広島大学総合科学部紀要IV理系編，第 28 巻，109 － 126 ページ，2002 年 12 月

回帰偏差値の考案とその効用：
地域間格差を相対評価する偏差値

日下部　眞一

広島大学総合科学部

The concept and usefulness of the regression-based score and its application:
A new index, “the regression-based score (R-score)” to measure the relative size of the local statistics

Shinichi KUSAKABE

Faculty of Integrated Arts and Sciences, Hiroshima University, 1-7-1,
Kagamiyama, Higashi-Hiroshima 739-8521, Japan

Abstract : Although the per capita statistics is usually and customarily used in the analysis on the relative evaluation of the statistics of the prefecture, these measures contain an artificial flaw due to the nonlinearity between the population number and the size of the statistics in question. For the purpose of overcoming these problems on nonlinearity, the regression-based score (R-score) was originated as a new indicator to measure the relative size of the statistics of the prefecture. This index is the application of the Studentized residuals, calculated based on the regression analysis. This index seems to be very useful in comparative evaluation on the size of the various statistics among the prefectures. The size of the various statistics of the prefecture was analyzed with this regression-based score (R-score), and the usefulness and the power of the new index was critically evaluated.

Keywords : Regression analysis, Regression-based score (R-score), Studentized residual.

１．はじめに

　都道府県の統計指標（たとえば『統計でみる県のすがた』など）を眺めてみると、かなり多くの指標（たとえば病院数、県民所得、土木費ほか）が人口当たりの数値で表されている。そして指標値と同時に、その人口当たりの数値の大小によって順位付けされている。例えば『統計でみる県のすがた』にのっている約 450 項目の指標値では、人口の最も少ない鳥取、島根、高知の３県のいずれかが１位か最下位の 47 位になっている場合がかなり多い（455 指標項目の内 95 項目）ことに気づかされる。もちろん、この現象が真実を表しているのかもしれないが、しかし、単なる人為的な統計操作によってこのような現象がでてくる可能性が十分考えられるし、もし、そうであれば、人口当たりの数値で表して指標の比較を行って相互評価する時には、人口の大小に依存した誤った評価を行って

2002 年 10 月１日受付；2002 年 11 月１日受理

いることになり、これは社会的にも大きな問題である。

ここ数年の行政改革では行政が行う様々な事業の評価が取り上げられている。そして、その評価は新しい事業企画へ考慮されることになる。この時の行政単位は、ふつう都道府県が行政単位であり、都道府県単位でさまざまな地域間比較が、できるだけ客観的な基準で行われることがのぞましい。こういう観点に立って都道府県のさまざまな統計指標をながめてみると地域比較という点においてはさらに大きな改良がなされるべきであるように思われる（大友　1997)。

筆者は、特定非営利活動促進法（NPO法）施行後の各都道府県認証ＮＰＯ数の地域間変異を考察し、各地域のNPOセクターの規模を客観的に相対評価できる“回帰偏差値”という評価値を考案した（日下部 2002)。これは、統計学的には回帰分析における残差分析において Studentized residual（スチューデント化残差）として理論化されている偏差であり、筆者のいう回帰偏差値は、このスチューデント化残差を10倍して50を加えた値である。この回帰偏差値は、通常の偏差値、つまり正規分布を基礎にした正規偏差値同様、きわめて有用であり、相対評価として有力な手法であると考えられるので、ここにその手法と適用例を詳細に検討して解説したい。

2．地域指標の大きさを相対評価するにあたっての課題

従来、慣習的に一人当たりの数値で様々な指標の大小の順位付けを行ってきた相互評価は、統計的には実測数で表現して、

Y（指標値）=a * X（人口数）

の式において、“さまざまな観測値aを比較する値”として相互評価しているわけで、“人口数に依存しない値を同等とすること”を暗黙の内に仮定していることを意味している。つまり100人に対する10個も1000人に対する100個も10000万人に対する1000個も、人口当たりの数値で行う相互評価では同じ0.1という評価になる。

いいかえれば、“人口でわった数値を比較分析の基準とし、それらの指標値を用いて相互評価を行うこと”は、実は、“人口と指標値の間に線形性が保証されている”ことを暗黙の内に想定しているのである。したがって基本的には、この仮定、“人口要因を独立変数としたときに指標値と人口との間に線形性が成り立つのか”、つまり“従属変数を指標値、独立変数を人口としたときに１次回帰直線の傾きが一定数であることを仮定してよいかどうか”を検討しなければならない。

しかし社会現象には“経験的”に、この傾きが一定でなく上に凸の指数関数的分布を示したり、下に凸の指数関数的分布を示したりする場合が多く見られる。例えば、経済学における生産要素に関する収穫逓減の法則は一般的に上に凸の指数関数で表されることが良く知られている。もし上に凸であれば、このような状況で人口当たりの数値でランキングや相互評価を行えば、人口が小さい地域が過大評価され、人口が大きい地域が過小評価されることになる。先に述べた、都道府県のさまざまな指標のランキングで人口が小さい県が上位によく並ぶのはこのような過大評価の効果がでてきてしまうためである。

本論は、都道府県ランキングなどの相対評価を行おうとする場合には、正であれ負であれこのような規模の効果を出来るだけ取り除いて相互評価をしたほうが指標値のより客観的な相対評価になると考え、このような影響をできるだけ克服する相対評価の指数、“回帰偏差値”を考案し、これを用いて統計指標値の相対評価を試みることを目的としている。

まずはじめにこのような課題を解決できる方法として回帰偏差値の手法を述べる。次に、正の規模の効果（効用逓減）を示す場合、規模の効果が見られない場合、負の規模の効果（効用逓増）を示す

場合の典型的な３つの例を用いて回帰偏差値の手法の有効性を示す。その後、自然科学の例として動物の脳の重量の進化の問題を、経済学の例として最近話題となっている公共事業費の地域格差の問題を具体的な回帰偏差値の適用例として検討する。

3．回帰偏差値の考案

3.1　規模の効果をしめす非線形性

１で述べた問題点を理解しやすいように図で示す。例として、あとでも詳しく解析する郵便局数を考える。図１は各都道府県の郵便局数を、横軸を各都道府県人口実数、縦軸を郵便局数で表した。中ほどに大きく分布からはずれた北海道の点が見られるが、分布は上に凸のゆるやかな指数関数的関係にあるように見える。つまり、人口と郵便局数の分布は緩やかではあるが非線形で、上に凸の指数関数的関係にある。

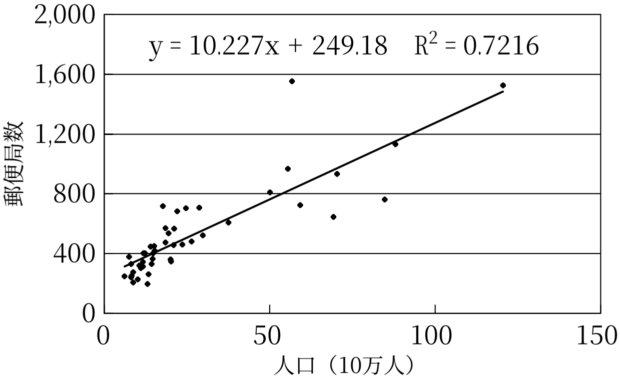

図１　各都道府県の郵便局数を、横軸に人口実数、縦軸に郵便局数で表した分布

慣用的に用いられてきた“ 人口当たりの郵便局数 ”というのは、この図１で各観測点と原点を結ぶ直線の傾きを意味している。したがって、このように分布に非線形性が見られるときに“ 人口当たりの数値 ”で相互評価を行えば、分布が上に凸の指数関数的分布を示すような時には、人口の小さい観測点が順位の上位を占め、逆に、下に凸の指数関数的分布を示すような場合は、人口の大きい観測点が順位の上位を占めるような傾向を示すことになる。

