分布、下位 80% は対数正規分布(都道府県、市町村)で近似できる一方、1 本のべき乗分布(都道府県)でも近似できるというように、2 つの異なる見方ができる定常分布になっていることを示した。1960、70 年代の変動が大きい過渡期を過ぎ、1980 年以降は、定常状態になっているのである。また、移動量についても、べき乗則が成り立つことを示した。

第 5 章では、自治体の成長モデルとして、加算ノイズを伴った乗算的確率過程を提案し、成長率と転入・転出がランダムであると仮定することによって実データを再現した。これは、どの自治体も大きくなる可能性と小さくなる可能性が平等にあることを意味する。まず、自治体の成長は、そのサイズに依存しないというジブラ則が成り立つことを確認した。これは、サイズの変化の大きさが、ある確率変数とサイズのかけ算で決まるということを意味している。ここで、第 1 章で挙げた問い 1 に対する答えが導かれる。成長率は、サイズとは無関係であるが、たまたま大きなサイズの県にたまたま大きな成長率がかけられたら、そのサイズの変化も大きくなるのである。この単純な乗算的確率過程をジブラ過程(Gibrat's process)という。このジブラ則に基づいた乗算的確率過程は、対数正規分布を生成することが知られているが、このジブラ過程にノイズを加えることによって、下位 80% は対数正規分布、上位 20% はべき乗分布で近似できるサイズ分布を生成できることを示した。さらに、この乗算的確率過程の変数によって、左右への分布シフト、分布幅の拡大縮小といった制御ができることを発見し、この制御性を用いて実データへのフィッティングを行った。シミュレーション結果のフィッティングの性能はよく、加算ノイズを伴った乗算的確率過程が、現実をうまく表現できるモデルであることを確認した。

また、実データとのフィッティング作業から、乗算ノイズと加算ノイズの間にトレードオフ関係があることを発見し、実際のデータとノイズの対応関係についても議論した。このように、加算ノイズを伴った乗算的確率過程は、自治体の成長モデルとして、現実の本質的なメカニズムを捉えているといえる。

第6章では、都道府県を壺に、人をボールに対応付けて、完全グラフとした壺間のネットワーク内をボールが移動するという人口移動モデルを提案した。移動先が選ばれる際のルールは、多くのボールが入っている壺ほど、次の移動先として選ばれやすいという優先的選択メカニズムに基づいている。

各壺にボールを引きつける力(誘引度)を設定し、これを操作することによって、実データへのフィッティングを行った。ここでもフィッティングの性能はよく、壺モデルが、現実をうまく表現できるモデルであることを確認した。さらに、定常状態後のボールの移動量のサイズ分布が、実際の移動量のサイズ分布とよくフィットすることを確認した。つまり、わずかな誘引度の差と優先的選択という単純なルールから、実際の人口サイズ分布が生成できるのである。

以上、複雑な人口現象に対して、そのサイズ分布に規則性を見出だし、その規則性を再現できる確率モデルについて検討した。さらに、実データとのフィッティング法を提案して、比較・検討を行った。しかしながら、第1章で挙げた問い2には、未だ十分に答えられていない。ジブラ則が成り立っている以上、どのサイトが成長し、どのサイトが成長しないかを、あらかじめ予測することはできない。そもそもあるサイトが成長し、あるサイトが衰退するのに、特別な理由は要らないのである。総人口とサイト数が与えられれば、あらかじめあるべきサイズ分布へと落ち着くことは決まっているようなのである。平等な成長と自由な移動が保障された社会においては、地域人口の格差が必然的に生じるという帰結を導くことができる。そのような分布になった理由を、経済的、政治的、地理的な解釈から説明はできるが、それは後付けでしかない。むしろ、ある範囲で合理的な説明ができる仮説同士が矛盾するという、要素還元主義のジレンマに陥る原因となる。人間の社会活動の多くに、裾野の長い分布が見られることから、むしろ、あるべき分布に落ち着くように、さまざまな要因が発生すると解釈するほうが、自然である。

このように解釈すると、社会政策への適用の可能性が出てくる。個々の自治体の発展と衰退という詳細は分からなくても、総人口、サイト数が分かっているならば、成長率と転入・転出量のどちらか一方のトレンドからもう一方を予測することができる。例えば、今後出生率がさらに低下すると仮定して、現在の定常分布を維持するためには、移

民をどの程度受け入れなければならないのか、といった問題を考える目安になる。

7.2 今後の研究課題

今回、人口の分布に着目して、規則性とその生成メカニズムについて検討したが、この規則性は、人口の分布だけではなく人間の社会活動全般、さらには自然現象と多くのものに当てはまるものでもあった。このことは、今回議論したモデルおよび実データとのフィッティング手法が、他の多くの現象にも適用できることを意味する。今後、社会、自然界を問わず、多くの現象にこのモデルを適用し、これらの現象を説明できる変数の意味を考えていくことが、大きな課題である。

また、成長モデルとして提案した加算ノイズを伴った乗算的確率過程は、相互作用を想定しないものであった。次に、移動モデルとして提案した壺モデルは、成長を想定しないものであった。これら 2 つのモデルは、それぞれ単独で、現実のデータを再現することができたが、成長しながら相互作用もしているのが本当の社会である。今後、これら 2 つのモデルを融合させることが課題である。2 つのモデルの融合の方向性の 1 つとして、加算ノイズを伴った乗算的確率過程の加算ノイズの項に、壺モデルの移動確率を入れ込むことも考えられよう。

さらに、単純な基本メカニズムから実際のサイズ分布を再現するという本研究は、人口現象をグローバルに理解するための一步にすぎない。予測不能なミクロな詳細を確率的な変数として置き換えることによって全体像の理解につなげることができたが、今後はこの全体像を念頭に置きつつ、経済的要因などの詳細の解明を進めていきたい。

付録

付録1 コーホート要因法

コーホート要因法(cohort-component method)は、ある時点の性・年齢別人口を基準人口とし、年齢が5歳階級別の移動率、生存率、出生率を与えて5年先の性・年齢別人口を推計し、そのプロセスを繰り返すことによって、どこまでも将来の人口を推計することが可能な方法である。移動率は、純移動率を用いる場合がほとんどであるが、転入率と転出率に分けて設定することも可能である。純移動率を設定する場合には、センサス人口と生命表の静止人口を用いて過去の動向を把握し、その傾向を勘案して推計のパラメータとして与えるのが一般的である。

ある地域*i*における*t*年の年齢階級*a* (たとえば 15～19 歳)の男子または女子の人口が、*t*+5年に年齢階級*a*+5 (上記の例では 20～24 歳)に移行するときの純移動率 ${}_i m_a^t$ は、以下のようになる。

$${}_i m_a^t = ({}_i P_{a+5}^{t+5} / {}_i P_a^t) - {}_i S_a^t \quad (\text{A1.1})$$

ここで、 ${}_i P_a^t$ は地域*i*における*t*年の年齢*a*の男子または女子の人口であり、*t*+5年の年齢階級*a*+5の人口 ${}_i P_{a+5}^{t+5}$ で割った値をコーホート変化率と呼ぶ。コーホート変化率は、移動と死亡による変化の両方を含んだ率である。第2項の ${}_i S_a^t$ は地域*i*における*t*年の年齢*a*の男子または女子の人口が、*t*+5年に年齢階級*a*+5に移行するときの生存率で、通常、生命表⁵⁸の関数である静止人口を用いて*t*年と*t*+5年それぞれの生存率を求め、この平均値を用いる。このように、純移動率はコーホート変化率から生存率を引いた値である。その値が0より大きければ転入超過、小さければ転出超過を意味する。

推計に用いる純移動率は、過去の純移動率を踏まえて設定される。ある時点の人口から5年先の人口を推計する際、(A1.1)を変形して、

⁵⁸ 生命表は、ある時期における死亡状況(年齢別死亡率)が今後も変わらないとしたとき、各年齢に達したものが平均してあと何年生きられるかを、死亡率、生存数、平均余命等の生命関数によって表現したものである。

$${}_iP_{a+5}^{t+5} = {}_iP_a^t ({}_im_a^t + {}_iS_a^t) \quad (\text{A1.2})$$

となる. t 年から $t+5$ 年の間の出生数 ${}_iB_a^t$ は,

$${}_iB_a^t = \sum_{a=15}^{45} (5/2) ({}_iP_a^t + {}_iP_a^{t+5}) {}_if_a^t \quad (\text{A.1.3})$$

となる. ここで, ${}_iP_a^t$ は i 地域における t 年の年齢階級 a ($a \sim a+4$ 歳) の女子人口で

ある. ${}_if_a^t$ は i 地域における $t \sim a+5$ 年の年齢階級 a の女子の出生率であり, 実績値は $t \sim t+5$ 年の年齢階級 a の女子の出生率の平均値を用いることが一般的である.