人口当たりの郵便局数をもとに順位付けして都道府県の相互評価をすることは、人口と郵便局数に見られる図１のような分布の非線形性をこみにして評価していることになる。しかし、地域間の郵便局数についての相対評価を行いたい場合には、この非線形効果を除いて相互評価を行なったほうが相対評価としてはよいというのが本論の主張で、これを除いて比較する方法が“ 回帰偏差値 ”の手法である。

まず、理解しやすいように今まで用いられてきた、人口当たりの郵便局数をプロットしてみる。図３は、横軸を人口（10万人）、縦軸を人口（10万人）で郵便局数をわった値として 47 都道府県の値の分布を示したものである。

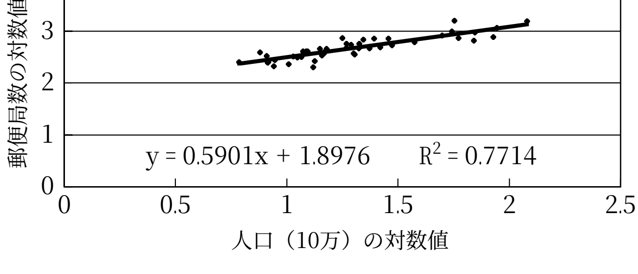

図２　各都道府県の郵便局数を、横軸に人口実数、縦軸に人口（10 万）あたりの郵便局数で表した分布

この図３を見て気づかされるのは、まず、観察点の分布の不均一性である。したがって、人口当たりの数値でもって相互評価を行っているということは、このような分布の不均一性、統計学のことばでいえば分散の不均一性を考慮

していないことにまず注意しておく必要がある。次に、統計的に有意となる直線回帰性（この場合は負の相関）が認められる。実際に直線回帰を行い回帰直線の傾きを求めると -0.2323 となる。この傾きの０からの統計的有意性を検定するとｔ値 (自由度は 45) は 6.04 となり高度に有意である。相関の度合いを示す相関係数の値は、0.4477 であり統計的に高度に有意で（t = 3.25, 自由度 45)、統計的に有意な正の規模の効果が見られるのである。したがって、このような分布を示す場合に、人口当たりの数値の大小でもって順位付けする場合には、当然、図３における人口約 300 万以下の小さい地域が上位を占めるという結果になってしまう。

それでは、どのようにしてこのような分散の不均一性を除き、非線形的効果を除くことが可能であろうか。これらを克服するために、まず、両対数変換を行う。図１と同じデータを両対数変換して示したのが図３である。対数は実測値との対応が容易なように 10 を底とする対数を用いた。人口数は 10 万人あたりで表している。

対数変換により分散は安定化し、回帰直線の周辺における点の分布が正規的になり等分散性がたもたれる。この図における回帰直線の傾き 0.5901 は t = 8.56（自由度 45）で、統計的に高度に有意に１とは異なっている。この傾きがもし１であれば、両対数変換しないときに線形関係であることを意味しており、対数変換後の傾きが１より大きいということは両対数変換しないときに下に凸の指数関数的関係があることを意味している。この傾きが１からずれるにつれて非線形性の効果は大きい。

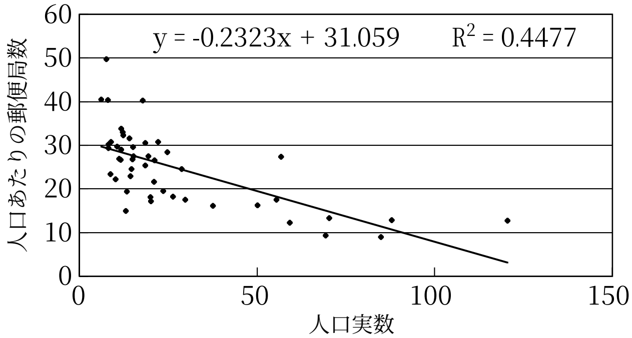

図３　各都道府県の郵便局数と人口（10 万人）の両対数変換値の分布図

3.2　回帰偏差値の概念と手法

対数変換後の郵便局数と人口の分布は、観測点間の距離は実数の分布に比べると変わっているが、観測点相互の位置関係は保たれたままである。したがって、分布としては同型で線形近似化され分散が安定化されたことになる。それでは、どのようにしたらこれらの観測点を相対評価に耐え得る値にできるであろうか？それは、両対数変換値に引いた回帰直線を基準点として考えることである。つまり回帰直線を基準として、回帰直線による予測値から各観測点間の距離、残差（偏差）を相対評価の値として用いることである。しかし、統計学的には分布の中心である平均値からずれるにつれて分散が大きくなることが知られている。このような分布を、分布の同型性を保ったまま平均０、分散１の分布に変換するのが影響力係数を用いた変換で、統計学的には“残差のスチューデント化（studentized residual）”として知られている（Myers 1990, Cook and Weisberg 1994, 1999）。

直感的にわかりやすく表現すると図３の分布図で、回帰直線がｘ軸になるように、やや右回転させ回帰直線上の横長い楕円状の観測点の分布をｘ（人口）の平均値上に投影して、その分布が分散１、平均０になるように押し縮めてやる変換であると考えればよい。分布の同型性を保ったままの変換であるから厳密には正規分布とはいえないのであるが、標本数が多いと独立性の欠如を無視しても差し支えないので、たがいに独立に正規分布に従うとみなしてもよい（Chatterjee and Price 1977）といわれている。このスチューデント化残差は尺度フリーで、標準化しない残差よりも客観的な残差分析に適していると評価されていて今日では広く普及し始めている（Myers 1990）。

3.3　回帰偏差値をもとめる手順

新しい指数を考案した基本的な考えは、人口規模から期待される郵便局数からの現実の郵便局数の残差を客観的に標準化することにある。そして、この期待される郵便局数というのが人口への１次回帰によって説明される値、つまり１次回帰式によって推定される予測値である。標準化の方法は、回帰分析で得られた標本回帰直線からの残差 (d_{YX}) を、回帰分散分析表の不偏分散の推定値の平方根 (s) と影響力係数 (leverage coefficient) (h) を用いた値で基準化する。影響力係数 (leverage coefficient) (h) と基準化した偏差 (*Stdev*) は、以下の式で求められる。

$$h_i = 1/n + \left(X_i - \overline{X}\right)^2 / \sum x^2$$

$$Stdev = d_{YX} / \left(s_{YX}\sqrt{1-h_i}\right)$$

h_i ：都道府県(i)の影響力係数（ハット行列の対角要素）

X_i ：都道府県(i)の人口の対数値（独立変数）

$\overline{X}$ ：都道府県人口の対数値（独立変数）の平均値

$x = X_i - \overline{X}$

n : 都道府県数（標本数）

$d_{YX} = Y - \hat{Y}$　(回帰直線からの偏差)

Y : 郵便局の対数値（従属変数）

$\hat{Y}$ ：回帰直線にもとづく Y の推定値

$s_{YX} = \sqrt{\sum d_{YX}^2 / (n-2)}$ ：回帰による誤差分散の不偏推定量の平方根

この方法については、例えば Sokal and Rohlf 1973（日本語訳 311 ページ）や Myers (1990)、Cook and Weisberg (1999) を参照するとよい。

通常の正規分布に基づく偏差値が一般に流布しているのと同じように、平均50、分散10の偏差値で表した方が統計学にあまり通じていない一般にはわかりやすいと考えられる。したがって、この標準化残差 (*Stdev*; Studentized residual) を 10 倍し 50 を加えると、得られる値は、平均 50、分散 10 の統計量になる。このような量を従来の正規分布に基づいて導かれた偏差値、すなわち“ 正規偏差値 ”に対して、回帰分析によって導かれた偏差値という意味で“ 回帰偏差値 ” (R-score) と呼ぶことにする。

４．回帰偏差値による解析と結果

『統計でみる県のすがた』には約450項目の統計指標が載せられている。これらの中から１３指標を選び、都道府県人口とそれら指標を両対数変換し単回帰分析を行った。この回帰直線の傾きが１より大きいときは、実数で表した両者の関係が規模の効果としては下に凸の指数関数的関係（規模の効用逓増）を意味しており、逆に、１より小さいときは、実数で表した両者の関係が規模の効果としては上に凸の指数関数的関係（規模の効用逓減）を意味している。傾きがほぼ１の場合は、実数で表した両者の関係に規模の効果が見られず、直線関係であることを意味している。これらを表す解析結果を表１にしめす。この表１の４列目には回帰直線の傾き（b）の傾き１からの有意性、５列目には回帰直線の傾き（b）の傾き０からの有意性をｔ値で表している。

この表の最初の５項目は、実数で表した指標と人口の関係が規模の効果としては上に凸（効用逓減）を意味している。次の４項目は規模の効果が見られない場合で、回帰直線の傾きはほぼ１である。最後の４項目は、実数で表した指標と人口の関係が規模の効果としては下に凸（効用逓増）を意味していて、回帰直線の傾きは１より有意に大きい。老人ホーム数は老人ホーム数と65歳以上人口との関係で解析したが、これ以外は全て問題指標と都道府県の総人口との関係で解析した。

表１　都道府県の統計指標値の規模の効果

	切片：a	傾き：b	傾きbの標準誤差	t値(傾きbの1からの有意性)	t値(傾きbの0からの有意性)	AIC
（1）回帰直線の傾きが１から有意に小さい指標						
郵便局数	1.8976	0.5901	0.0479	-8.5584	12.3218	-212.62
GSスタンド数	2.1870	0.6348	0.0458	-7.9662	13.8469	-216.73
救急車数	1.1012	0.6848	0.0393	-8.0185	17.4211	-231.19
病院数	1.1409	0.7711	0.0673	-3.3992	11.4538	-180.60
老人ホーム数	0.0583	0.7840	0.0477	-4.5250	16.4207	-226.76
（2）回帰直線の傾きがほぼ１である指標						
薬局数	1.5437	0.9988	0.0378	-0.0328	26.4336	-234.91
電力量	1.8013	1.0182	0.0368	0.4954	27.6681	-237.38
火災死傷者数	1.7614	1.0493	0.0391	1.2587	26.8077	-231.59
交通事故数	1.7447	1.0730	0.0475	1.5384	22.6021	-213.44
（3）回帰直線の傾きが１より有意に大きい指標						
書籍額	3.0935	1.1124	0.0385	2.9232	28.9304	-233.26
旅券数	1.3370	1.1945	0.0413	4.7091	28.9163	-226.52
郵便物数	3.7213	1.2708	0.0600	4.5101	21.1656	-191.37
ガス	0.5540	1.9773	0.1180	8.2838	16.7596	-127.87

これら13項目の中から規模の効果が効用逓減の場合として郵便局数、規模の効果が見られない場合として薬局数、そして、効用逓増の場合として一般旅券発行数の場合を選んで詳細な解析を行い、回帰偏差値の効用と有効性を確認していく。

4.1　郵便局数の変異

郵便局数と人口との関係を実数値で表現すると図１のようになる。図中の点の分布で一つ大きく離れているのは北海道を示す点である。この図からは郵便局数と人口との傾向はつかみにくいが、図３に示したように縦軸を人口あたりの郵便局数で表すと傾向がはっきりしてくる。人口あたりの郵便局数は人口と明らかに負の相関を示す。図３で直線回帰をすると傾きは、-0.2323となり、統計的に高度に有意に（t = 6.04）傾きは０とは異なる。これは図２に示すように両対数変換すると直線回帰の傾きが１より小さいことを意味する。

3.3で述べた手法に基づいて計算した回帰偏差値を表２に示す。偏差値が大きい方に 1* σ以上離れているのは、大きい方からいうと北海道、鹿児島、島根、長野、新潟、熊本の順である。逆に、偏差値が小さい方に 1* σ以上離れているのは、小さい方からいうと沖縄、埼玉、神奈川、滋賀、佐賀、

香川、群馬、栃木の順である。

通常、地域比較に用いられる人口当たりの郵便局数を表２の３列目に表している。この数値は図３の縦軸と同じで、図３の点の分布を縦軸のみで表したのが、これらの数値に相当する。これら数値の大小を順位づけると回帰偏差値の順位とは大きく異なって表２の４列目に示したようになる。図１では大きく分布から離れていた北海道は、この順位付けでは20位となり分布からのずれを全く気づかせないような順位で表現されることになる。もちろん、東京の違いも大きいが新潟、広島が偏差値では上位10位以内に入ってくるのも大きい。

表２　郵便局数の回帰偏差値と人口当たりの郵便局数のランキング

郵便局数	回帰偏差値	順位 (a)	人口当たりの値	人口当たり順位(b)	(a) - (b)	人口順位
北海道	75.57	1	27.29	20	-19	7
青　森	47.20	29	24.53	26	3	28
岩　手	57.05	9	31.50	8	1	30
宮　城	45.58	33	19.45	32	1	15
秋　田	56.93	10	33.73	5	5	35
山　形	55.80	15	32.23	7	8	33
福　島	56.81	11	26.52	24	-13	17
茨　城	45.02	35	17.45	36	-1	11
栃　木	39.48	40	18.00	35	5	20
群　馬	37.44	41	17.09	38	3	19
埼　玉	32.37	46	9.27	46	0	5
千　葉	41.59	39	12.20	45	-6	6
東　京	55.93	14	12.63	44	-30	1
神奈川	34.33	45	8.96	47	-2	3
新　潟	62.20	5	28.31	17	-12	14
富　山	46.25	31	26.85	21	10	38
石　川	50.36	24	28.96	16	8	36
福　井	45.89	32	30.16	12	20	43
山　梨	47.93	28	30.74	9	19	41
長　野	63.69	4	30.71	10	-6	16
岐　阜	47.98	27	21.58	31	-4	18
静　岡	45.52	34	16.06	40	-6	10
愛　知	48.19	26	13.25	42	-16	4
三　重	52.59	19	25.36	25	-6	23
滋　賀	35.57	44	19.36	33	11	31
京　都	44.59	36	18.15	34	2	13
大　阪	50.76	22	12.82	43	-21	2
兵　庫	55.80	16	17.42	37	-21	8
奈　良	43.86	38	22.87	29	9	29
和歌山	49.61	25	29.63	13	12	39
鳥　取	53.21	18	40.46	2	16	47
島　根	65.80	3	49.67	1	2	46
岡　山	56.65	12	27.37	19	-7	21
広　島	58.77	7	24.52	27	-20	12
山　口	52.51	20	27.42	18	2	25
徳　島	44.47	37	29.25	15	22	44
香　川	36.35	42	22.09	30	12	40
愛　媛	51.03	21	26.72	22	-1	27
高　知	57.98	8	40.29	3	5	45
福　岡	50.69	23	16.13	39	-16	9
佐　賀	36.00	43	23.38	28	15	42
長　崎	55.52	17	29.53	14	3	26
熊　本	60.38	6	30.50	11	-5	22
大　分	56.37	13	32.92	6	7	34
宮　崎	46.57	30	26.58	23	7	37
鹿児島	71.27	2	40.15	4	-2	24
沖　縄	24.09	47	14.87	41	6	32