以上で述べたように, コーホート要因法による地域の将来人口推計を実施するためには, 当該地域における過去の性・年齢別人口, 生命表, 母親の年齢 5 歳階級別出生数などの資料が必要になる. 日本の場合, 都道府県より小さい地域では, 人口動態統計に基づくこれらのデータが入手困難なことから, コーホート要因法を用いることは容易ではない. なお, コーホート要因法による推計を行った場合, 推計結果として, 将来の性・年齢別人口以外に, 期間ごとの出生数, 死亡数, 純移動数を得ることができる.

付録 2 地域人口分布の推移

日本の地域人口分布は、これまでどのように推移してきたのであろうか。小山は、このような問いに答えるために、古人口学の手法を用いて、統計データのない縄文時代からの地域人口を推計した[小山 1984]。この推計をもとに鬼頭は、日本の地域を16に分けて、縄文早期からの地域人口推計をまとめている[鬼頭 2005]。図 A2.1 は、16 地域と都道府県の対応関係を地図上で表したものである。この 16 地域の推計人口を、日本全人口に占める割合に変換して色わけ(15%以上:赤, 12~15%:黄, 9~12%:緑, 6~9%:水色, 3~6%:青, 3%以下:灰色)し、時代別の地域人口分布として可視化したのが図 A2.2 である。

図 A2.2 について、左上から順に説明する。縄文時代(4300B.P.⁵⁹)においては、南関東と東山で 15%以上と、**東**に人口が集中しているのに対し、西日本にはほとんど人がいない。それが、弥生時代(1800B.P.)になると、15%以上の地域がなくなり、西日本の方にも人が住むようになった。そして、平安(1150)、戦国(1600)と人口の中心的地域が**東へ西へ**と移動した後、江戸時代(1721)には青と灰色の地域が少なくなり、**地域分散**となる。最後に、近代(1995)は東京**一極集中**の傾向が強まっている。この長期的な地域人口分布の推移から、日本の人口は関東から徐々に分散していき、近代になってまた関東に集中していることがわかる。また、一度栄えた地域が、そのままずっと繁栄を維持できるとは限らないことも分かる。近代において、繁栄の極みにある東京の今後はどうなるのであろうか。近代以前と近代で、決定的に異なる点は、人々の移動能力である。誰もが自由に素早く移動できるようになった近代は、別の世界であるともいえる。