回帰偏差値による順位と人口当たりの数値による順位との違いを検討するために、これらの順位の差（(a) - (b)）をとって表２の５列目に表す。この値が正であることは回帰偏差値で表した順位の方が大きくなった（順位が下がった）ことであり、人口当たりの数値による順位で過大評価されていたことを示している（例えば福井、山梨、鳥取、徳島、佐賀など）。逆に、この値が負であることは回帰偏差値で表した順位の方が小さくなった（順位が上がった）ことであり、人口当たりの数値の比較による順位で過小評価されていたことを示している（例えば北海道、福島、東京、愛知、大阪、兵庫、広島、福岡など）。

これらの過大評価・過小評価が都道府県人口の大きさとどのような関係にあるかを見るために都道府県人口の大きさの順位を６列目に示す。この人口の順位を横軸にとって縦軸に回帰偏差値による順位と人口当たりの数値による順位との差（(a) - (b)）をとって図に示せば有意な正の相関が見られる。つまり順位の変化が人口の大きさに依存して起こっていることを示している。したがって、人口に対して上に凸の規模の効果をもつような変量では、これを人口当たりの数値で相互評価を行えば、人口が大きい地域が過小評価され、人口が小さい地域が過大評価されることが明らかに理解される。

4.2 薬局数の変異

都道府県毎の薬局数についての解析も上と同じように人口との回帰分析を行う。

まず実数で表した薬局数と人口との関係は図５で示すように、ほぼ直線回帰でき、相関の度合いもきわめて高い (R^2= 0.9491)。したがって図６で示すように、縦軸に人口当たりの数値で表現すれば、点の分布はほぼ水平に分布する。両対数変換して表すと図７で示すように傾きはほぼ 1 (0.9988) で、相関の度合いも強い (R^2= 0.9395)。

回帰偏差値を求めて表３に示す。人口当たりの薬局数を表３の３列目に示す。この数値は図６の縦軸と同じで、図６の点の分布を縦軸のみで表したのが、これら表３の数値である。これら数値の大小を順位づけると偏差値の順位とほぼ同じであることがわかる。

回帰偏差値による順位と人口当たりの数値による順位との違いを表３の５列目に表しているが、順位の変化は４カ所で順位の交代があっただけである。したがって、人口の順位を横軸にとって、縦軸に回帰偏差値による順位と人口あたりの数値による順位との差をとって図８に示せば分布点がほとんど図のＸ軸上にのってしまう。

これら薬局数の解析で、実数値で表して人口に対してほぼ直線性を示すような変量では、回帰偏差値による順位と人口当たりの順位の間には大きな変化が見られないことが明らかである。

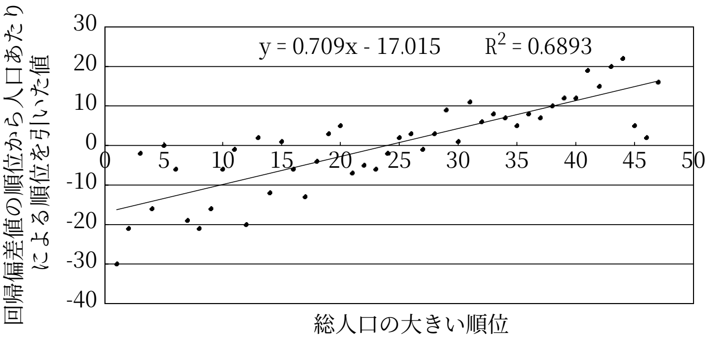

図４ 郵便局数の回帰偏差値順位と人口当たりの数値による順位の差の分布 （郵便局数）

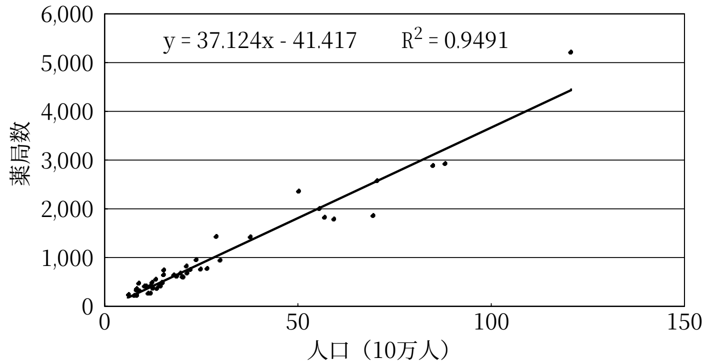

図５ 実数で表した薬局数と人口との関係

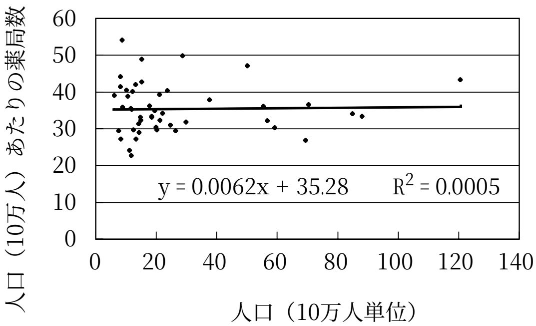

図６ 人口あたりの薬局数と人口との関係

4.3 一般旅券発行数の変異

一般旅券発行数と人口との関係を実数値で表現すると図９のようになる。図中の点の分布の中程で下の方に一つ大きく離れているのは北海道を示す点である。この図からは一般旅券発行数と人口との傾向はつかみにくいが、図 10 に示したように縦軸を人口当たりの一般旅券発行数で表すと傾向がはっきりしてくる。人口当たりの一般旅券発行数は人口と明らかに正の相関を示す。図10で直線回

帰をしてみると、傾きは0.2646となり、０から統計的に高度に有意に（t = 6.03)０とは異なっている。これは、図11に示すように両対数変換すると直線回帰の傾きが１より有意に（t = 4.7091）大きいことを示している。

回帰偏差値を求めて表４に示す。偏差値が大きい方に 1* σ以上離れているのは、大きい方からいうと奈良、滋賀、山梨、京都、福井、鳥取の順である。逆に、偏差値が小さい方に 1* σ以上離れているのは、小さい方からいうと秋田、青森、北海道、岩手、新潟、福島の順である。

人口当たりの一般旅券発行数を表４の３列目に表す。この数値は図10の縦軸と同じで、図10の点の分布を縦軸のみで表したのが、これらの数値である。これら数値の大小を順位づけると偏差値の順位とは大きく異なって表４の４列目に表したようになる。図９で分布から大きく離れていた北海道は、この順位付けでは４５位となっている。

回帰偏差値による順位と人口当たりの数値の比較による順位との違いを検討するために、これらの差（(a) - (b)）をとって表４の５列目に表す。この値が正であることは回帰偏差値で表した順位の方が大きくなった（順位が下がった）ことであり、人口当たりの数値による順位で過大評価されていたことを示している（例えば、茨城、埼玉、千葉、静岡、愛知、大阪、兵庫など)。逆に、この値が負であることは回帰偏差値で表した順位の方が小さくなった（順位が上がった）ことであり、人口当たりの数値による順位で過小評価されていたことを示している（例えば、福井、山梨、和歌山、鳥取、徳島、香川、佐賀、大分など)。

これらの過大評価・過小評価が人口の大きさとどのような関係にあるかを見るために都道府県人口の大きさの順位を６列目に示す。この人口の順位を横軸にとって縦軸に回帰偏差値の順位と人口当たりの数値による順位との差をとって図12に示せば有意な負の相関が見られる。つまり、順位の変化が人口の大き

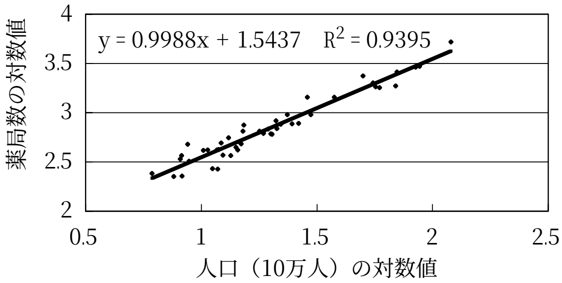

図７　両対数変換した薬局数と人口との関係

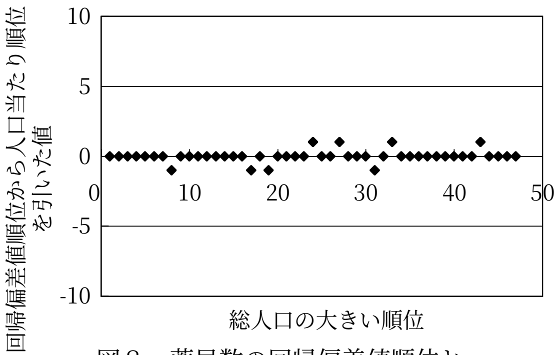

図８　薬局数の回帰偏差値順位と人口当たり順位の差の分布

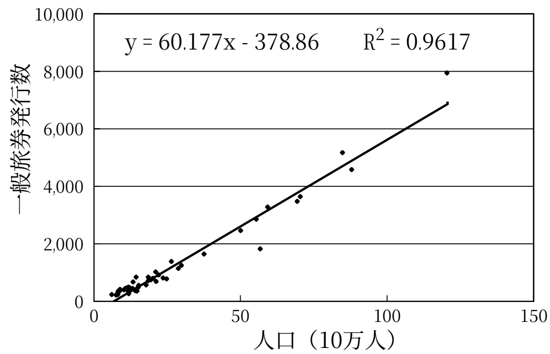

図９　実数で表した一般旅券発行数と人口との関係

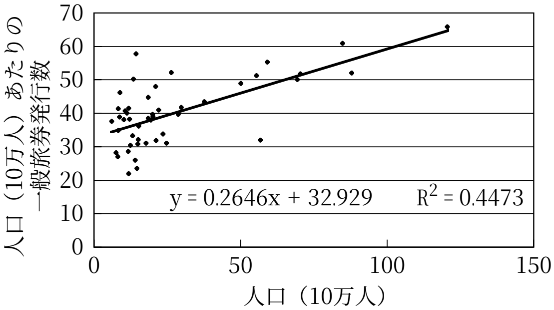

図10　人口当たりの一般旅券発行数と人口との関係

さに依存して起こっていることを示している。

これら一般旅券発行数の解析によって、人口に対して下に凸の規模の効果をもつような変量では、これを人口当たりの数値で相互評価を行えば、人口が大きい地域が過大評価され、人口が小さい地域が過小評価されることが明らかである。

５．　回帰偏差値の適用例

5.1　動物の脳化指数：脳重量と生体重との相関

Weisberg (1985)が、かれの著書の回帰分析の残差分析で引用している動物の脳の重量と生体重の一覧表は動物における脳の進化を考える上で大変興味深い課題を提起している。このデータをもとに目的変数を脳重量、独立変数を生体重として両対数変換した値の回帰分析を行って図13に示す。生体重に対する脳重量の回帰偏差値を計算して表５に示す。これは、脳重量の実測値が、生体重で期待される脳の重量よりも大きいか小さいかの回帰偏差値を示す指数である。したがって、進化的観点から見れば、動物一般で期待されている以上に脳の発達が見られるかどうかの目安になるので“脳化指数”とでも表現できる指数である。図13の回帰直線から上の方にはなれて見える点がヒトの脳重量・生体重の点であることは容易に理解できる。ヒトの偏差値は78であり、平均から2*σ（20）以上はなれている。偏差値で、60以上である11種の動物で8種類がサルの仲間である。これはサルの仲間で脳の発達が平均以上に加速していたこと示しており、進化学的に興味深い課題を提出する。

表３　薬局数の回帰偏差値による順位と人口当たりの数値による順位との差

薬局数	回帰偏差値	順位 (a)	人口当たりの値	順位 (b)	(a) - (b)	人口順位
北海道	45.71	32	32.17	32	0	7
青　森	47.29	29	33.13	29	0	28
岩　手	44.22	34	31.29	34	0	30
宮　城	57.84	11	40.34	11	0	15
秋　田	50.58	22	35.24	22	0	35
山　形	41.34	39	29.66	38	1	33
福　島	45.80	30	32.20	31	-1	17
茨　城	45.01	33	31.73	33	0	11
栃　木	42.67	36	30.37	36	0	20
群　馬	41.35	38	29.63	39	-1	19
埼　玉	35.65	45	26.82	45	0	5
千　葉	42.26	37	30.21	37	0	6
東　京	62.50	6	43.24	6	0	1
神奈川	48.74	25	34.00	25	0	3
新　潟	43.59	35	30.90	35	0	14
富　山	30.11	46	24.09	46	0	38
石　川	26.74	47	22.61	47	0	36
福　井	36.39	44	27.14	43	1	43
山　梨	51.43	20	35.81	20	0	41
長　野	48.93	24	34.15	24	0	16
岐　阜	56.41	13	39.28	13	0	18
静　岡	54.40	16	37.78	16	0	10
愛　知	52.77	17	36.58	17	0	4
三　重	47.63	26	33.33	26	0	23
滋　賀	36.52	43	27.10	44	-1	31
京　都	41.05	40	29.46	40	0	13
大　阪	47.53	27	33.28	27	0	2
兵　庫	51.98	18	36.08	19	-1	8
奈　良	39.97	42	28.90	42	0	29
和歌山	55.73	15	38.79	15	0	39
鳥　取	56.10	14	38.99	14	0	47
島　根	40.76	41	29.43	41	0	46
岡　山	49.94	23	34.80	23	0	21
広　島	69.13	2	49.77	2	0	12
山　口	68.10	3	48.89	3	0	25
徳　島	62.67	5	44.05	5	0	44
香　川	58.03	10	40.47	10	0	40
愛　媛	45.79	31	32.22	30	1	27
高　知	59.31	9	41.40	9	0	45
福　岡	66.44	4	47.09	4	0	9
佐　賀	73.72	1	54.05	1	0	42
長　崎	60.88	7	42.72	7	0	26
熊　本	47.32	28	33.14	28	0	22
大　分	57.55	12	40.13	12	0	34
宮　崎	50.93	21	35.47	21	0	37
鹿児島	51.91	19	36.11	18	1	24
沖　縄	60.03	8	42.03	8	0	32