⁵⁹ B.P.は「Before Present」の略で、年代の表記号(考古学用語)である。

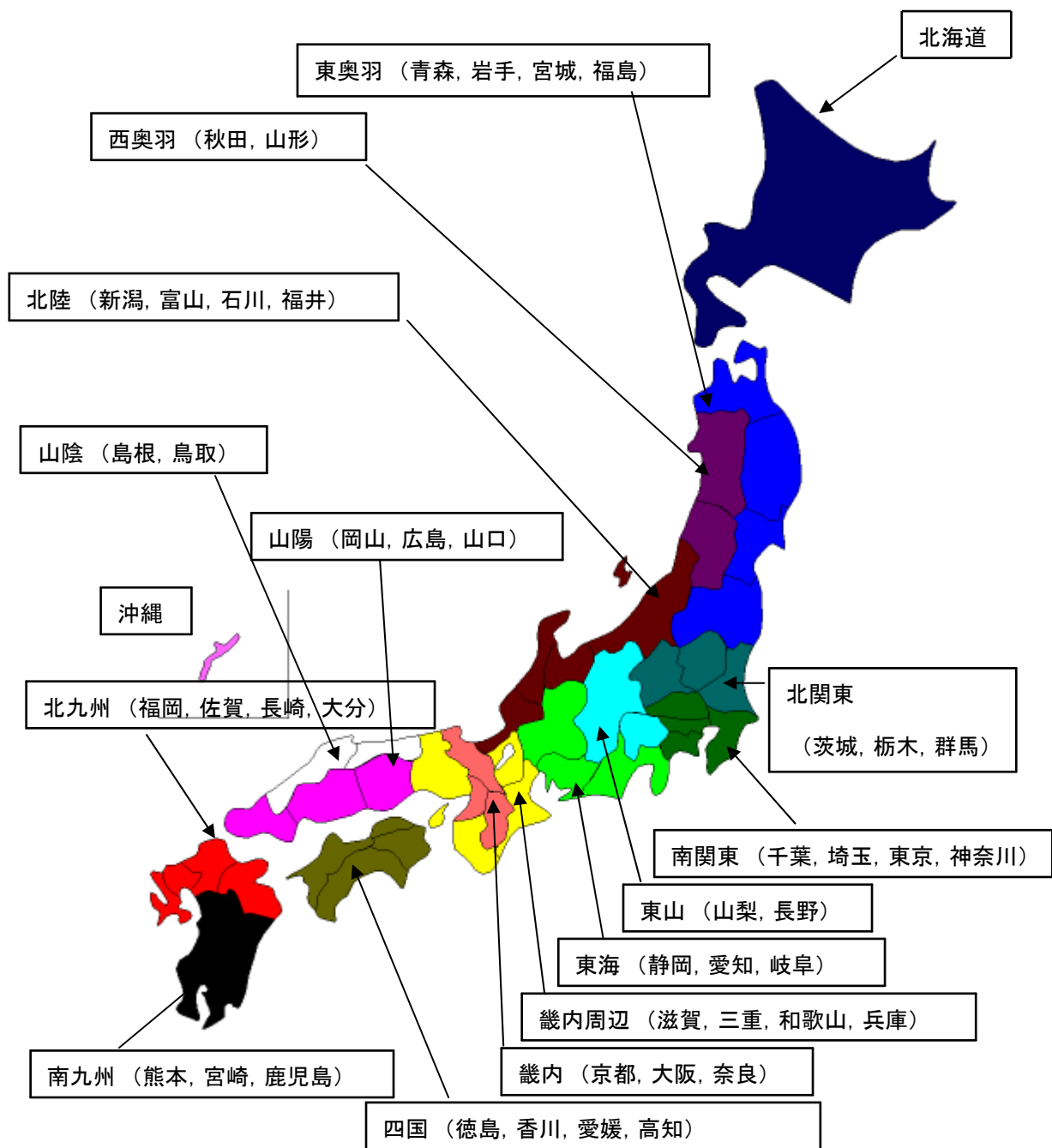


図 A2.1 16 地域の国別(府県別)構成

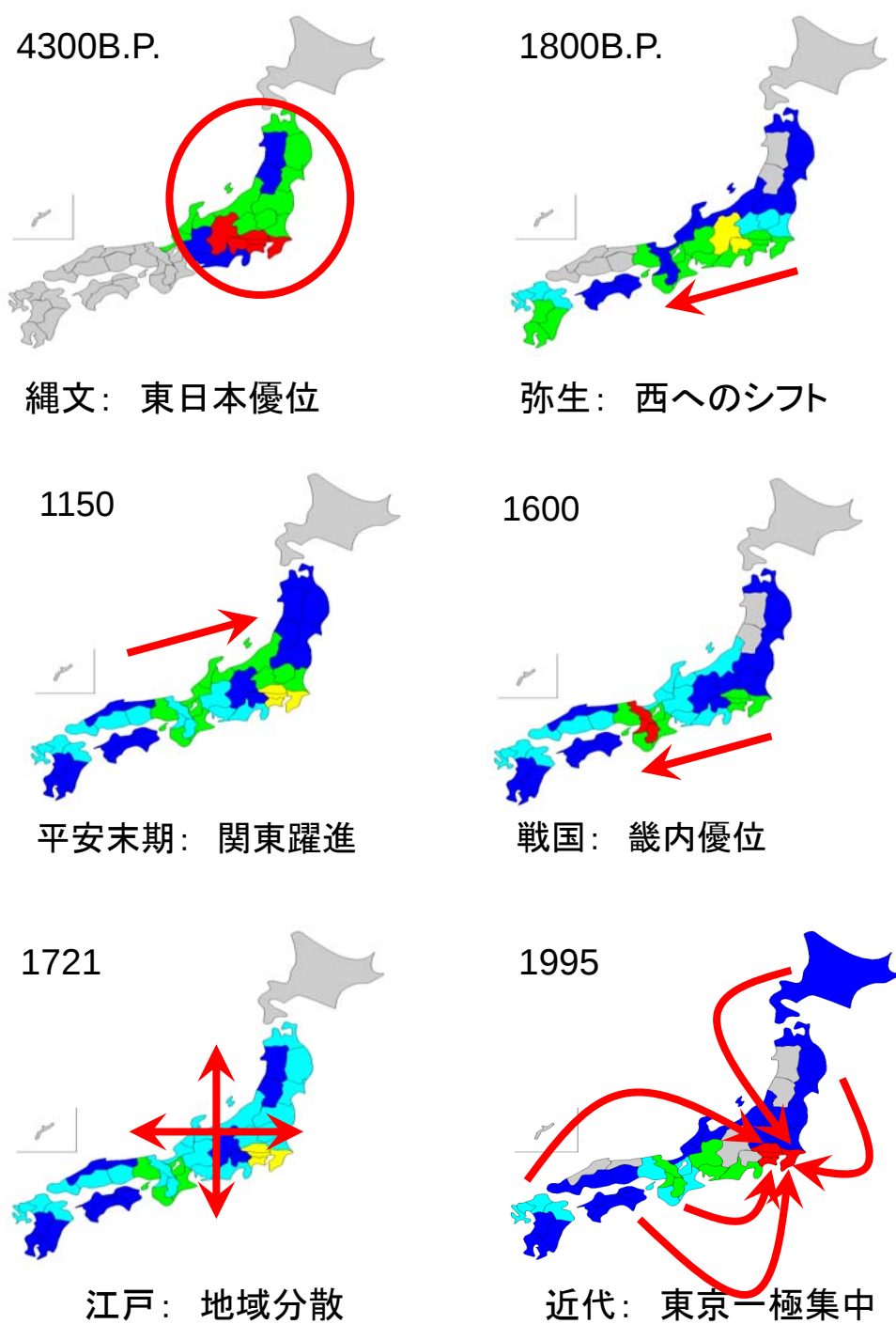


図 A2.2 地域人口分布の変遷⁶⁰

15%～:■, 12～15%:■, 9～12%:■, 6～9%:■, 3～6%:■, ～3%:■

⁶⁰ 人口から読む日本の歴史[鬼頭 2005]の p16,17 のデータをもとに作成.

付録3 人口移動ネットワークの可視化

地域間では, さまざまなフロー(流れ)が, さまざまなスケールにおいて展開している. 資本・財・サービス・情報などの移動は, このフローとしてとらえられ, これらが流通・通信などを通じて流れている. 人口移動(migration)も, こういったさまざまなフローのなかの一つとしての性格をもっている.

人口移動には, 空間的な移動である地域移動と職業間・社会階層間などを移動する社会移動があるが, 本研究では前者を指す. 人口移動の分析を行うためには, ある期間内に観測された移動数(量)と移動の方向が必要である. このようなデータは, 都道府県間の人口移動データについて記載された住民基本台帳人口移動報告書しかないので, 人口移動の分析は都道府県のみとする. 人口移動のデータは, 行に出発地(origin), 列に到着地(destination)を配置した origin-destination 行列(以下, OD 行列と呼ぶ)として提示される. 都道府県間の人口移動の場合, 47×47 の行列で示される. 出発地を i , 到着地を j とすると, i から j への移動流は, M_{ij} と表現できる.

OD 行列は, 転入数と転出数が並んだだけの表であるが, この行列にさまざまな加工を加えることで, 人口移動現象を目に見える形にすることができる.

3.1 影響力の地図

OD 行列に条件付けをして色をつけると, 人口移動の構造が見えてくる. 図 A3.1 は, 表 A3.1 の条件付けにしたがって色をつけた都道府県人口移動の骨格である. 北海道から沖縄まで北から南の順で県の頭文字が並んでいる.