表4　一般旅券発行数の回帰偏差値による順位と人口当たりの数値による順位の差

	回帰偏差値	順位 (a)	人口当たりの値	順位 (b)	(a) - (b)	人口順位
北海道	29.9	45	31.9	36	9	7
青　森	28.0	46	23.4	46	0	28
岩　手	33.4	44	25.9	45	-1	30
宮　城	41.5	37	33.8	33	4	15
秋　田	26.7	47	21.9	47	0	35
山　形	42.4	36	30.4	41	-5	33
福　島	39.6	42	31.8	37	5	17
茨　城	49.6	28	41.7	16	12	11
栃　木	50.8	23	39.6	22	1	20
群　馬	50.0	27	39.0	24	3	19
埼　玉	50.5	24	50.1	10	14	5
千　葉	57.0	14	55.3	4	10	6
東　京	59.2	8	65.8	1	7	1
神奈川	58.5	9	60.9	2	7	3
新　潟	36.9	43	31.0	38	5	14
富　山	56.9	15	40.0	21	-6	38
石　川	58.2	11	41.5	17	-6	36
福　井	61.5	5	41.3	18	-13	43
山　梨	66.2	3	46.1	13	-10	41
長　野	51.5	22	40.9	19	3	16
岐　阜	59.8	7	48.0	12	-5	18
静　岡	49.3	29	43.4	15	14	10
愛　知	52.0	21	51.7	7	14	4
三　重	57.5	13	44.7	14	-1	23
滋　賀	66.3	2	50.2	9	-7	31
京　都	61.6	4	52.1	5	-1	13
大　阪	50.1	26	52.0	6	20	2
兵　庫	53.8	18	51.2	8	10	8
奈　良	72.5	1	57.8	3	-2	29
和歌山	58.2	10	40.7	20	-10	39
鳥　取	59.8	6	37.6	30	-24	47
島　根	43.1	35	28.1	43	-8	46
岡　山	48.9	30	37.9	29	1	21
広　島	47.4	32	39.6	23	9	12
山　口	48.9	31	36.1	31	0	25
徳　島	53.0	19	34.8	32	-13	44
香　川	55.4	16	38.1	28	-12	40
愛　媛	41.2	38	30.7	40	-2	27
高　知	40.5	39	27.0	44	-5	45
福　岡	52.5	20	48.9	11	9	9
佐　賀	57.9	12	38.9	25	-13	42
長　崎	43.2	34	32.1	35	-1	26
熊　本	50.2	25	38.5	26	-1	22
大　分	53.8	17	38.2	27	-10	34
宮　崎	40.0	40	28.6	42	-2	37
鹿児島	40.0	41	31.0	39	2	24
沖　縄	46.3	33	33.3	34	-1	32

5.2　公共事業費

経済学的観点からも地域社会学的観点からも重要な財政問題の一つである公共事業費について考えてみる。8月18日の朝日新聞は「地方へのカネ　活力そいだ」との見出しのもと、「県別の一人当たり公共投資額（98年度）の上位は島根、北海道、秋田などが占める。・・・」とし、「バブル崩壊で失業者が増えたのは都市圏だったのに、90年代の公共投資は自民党支持基盤を中心にした地域への所得分配政策だった」とする、BNPバリバ証券の河野龍太郎の見解を載せている。

この主張となったのは、河野龍太郎が『金融ビジネス』にのせた論文（2001）であるが、本論の課題である“相対評価をいかにおこなうか”という点できわめて興味深い課題を含んでいるので検討してみたい。結論を先に言うと、統計分析をすべて一人当たりの数値で行っているために、人為的な統計的操作によって間違った解釈に導かれた例として興味深いのである。

まず、はじめに、98年度公共事業費の回帰

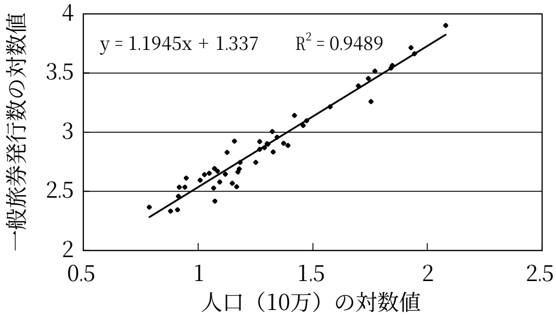

図11　両対数変換した一般旅券発行数と人口との関係

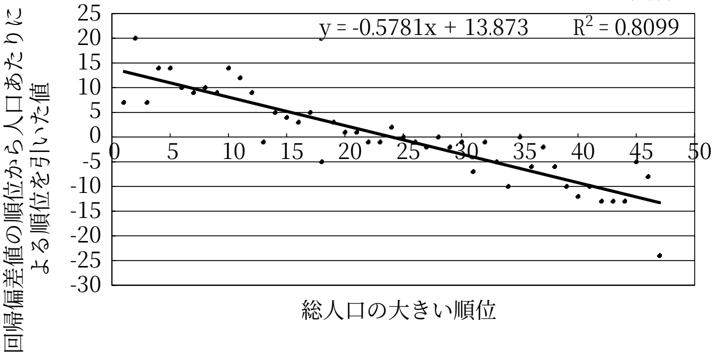

図12　回帰偏差値による順位と人口あたりによる順位の変化（一般旅券発行数）

分析を行う。データは『民力2002年版』の国土交通省がまとめた“一般公共事業費”を用いた。図14は、実数での分布図である。両対数変換して表すと図15で示すように回帰直線の傾きは0.7958で、上に凸の規模の効果が逓減している。ということは、先の郵便局の解析で考察したように、人口が大きい地域が過小評価され、小さい地域が過大評価される傾向になっているのである。

これは、回帰偏差値を求めると明瞭にわかる。表6は、朝日新聞が図として掲載していた“一人当たり公共事業費”のランキングと回帰偏差値によるランキングを対照的に表したものである。表6の右に示した一人当たり公共事業費のランキングを見ると“都市対地方”という図式にのっとった明瞭なちがいが見られる。20位、21位の岐阜、兵庫あたりまでは地方と感じさせる県ばかりが並んでいる。逆に、下位10県には、埼玉、東京、大阪、愛

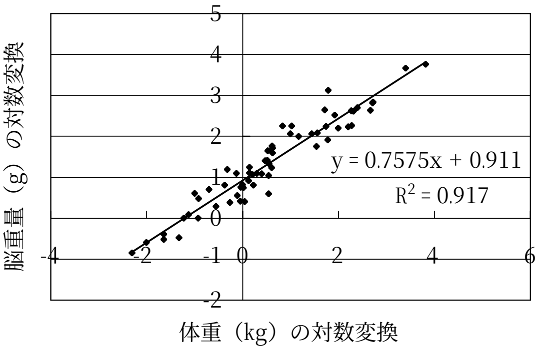

図13　脳と生体重との相関

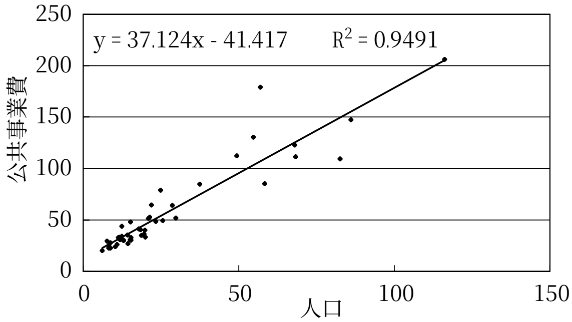

図14　公共事業費と人口との実数での相関

表５　脳化指数

動物名	回帰偏差値
Human	78.25
Rhesus monkey	73.43
Baboon	68.72
Owl monkey	67.21
Ground squirrel	64.83
Chimpanzee	64.30
Patas monkey	62.92
Vervet	62.55
Arctic fox	61.07
Galago	60.53
Red fox	60.42
Rock hyrax	59.02
Mole rat	58.59
Genet	57.21
Racoon	56.69
Roe deer	56.39
Asian elephant	55.88
Chinchilla	55.84
Gray seal	54.62
Echidna	54.13
Cat	53.43
Slow loris	52.47
Goat	51.89
Eastern American mole	50.67
Star-nosed mole	50.48
Sheep	50.35
Little brown bat	50.06
Donkey	49.66
Rock hyrax	49.66
Phalanger	49.57
Gray wolf	49.48
Lesser short-tailed shrew	49.25
okapi	48.74
Tree hyrax	48.37
African elephant	48.25
Gorilla	48.09
Mouse	47.70
African giant pouched rat	46.98
Mountain beaver	46.66
Arctic ground squirrel	45.78
Rabbit	45.72
Yellow-bellied marmot	45.36
Giraffe	45.23
Horse	44.85
Guniea pig	43.94
Big brown bat	43.49
Rat	42.92
Golden hamster	42.89
Jaguar	42.34
North American opossum	40.56
European hedgehog	40.51
Nine-banded armadillo	40.45
Crow	39.72
Kangaroo	38.99
Desert hedgehog	38.95
Giant armadillo	38.40
Brazilian tapir	38.22
Pig	37.12
Musk shrew	36.81
Tenrec	34.77
Tree shrew	32.00
Water opossum	25.88

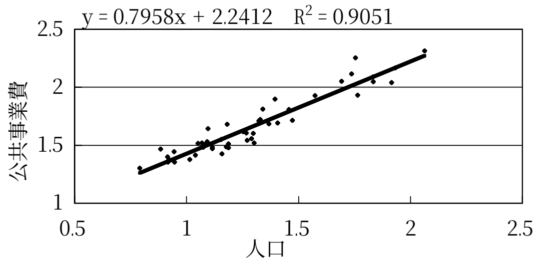

図 15　公共事業費と人口との両対数変換での相関

表６　一人あたり公共事業費、
公共事業費回帰偏差値

順位		回帰偏差値	順位		一人当たりの値
1	北海道	78.51	1	島　根	0.38
2	新　潟	69.10	2	山　形	0.35
3	山　形	67.11	3	鳥　取	0.32
4	島　根	66.61	4	新　潟	0.32
5	長　野	63.83	5	愛　媛	0.31
6	愛　媛	63.55	6	北海道	0.31
7	兵　庫	62.98	7	山　梨	0.31
8	福　岡	59.21	8	高　知	0.30
9	山　梨	57.55	9	長　野	0.29
10	富　山	55.98	10	富　山	0.29
11	静　岡	55.71	11	徳　島	0.28
12	東　京	55.55	12	石　川	0.28
13	鳥　取	55.26	13	大　分	0.27
14	高　知	55.21	14	福　井	0.27
15	石　川	54.62	15	秋　田	0.26
16	福　島	54.05	16	宮　崎	0.26
17	大　分	53.56	17	佐　賀	0.25
18	岐　阜	53.37	18	岩　手	0.25
19	広　島	52.08	19	福　島	0.25
20	秋　田	51.70	20	岐　阜	0.24
21	徳　島	51.48	21	兵　庫	0.24
22	埼　玉	50.28	22	和歌山	0.24
23	大　阪	49.84	23	沖　縄	0.23
24	岩　手	49.83	24	鹿児島	0.23
25	宮　崎	49.75	25	香　川	0.23
26	福　井	49.08	26	福　岡	0.23
27	鹿児島	48.66	27	滋　賀	0.23
28	佐　賀	46.19	28	静　岡	0.23
29	三　重	46.01	29	広　島	0.22
30	宮　城	45.87	30	三　重	0.22
31	沖　縄	45.50	31	山　口	0.21
32	愛　知	44.68	32	宮　城	0.21
33	和歌山	44.59	33	青　森	0.20
34	滋　賀	44.27	34	栃　木	0.20
35	京　都	42.72	35	長　崎	0.19
36	香　川	42.54	36	京　都	0.19
37	山　口	42.19	37	熊　本	0.19
38	栃　木	42.11	38	奈　良	0.18
39	青　森	39.97	39	岡　山	0.18
40	茨　城	38.80	40	埼　玉	0.18
41	長　崎	37.91	41	東　京	0.18
42	熊　本	37.60	42	茨　城	0.17
43	岡　山	37.44	43	大　阪	0.17
44	千　葉	37.01	44	群　馬	0.16
45	神奈川	35.06	45	愛　知	0.16
46	奈　良	34.39	46	千　葉	0.15
47	群　馬	31.67	47	神奈川	0.13