表 A3.1 色付けの条件

移動量	ランク
0～80	0
81～800	1
801～8000	2
8001～800000	3
800001～	4

4	北	青	岩	宮	秋	山	福	茨	群	埼	千	東	神	新	富	石	福	山	長	岐	静	愛	三	滋	京	大	兵	奈	和	鳥	岡	山	徳	香	愛	高	福	佐	長	能	大	宮	鹿	神					
北	0	2	2	2	2	1	2	2	2	1	2	2	3	3	2	1	1	1	1	1	2	2	1	1	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
青	2	0	2	2	2	1	1	1	1	1	2	2	2	2	1	1	1	0	0	1	1	1	1	0	1	1	1	0	0	0	0	0	1	1	0	0	0	0	1	0	1	0	0	0	0	0	1	1	1
岩	2	2	0	2	2	1	2	1	1	1	2	2	2	2	1	0	0	0	1	1	0	1	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	
宮	2	2	2	0	2	2	2	2	2	1	2	2	3	3	2	2	1	1	0	1	1	2	2	1	1	1	2	1	1	1	1	1	1	1	0	0	0	1	1	0	0	1	0	0	1	1	1		
秋	2	2	2	2	0	2	1	1	1	1	2	2	2	2	1	0	1	0	0	1	0	1	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0		
山	1	1	1	2	1	0	2	1	1	1	2	2	2	2	1	1	1	0	1	1	0	1	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
福	2	1	2	2	1	2	0	2	2	1	2	2	2	2	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	1		
茨	2	1	1	2	1	1	2	0	2	2	2	3	3	2	1	1	1	1	1	1	2	2	1	1	1	2	1	1	1	0	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1		
群	1	1	1	2	1	1	2	2	0	2	2	3	3	2	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	0	1	1	0	1	0	1	0	1	1	1	1			
埼	2	2	2	2	2	2	2	2	2	2	2	0	3	3	2	1	1	1	2	2	1	2	2	1	1	2	2	2	1	1	1	1	2	1	1	1	1	1	2	1	1	1	1	1	2	2			
千	2	2	2	2	2	2	2	2	2	2	3	0	3	3	2	1	1	2	2	1	2	2	1	1	2	2	2	1	1	1	1	2	2	2	1	1	1	2	2	1	1	2	1	1	2	2			
東	3	2	2	2	2	2	2	2	2	3	0	3	3	2	1	1	2	2	2	2	3	3	2	2	2	2	2	2	2	2	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2		
神	2	2	2	2	2	2	2	2	2	2	3	3	4	0	2	2	2	1	2	2	3	2	2	2	2	2	2	2	2	1	1	1	2	2	2	1	1	2	1	2	1	2	1	2	2	2			
新	2	1	1	2	1	1	1	1	2	2	3	2	0	2	1	1	1	2	1	1	2	1	1	1	1	1	1	0	0	0	1	1	1	0	1	0	0	1	0	0	0	0	0	1	1	1			
富	1	0	0	1	0	0	1	1	1	1	1	2	2	0	2	1	0	1	1	2	1	1	2	1	1	2	1	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	0	0	1			
石	1	1	0	1	1	0	1	1	1	1	2	2	2	1	2	0	2	0	1	1	1	2	1	1	2	2	1	0	0	1	1	1	0	1	1	0	1	1	0	1	0	1	0	1	1	1			
福	1	0	0	0	0	0	0	1	0	0	1	1	2	1	1	2	0	0	1	1	1	2	1	1	2	2	1	1	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0			
山	1	1	1	1	0	0	1	1	1	1	2	2	2	2	1	1	0	0	2	1	0	0	2	1	1	1	1	0	0	0	1	1	0	0	1	1	0	0	1	0	1	0	0	1	1	1			
長	2	1	1	1	1	1	1	1	1	2	2	2	2	2	1	1	1	2	0	2	3	2	1	1	2	2	1	1	1	0	0	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1			
岐	2	1	1	2	1	1	2	1	1	2	2	3	3	1	1	1	2	2	0	3	2	1	2	2	0	3	2	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1			
静	2	1	1	2	1	1	2	1	1	2	2	3	3	1	2	2	1	2	3	3	0	3	2	2	2	2	2	1	1	1	2	2	1	1	1	1	1	1	2	1	2	2	2	2	2	2			
愛	1	0	0	1	0	0	1	1	1	1	2	2	2	1	1	1	1	1	2	2	3	0	1	2	2	2	2	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
三	1	0	0	1	0	0	1	1	1	1	2	2	2	1	1	1	0	1	1	1	2	1	0	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1			
滋	2	1	1	1	1	1	1	1	1	2	2	2	2	1	1	2	2	1	1	2	2	1	0	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1			
京	2	1	1	2	1	1	2	1	1	2	3	3	1	2	2	2	1	2	2	2	3	0	3	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2			
大	2	1	1	1	1	1	1	1	1	2	2	3	2	1	1	1	1	1	1	2	2	2	3	0	2	2	2	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2			
兵	1	0	0	1	0	0	0	1	1	1	2	2	2	1	1	1	1	1	2	2	1	2	3	2	0	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1			
奈	1	0	0	0	0	0	0	1	0	0	1	1	2	1	0	0	0	0	0	0	1	1	1	1	2	2	1	0	0	2	2	2	1	1	1	1	1	0	0	0	0	0	0	0	0	0			
和	1	0	0	0	0	0	0	1	0	0	1	1	2	1	0	0	0	0	0	0	1	1	1	1	2	2	1	0	0	2	2	2	1	1	1	1	1	1	1	0	0	0	0	0	0	0			
鳥	1	0	0	0	0	0	0	1	0	0	1	1	2	1	0	0	0	0	0	0	1	1	1	1	2	2	1	0	0	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1			
岡	1	0	0	1	0	0	0	1	1	1	2	2	2	1	1	1	0	1	1	1	2	2	1	1	2	2	1	1	2	2	0	2	2	2	0	2	1	1	1	1	1	1	1	1	1	1			
山	1	1	0	1	0	0	1	1	1	1	2	2	2	1	1	1	1	1	1	1	2	1	1	2	2	1	1	2	2	1	1	2	2	0	1	1	1	1	2	1	2	1	1	1	1	1			
徳	1	0	0	0	0	0	0	1	0	0	1	1	2	2	0	0	0	0	0	0	1	1	1	1	2	2	1	1	0	1	1	1	0	1	1	0	2	2	1	1	0	1	0	1	1	1			
香	1	0	0	1	0	0	0	1	0	0	1	1	2	2	1	0	1	0	0	0	1	1	1	1	2	2	1	1	1	1	1	2	2	1	2	0	2	2	2	0	1	1	1	1	1				
愛	1	0	0	1	0	0	1	1	0	0	1	1	2	2	0	0	0	0	0	1	1	2	1	1	1	2	2	1	1	1	1	2	2	1	2	2	0	2	2	0	1	1	1	1	1				
高	2	1	1	1	1	1	2	1	1	2	2	3	2	1	1	1	1	1	2	2	1	1	2	2	2	1	1	1	1	1	2	2	2	1	1	2	1	0	2	3	3	2	2	2	2				
福	1	0	0	0	0	0	0	1	1	0	1	1	2	1	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0	1	1	1	0	0	0	1	0	3	0	2	2	1	1	1	1				
佐	1	1	0	0	0	0	0	1	0	1	1	2	2	2	0	0	0	0	0	0	1	1	1	1	2	2	1	0	1	1	1	2	1	1	1	1	1	3	2	0	2	2	1	2	1				
長	1	0	0	1	0	0	0	1	1	1	2	2	2	0	0	1	1	1	1	1	2	2	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	3	2	0	2	2	2	2	1				
能	1	0	0	0	0	0	0	1	1	1	1	2	2	0	0	0	0	0	0	0	1	1	1	1	2	2	1	0	0	1	1	1	1	1	1	1	1	3	1	2	2	0	2	2	1				
大	1	0	0	0	0	0	0	1	1	1	1	2	2	0	0	0	0	0	0	0	1	1	1	1	2	2	1	0	0	1	1	1	1	1	1	1	3	1	2	2	0	2	2	1					
宮	1																																																

この図から、都道府県間のつながりの強さが見えてくる。色が濃いほど深いつながりがあり、相手県に対して強い影響力を持っているので、移動関係の等高線と捉えることができる。

1. 東京、千葉、埼玉、神奈川を1つのブロックと考えると、このブロックは他のすべての県に対して影響力を持っている。
2. 同様に、京都、大阪、兵庫を1つのブロックと考えると、このブロックは以南の県に対して影響力を持っている。
3. 福岡は、九州のすべての県に対して影響力を持っている。
4. 対角線周辺が濃いことから、近県同士のつながりは強い。
5. 東北地方と関西以南のつながりは弱い

このように、図 A3.1 は、ある県同士の影響力について知るための地図として、役に立つ。

3.2 移動空間地図

人口の移動量が多い県同士ほど近くに存在すると考えてみると、実際の地図とは違う、移動空間での位置関係が見えてくる。図 A3.2, 3.3, 3.4, 3.5 は、それぞれ表 A3.2 の条件付けにしたがって都道府県間にリンクを張ったものである。図 A3.2 は 81 人以上①、図 A3.3 は 801 人以上②、図 A3.4 は 8001 人以上③の人口移動がある場合に、都道府県間にリンクを張っている。上側の図が移動空間地図、下側が実際の地図を表す。移動空間地図は、kamada-kawai アルゴリズム⁶¹[Kamada 1989]を用いて、

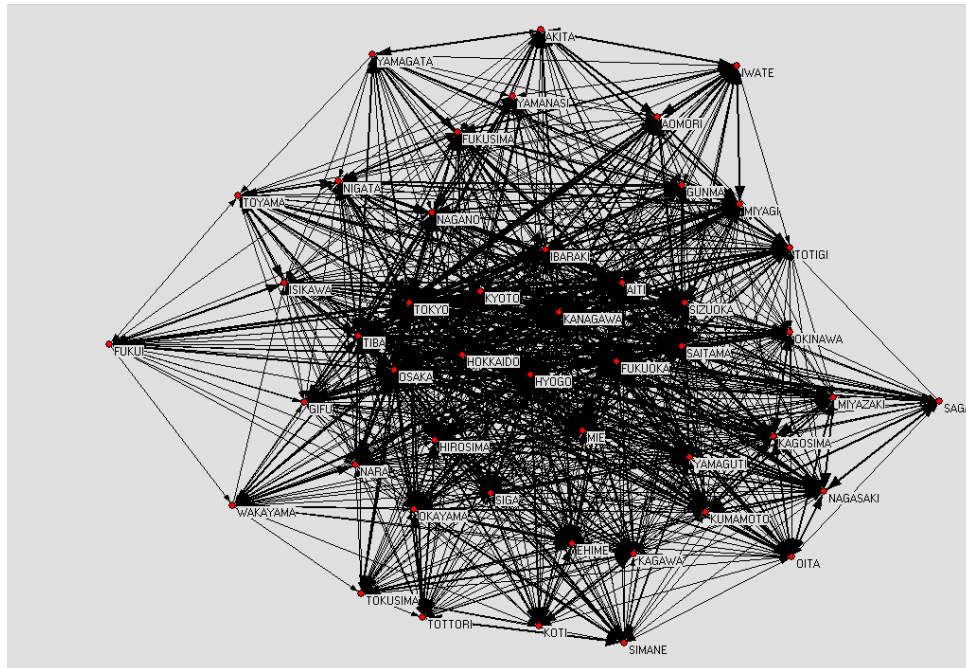
⁶¹ kamada-kawai アルゴリズムは、ノードがバネで結合されているというアナロジーを用いた、描画アルゴリズムである。バネはある長さのときエネルギーが最小になり、それより長くても短くてもエネルギーが大きくなる。エネルギーの総和が極小になるようにグラフを漸近的に変形していく。

県の位置の再配置を行ったものである. これに対し, 実際の県間地図の図は, 表 A3.2 の条件を満たしていれば, その県の県庁所在地同士にリンクを張ったものである.

表 A3.2 リンクを張るための条件

N0	表示条件 (移動人口)
①	81 人以上
②	801 人以上
③	8001 人以上

移動空間地図①



実際の地図①

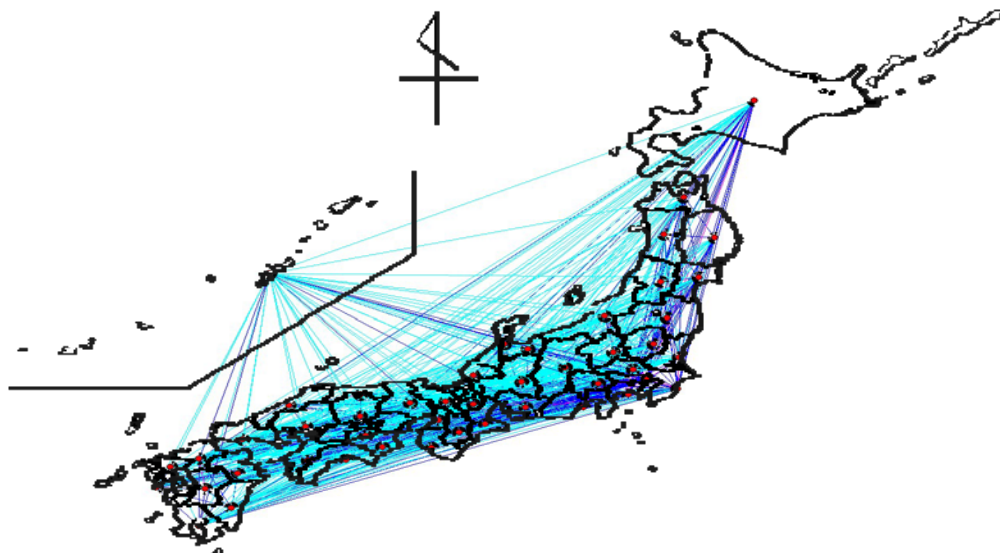
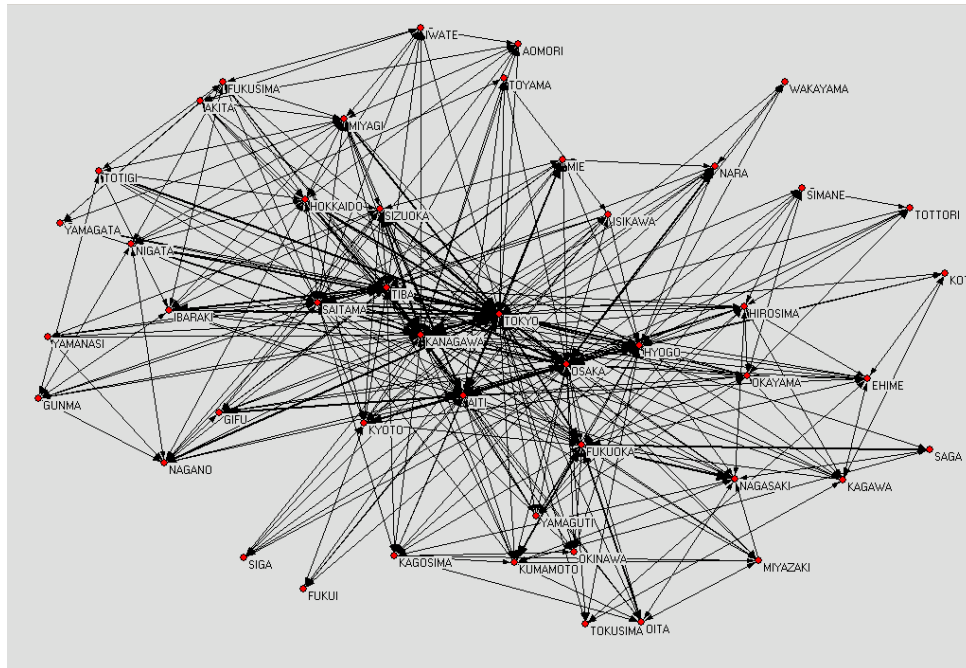


図 A3.2 移動空間地図と実際の地図①

移動空間距②



実際の地図②

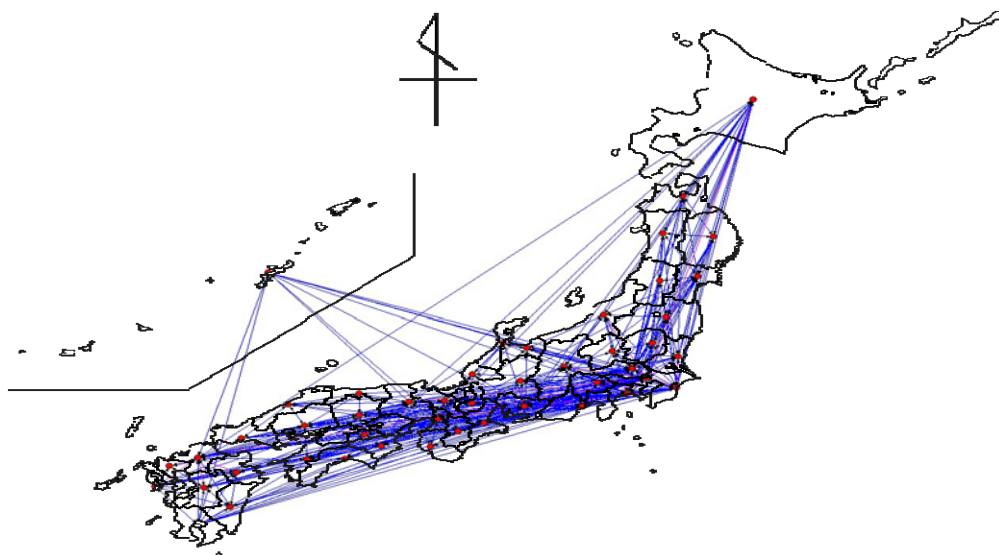
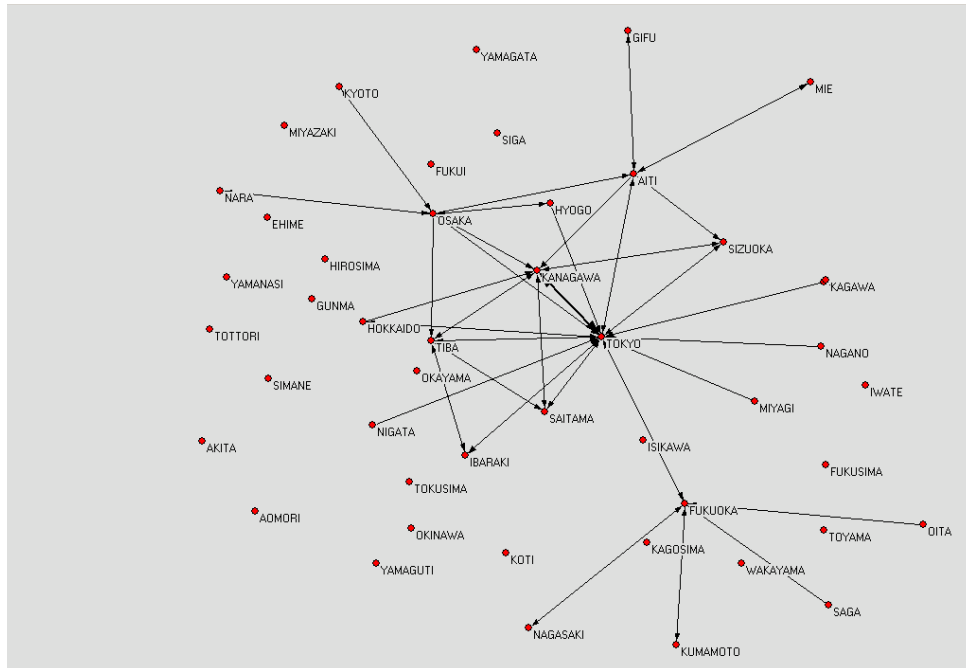


図 A3.3 移動空間地図と実際の地図②

移動空間地図③



実際の地図③

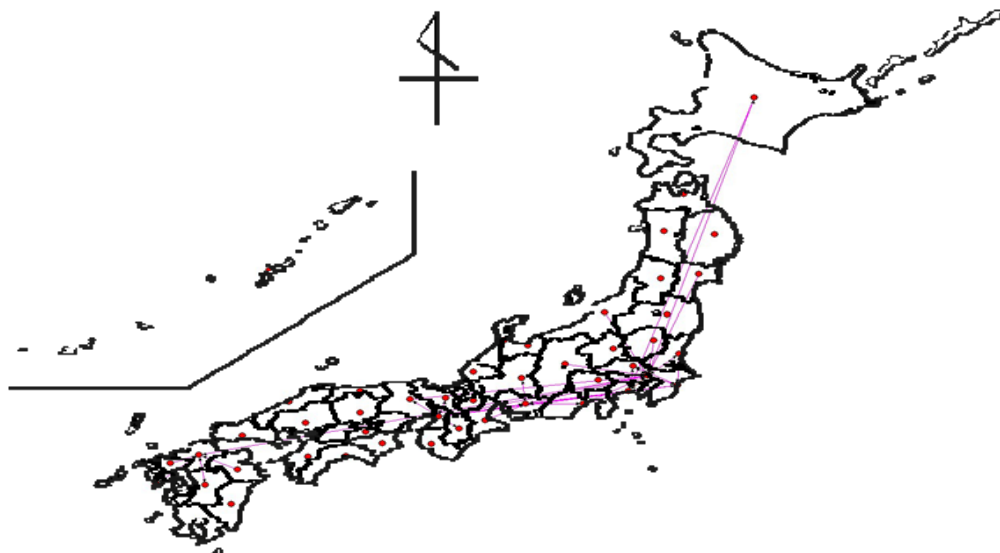
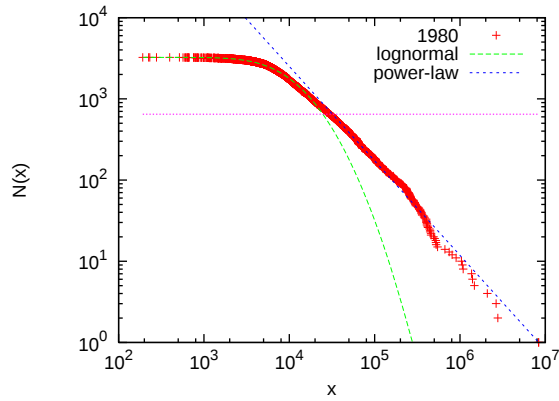


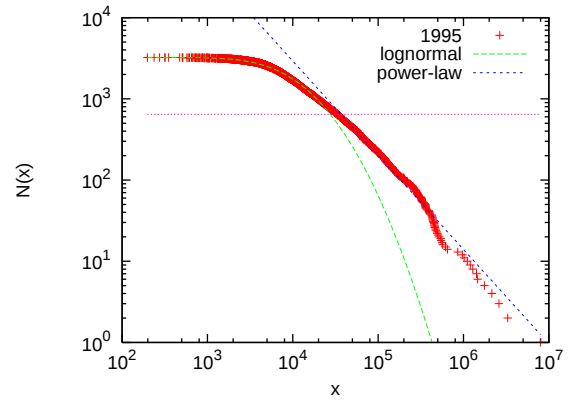
図 A3.4 移動空間地図と実際の地図③

付録 4 市町村のサイズ分布

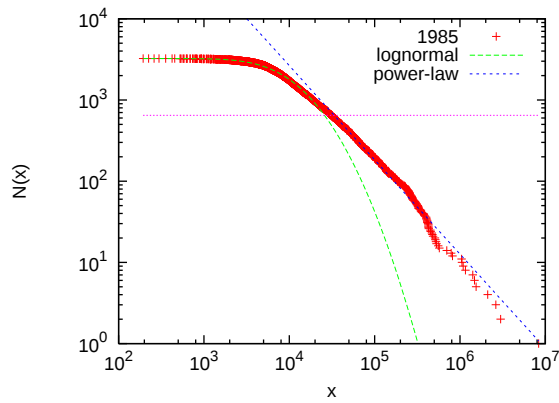
1980 年



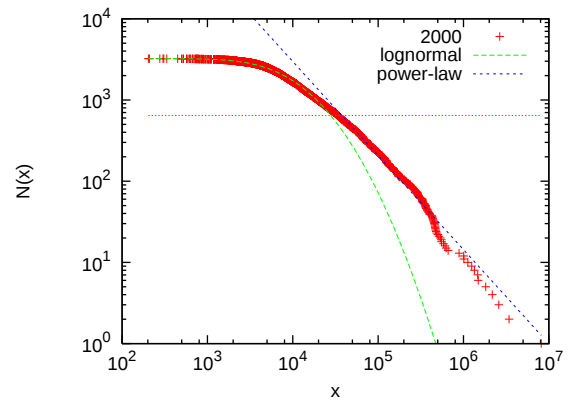
1995 年



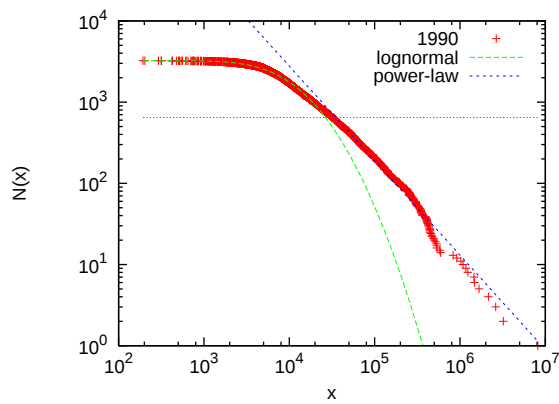
1985 年



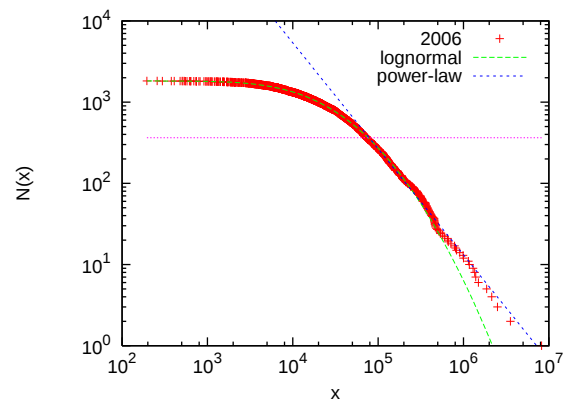
2000 年



1990 年



2006 年



謝 辞

本研究を行うにあたり、多くの方々にお世話になりました。この場を借りて、感謝の意を述べさせていただきたいと思います。

直接の御指導を頂いた指導教官である林幸雄准教授には、研究環境から学会発表の支援まで、さまざまな面でお世話になりました。また、研究者としての姿勢を背中で見せていただきました。心より深謝致します。

本論文の審査委員をお引き受けいただきました、中森義輝教授、吉田武稔教授、橋本敬准教授には、本論文を完成させるにあたり、さまざまな御助言を頂きました。そして、中央大の松下貢教授には、大変貴重なアドバイスを頂きました。深く感謝いたします。

防衛大の生天目章教授には、博士後期課程へ進むきっかけを頂くと共に、あたたく見守っていただきました。厚く御礼申し上げます。

情報科学研究科の岩垣剛助教には、本研究を進めるにあたり、有益な御助言と御鞭撻を賜りました。航空自衛隊の石西正幸さん、防衛大の佐藤浩准教授には、幾度となく励ましのお言葉をいただきました。また、本研究に関して幅広く議論していただきましたATRの相馬亘さん、中央大の國仲寛人さん、脇田順一さんには、心から感謝いたします。

最後に、いろいろな面においてあたたくサポートして頂いた両親、妻の早映子には、本論文の完成をもって謝意をあらわしたいと思います。

平成 20 年 3 月

富田 真治

参考文献

- [Abello 1998] J. Abello, A. Buchsbaum and J. Westbrook, “A Functional Approach to External Graph Algorithms,” in Proceedings of the 6th European Symposium on Algorithms(Springer, Berlin), 1998.
- [Auerbach 1913] F. Auerbach, “Das Gesetz Der Bevolkerungs konzentration,” Petermanns Geographische Mitteilungen, pp.73-76, 1913.
- [Bialas 1997] P. Bialas, Z. Burda and D. Johnston, “Condensation in the Backgammon model,” Nuclear Physics B, Vol. 493, 1997.
- [Borts 1965] G. H. Borts and J. L. Stein, 『Economic Growth in a Free Nuclear Physics B, Vol. 493, 1997Market』, Columbia Univ. Press, 1964. 中川久成, 坂下昇訳, 『地域経済の成長理論』, 勁草書房, 1965.
- [Broder 2000] A. Broder , R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins and J. Wiener, “Graph structure in the web,” Computer Networks 33, 309, 2000.
- [Bonabeau 1995] E. bonabeau and L. Dagorn, “Possible universality in the size distribution of fish schools,” Phys. Rev. E 51, R5220 - R5223, 1995.
- [Boyle 1999] P. Boyle and K. Halfacree, “Migration and gender in the deceloped world,” London: Routledge, 1999.
- [Buchanan 2005] M. Buchanan, 『複雑な世界, 単純な法則』, 草思社, pp.9-27, 2005.
- [Champernowne 1953] D. G. Champernowne, “A Model of Income Distribution,” Economic Journal, 63(250), pp. 318-475, 1953.
- [Champion 1989] A. G. Champion, “Counterurbanization: the changing pace and nature of population deconcentration,” London: Edward Arnold, pp.230-244, 1989.
- [Clauset 2007] A.Clauset, M. Young and K. S. Gleditsch, “On the Frequency of Severe Terrorist Events,” Journal of Conflict Resolution 51, 58, 2007.
- [Eaton 1997] J. Eaton and Z. Eckstein, “Cities and Growth: Theory and Evidence from France and Japan,” Regional Science and

- Urban Economics, 27(4-5), pp. 443-474, 1997.
- [Eechout 2004] J. Eechout, "Gibrat's Law for (All) Cities," AMERICAN ECONOMIC REVIEW, VOL. 94, NO.5, 1429-1451, 2004.
- [Fagan 1993] M. Fagan and C. F. Longin, "Migrating retirees: A source for economic development," Economic Development Quarterly 7: 98-106, 1993.
- [Fray 1988] W. H. Frey, "Migration and metropolitan decline in developed countries: a comparative study," Population and Development Review 14: 595-628, 1988.
- [Gabaix 1999] X. Gabaix, "Zipf's Law for Cities: An Explanation," The Quarterly Journal of Economics, 113(3), pp.739-767, 1999.
- [Gibrat 1931] R. Gibrat, "Les inégalités économiques (Economic Inequalities)," Paris, Sirey, 1931.
- [Glaeser 1995] E. Glaeser, J. Scheinkman, and A. Shleifer, "Economic Growth in a Cross-Section of Cities," Journal of Monetary Economics, XXXVI, 117-143, 1995.
- [Greenwood 1985] M. J. Greenwood, "Human migration: theory, models and empirical studies," Journal of Regional Science, 25, pp.251-544, 1985.
- [Hackett 1967] A. P. Hackett, 『70 Years of Best Sellers, 1895–1965』, (R. R. Bowker Company, New York), 1967.
- [Hauser 1959] P. Hauser and O.D. Duncan, 『The Study of Population: An Inventory and Appraisal』, University of Chicago Press, 1959.
- [Hirschman 1963] A. O. Hirschman, 『The Strategy of Economic Development』, Yale Univ. Press, 1958. 小島清監訳, 『経済発展の戦略』, 巖松堂, 1963.
- [Holme 2007] P. Holme, J. Karlin and S. Forrest, "Radial Structure of the Internet," Proceedings of the Royal Society A 463, 1231, 2007.
- [Huss 2006] M. Huss and P. Holme, "Currency and commodity metabolites: their identification and relation to the mod-ularity of metabolic networks," Preprint q-bio/0603038, 2006.
- [Illeris 1990] S. Illeris, "Counter-urbanization revisited: the new map of population distribution in central and north-western Europe," Norsk Geografisk Tidsskrift 44: 39-52, 1990.
- [Ishikawa 2001] Y. Ishikawa, "Migration turnarounds and schedule changes in

- Japan, Sweden and Canada,” *Review of Urban and Regional Development Studies* 13: 20-33, 2001.
- [Ito 2000] T. Ito, K. Tashiro, S. Muta, R. Ozawa, T. Chiba, M. Nishizawa, K. Yamamoto, S. Kuhara, Y. Sakaki, “Toward a protein-protein interaction map of the budding yeast: A comprehensive system to examine two-hybrid interactions in all possible combinations between the yeast proteins,” *Proc. Natl. Acad. Sci. USA*, 97:1143. 2000.
- [Kamada 1989] T. Kamada and S. Hawaii, “An Algorithm for Drawing Undirected Graphs,” *Information Processing Letters*, 31:7-5, 1989.
- [Kesten 1973] H. Kesten, “Random Difference Equations and Renewal Theory for Products of Random Matrices,” *Acta Mathematica*, CXXI 207-248, 1973.
- [Kobayashi 2006] N. Kobayashi, Y. Sasaki, O. Moriyama, S. Matsushita and M. Matsushita, “Empirical Studies of Random Multiplicative Stochastic Processes: Revisit to Lognormals,” *Nonlinear Phenomena in Complex Systems*, Vol.9, No.3 pp.276-282, 2006.
- [Levy 1996] M. Levy and S. Solomon, “Power laws are logarithmic Boltzman laws,” *Int. J. Mod. Phys. C*, Vol. 7, pp. 595-601, 1996.
- [Manrubia 1999] S. C. Manrubia and D. H. Zanette, “Stochastic multiplicative processes with reset events,” *Phys. Rev. E*, Vol. 59, pp. 4945-4948, 1999.
- [Mitzenmacher 2003] M. Mitzenmacher, “Dynamic Models for File Sizes and Double Pareto Distributions,” *Internet Mathematics* Vol. I, No. 3: 305-333, 2003.
- [Mera 1978] K. Mera, “Population Concentration and Regional Income Disparities: A Comparative Analysis of Japan and Korea,” in N. M. Hansen (ed.), *Human Settlement System*, Ballinger Publishing Co., pp.155-174, 1978.
- [Malthus 1978] T. R. Malthus, 『An essay on the principle of population, as it affects the future improvement of society』, London, 1978, 永井義雄訳, 『人口論』, 中央公論, 2006.
- [Myrdal 1959] K. G. Myrdal, 『Economic Theory and Under-developed

- Regions』, Gerald Duckworth & Co, 1957. 小原敬司訳, 『経済理論と低開発地域』, 東洋経済新報社, 1959.
- [Newman 2005] M. E. J. Newman, "Power laws, Pareto distributions and Zipf's law," *Contemporary Physics* 46, 323-351, 2005.
- [Ohkubo 2005] J. Ohkubo and M. Yasuda, "Preferential urn model and nongrowing complex networks," *Phys. Rev. E* 72(6), pp.1-4, 2005.
- [Plane 1989] D. A. Plane, "Population migration and economic restructuring in the United States," *International Regional Science Review* 12: 263-280, 1989.
- [Ritort 1995] F. Ritort, "Glassiness in a Model without Energy Barriers," *Phys. Rev. Lett.* 75, 1190 – 1193, 1995.
- [Redner 1998] S. Redner, "How popular is your paper," *European Physics Journal B* 4, 131, 1998.
- [Rogers 1997] A. Rogers, and S. Rajbhandary, "Period and cohort age patterns of US migration, 1948-1993: are American males migrating less?," *Population Research and Policy Review* 16: 513-530, 1997.
- [Rozman 1990] G. Rozman, "East Asian Urbanization in the Nineteenth Century," in Van der Woude et al., pp. 61-73, 1990.
- [Sasaki 2007] Y. Sasaki, H. Kuninaka, N. Kobayashi and M. Matsushita, "Characteristics of Population Distributions in Municipalities," *J. Phys. Soc. Jpn.*, Vol. 76, No. 7, 2007.
- [Sato 2004] A. Sato, "Explanation of power law behavior of autoregressive conditional duration processes based on the random multiplicative process," *Phys. Rev. E* 69, 047101, 2004.
- [Sornette 1997] D. Sornette and R. Cont, "Spontaneous scaling emergence in generic stochastic systems," *J. Phys.* Vol. 17, pp. 431-444, 1997.
- [Simon 1996] H. A. Simon, 『Models of My Life』, The MIT Press, p275, 1996.
- [Smith 2003] F. A. Smith, S. K. Lyons, S. K. M. Ernest, K. E. Jones, D. M. Kaufman, T. Dayan, P. A. Marquet, J. H. Brown and J. P. Haskell, "Body mass of late Quaternary mammals," *Ecology* 84, 3403, 2003.

- [Streib 2001] G. F. Streib, “Retirement community,” In the encyclopedia of aging (3rd edition), eds. G. L. Maddox, R. C. Atchley, J. G. Evans and R. B. Hidson, 890-893, New York: Springer Publishing Company, 2001.
- [Takayasu 1997] H. Takayasu, A. Sato and M. Takayasu, “Stable Infinite Variance Fluctuations in Randomly Amplified Langevin Systems,” Phys. Rev. Lett. 79(6), 1997.
- [Takayasu 1986] H. Takayasu and I. Nishikawa, “Proceedings of the First International Symposium for Science on Form,” KTK Science Publisher, Tokyo 1986.
- [United Nations 1998] United Nations, 『World Population Prospects』, The 1998 Revision, Vol.1, New York, 1998.
- [Vining 1982] D. A. Vining and R. Pallone, “Migration between core and peripheral regions: a description and tentative explanation of the patterns in 22 countries,” Geoforum 13: 339-410, 1982.
- [Watts 2004] D. J. Watts, 『スモールワールド・ネットワーク』, 阪急コミュニケーションズ, pp24, 2004.
- [Willinger 1998] W. Willinger and V. Paxson, “Where Mathematics Meets the Internet,” Notices of the AMS, Vol 45, No.8, P961-970, 1998.
- [Zanette 1997] D. Zanette and S. Manrubia, “Role of intermittency in urban development: a model of large-scale city formation,” Phys. Rev. Lett. 79: 523-526, 1997.
- [Zipf 1949] G. Zipf, “Human Behavior and the Principle of Least Effort,” Cambridge, MA: Addison-Wesley, 1949.
- [青山 2007] 青山秀明, 家富洋, 池田裕一, 相馬亘, 藤原義久, 『プレートファームズ』, 日本経済評論社, 2007.
- [安達 1970] 安達生恒, “過疎の実態: 過疎とは何か, そこで何がおきているか,” 『ジュリスト』, 有斐閣, 455 号, 1970.
- [阿藤 2004] 阿藤誠, 早瀬保子編, 『ジェンダーと人口問題』, 原書房, 2004.
- [石川 2001] 石川義孝, 『人口移動転換の研究』, 京都大学学術出版会, pp.29-36, 2001.
- [大淵 1987] 大淵寛, “人口学の対象と方法,” 南亮三郎, 上田正夫編, 『人口学の方法』, 千倉書房, 1987.
- [兼清 2002] 兼清弘之, “9. 科学としての人口学,” 日本人口学会編,

- 『人口大事典』, 培風館, 316-320, 2002.
- [鬼頭 2005] 鬼頭宏, 『人口から読む日本の歴史』, 講談社学術文庫, 2005.
- [久保 2003] 久保亮五, 『統計力学』, 共立出版, 2003.
- [小山 1984] 小山修三, 『縄文時代 コンピュータ考古学による復元』, 中央公論社, 1984.
- [斎野 1987] 斎野岳郎, “わが国の人口移動圏とその変化—1971～1980年—,” 東北地理 39: 34-35, 1987.
- [酒井 1991] 酒井高正, “都道府県間人口移動の規定要因の分析,” 奈良大学紀要 19: 113-130, 1991.
- [菅沼 2005] 菅沼栄一郎, 『村が消えた—平成大合併とは何だったのか』, 祥伝社新書, 2005.
- [相馬 2006] 相馬亘, “経済物理における乗算的確率過程 Multiplicative Stochastic Process in Econophysics,” 情報論的学習理論ワークショップ, 2006
- [舘 1960] 舘稔, 『形式人口学』, 古今書院, 1960.
- [田原 2007] 田原裕子, “合衆国におけるリタイアメントコミュニティ産業の発展—デル・ウェップのサンシティ・アリゾナを中心に,” 國學院経済学 55 (2): 105-126, 2007.
- [丹保 2004] 丹保憲仁, 『人口減少下の社会資本整備 拡大から縮小への処方箋』, 土木学会, 2004.
- [林 2007] 林幸雄, 『ネットワーク科学の道具箱』, 近代科学社, 2007.
- [早瀬 2002] 早瀬保子, 『途上国の人口移動とジェンダー』, 明石書店, 2002.
- [松下 2002] 松下貢, 『フラクタルの物理(I)』, 裳華房, 2002.
- [南 1964] 南亮三郎, 『人口学総論—人口原理の研究』, 千原書房, 1964.
- [向井 2007] 向井正, “惑星間塵の研究の現状と課題”, エアロゾル研究 Vol. 22, No. 2 pp.119-124, 2007.
- [渡辺 1978] 渡辺良雄, “大都市居住と都市内部人口移動,” 総合都市研究 4: 11-35, 1978.
- [矢野 1996] 矢野桂司, “1980 年代後半の東京大都市圏における都市内部人口移動,” 総合都市研究 59: 35-47, 1996.
- [山口 1989] 山口喜一, 『人口分析入門』, 古今書院, 1989.

研究業績

[学術論文]

- [1] S, Tomita and Y. Hayashi, “A Controllable Model of Random Multiplicative Process for Entire Distribution of Population,” *Physica A*, Vol. 387 1345–1351 2008.

[国際会議]

- [2] S, Tomita and Y. Hayashi, “Spatial analysis of centralization and decentralization in the population migration network,” *Proc. of INFORMATION VISUALIZATION 2006*, pp139-142, 2006.
- [3] S, Tomita and Y. Hayashi, “Controllable Stochastic Models for Entire Distribution of Population,” *Proc. of the 11th Asia-Pacific Workshop on Intelligent and Evolutionary Systems*, 2007.

[国内研究会]

- [4] 富田真治, 林幸雄, “集中と分散の空間構造分析 —人口移動ネットワークから—,” ネットワーク生態学研究会 第1回サマースクール
- [5] 富田真治, 林幸雄, “県間人口移動のネットワーク分析,” ネットワーク生態学 第2回シンポジウム
- [6] 富田真治, 林幸雄, “確率モデル制御による人口分布の再現,” ネットワーク生態学研究会 第3回サマースクール