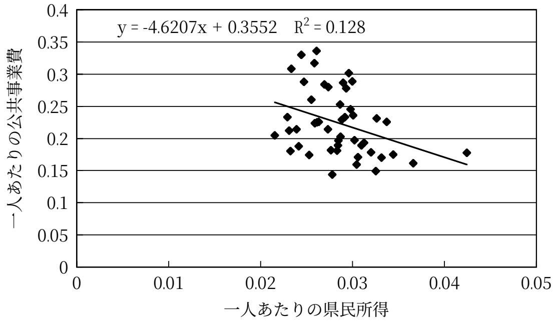

図 16（a）一人当たり公共事業費と
一人当たり県民所得との相関

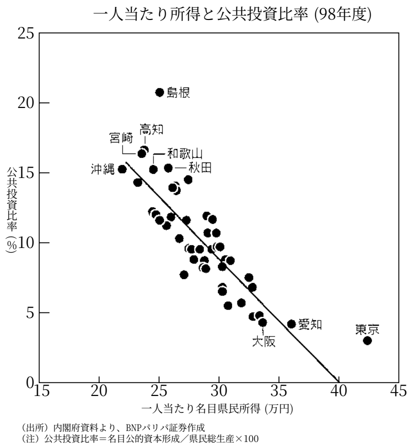

図 16（b）河野が示した一人当たり総資本形成
と一人当たり公共事業費との逆相関図

知、千葉、神奈川などの人口が多い都道府県が並んでいる。まさに、公共事業費が地方に厚く、都市に薄いという強い印象を与えさせる。しか

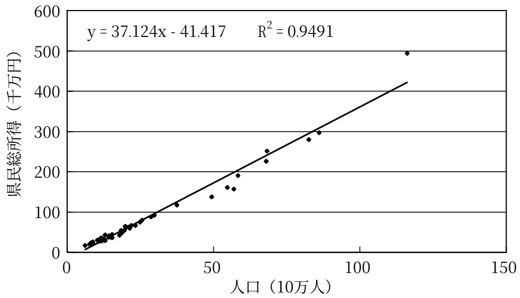

図 17　県民総所得と人口との実数での相関

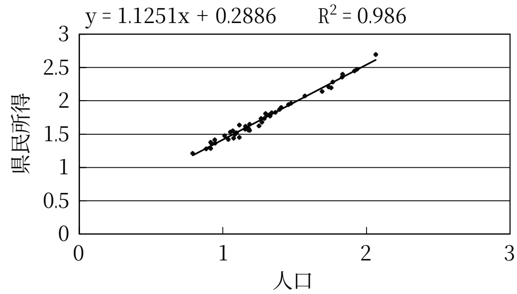

図 18　県民総所得と人口との両対数変換での相関

し、表 6 の左側に示した回帰偏差値を見ると全く異なる印象を抱かされる。上位 10 位に兵庫、福岡が入り、東京が 12 位になる。下位 10 位の中には神奈川、千葉も入っているが人口が少ないと思われる県もかなり入っている。

したがって回帰偏差値でみるかぎり公共事業費についての“ 都市対地方 ”という構図は明確には認められない。逆に、一人当たりの公共事業費のランクでは“ 規模の効果 ”のために必然的に人口の多い都道府県の順位が下位になってしまうことを明瞭に示している。

河野龍太郎（2001）が指摘していたもう一つの根拠は「一人当たりの総資本形成と一人当たりの公共事業費との間に負の相関が見られること」であった。したがって、「公共事業費は地方の活力を殺ぐ結果になった」という彼の結論が導かれることになる。

この魅力的な結論は確かに真実である可能性もあろうが､かならずしも彼の説を証明するものではない。つまり彼が示した負の相関は、彼の説が真実であろうと無かろうと、実は人為的な統計的操作によって生じるのである。以下に、その根拠を説明する。

総資本形成がデータとして得られないので、代わりに県民総所得の値を用いる。つまり、“ 一人当たりの県民所得と一人当たりの公共事業費に相関があるか ”を検討する。先に結果を示すと、図 16 (a)でわかるように、明らかに負の相関が見られ、河野の指摘した図 16(b)と一致するのである。しかし、これは、人為的統計的操作によって生じたものである。というのは、規模の効果が上に凸の変量

と下に凸の変量との間で相関をとるとどのような変量間でも必然的に負の相関になってしまうのである。

公共事業費は図 15 で示したように人口に対して上に凸の規模の効果を示していた。筆者の指摘が正しいならば県民所得は人口に対して下に凸の規模の効果を示しているはずである。これを確認するために、県民所得と人口との関係を検討する。図17は実数での県民所得と人口との相関で、やや下に凸の傾向を示している。さらに、両対数変換した県民総所得と人口との相関をみると図 18 のようになり、回帰直線の傾きは 1.1251 で、統計的に有意に 1 より大きく、確かに下に凸の規模の効果を示していることがわかる。

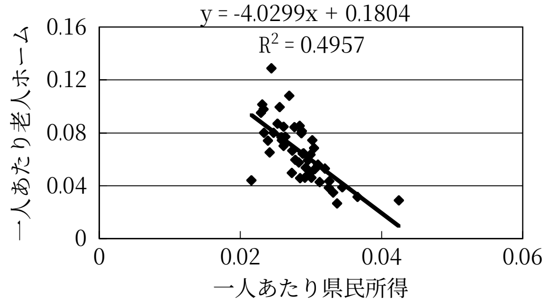

図 19　一人当たり県民所得と一人当たり老人ホーム数

表７　みかけ上の負の相関がおこる場合

統計指標の組み合わせ	相関係数
一人あたり県民所得と一人あたり老人ホーム数	-0.7041
一人あたり県民所得と一人あたり病院数数	-0.5318
一人あたり県民所得と一人あたりガソリンスタンド数	-0.6125
一人あたり県民所得と一人あたり郵便局数	-0.6585

筆者が指摘するように“ 規模の効果が上に凸の変量と下に凸の変量との間で相関をとると必然的に負の相関を示すことになる ”ことが示された。さらに、これが普遍的事実であることを示すために表７に、このような見かけ上の負の相関がおこる例を、今まで解析してきた例の中から示した。一人当たり県民所得と一人当たり老人ホーム数、一人当たり県民所得と一人当たり病院数数、一人当たり県民所得と一人当たりガソリンスタンド数、一人当たり県民所得と一人当たり郵便局数の４つの例で、すべて上に凸の規模の効果を示す量と下に凸の規模の効果を示す量との間には有意な負の相関が検出される。すべての場合で、相関に意味があるようには思えない組み合わせであるが、単純に統計的操作から見かけ上の負の相関が導かれる。この中で、一例として図19に一人当たり県民所得と一人当たり老人ホーム数の負の相関を図示する。

さて、それでは公共事業費と県民所得の関係を正しくはどのように理解すれば良いのであろうか？このためには、公共事業費の人口に対する回帰偏差値を考えたように、県民所得の人口に対する回帰偏差値を考えると良い。河野が指摘するように、もし、公共事業費の効果がなかったとすれば公共事業費の回帰偏差値と県民所得の回帰偏差値との間に相関がないか、負の相関が見られてもよいはずであろう。これを検討するために表８に県民所得の回帰偏差値を表し、図20に公共事業費の回帰偏差値と県民所得の回帰偏差値との相関を図示する。相関は弱く、回帰直線の傾きはやや正である。統計的検定では、この傾き(0.1364)は有意に０とは異なっている。したがって、公共事業費と県民所得の間には弱いながらも正の相関が検出されるのである。

公共事業費についての問題としては、朝日新聞や河野が主張するような“ 都市に厚く地方に薄い公共事業費が地方の資本形成にマイナスの効果をもたらしている ”というような結論を導くことはできない。しかも、回帰偏差値の相関からは、有意な弱い正の相関が見られたことは、“ 公共事業費は、都市、地方を問わず全国に一様に、資本形成にわずかなながらも県民所得を高める効果をもたらしている ”という結論になる。

６．総合考察

本論の主眼は、考案した回帰偏差値の効用を確認し、その有効性を検討することにある。このため都道府県の統計指標から人口増加に対して正の規模の効果（効用逓減）を示す場合（郵便局数）、規模の効果を示さない場合（薬局数）、負の規模の効果（効用逓増）を示す場合（一般旅券発行数）の三つの例を具体例として用いて詳細な解析を行った。そして従来の人口当たりの値でもって相互評価を行う場合には、人口の大きい地域や小さい地域は、その規模の効果の大小・正負に応じて過大評価や過小評価されること、したがって従来の人口当たりの値でもって順位付けする場合は相対評価が誤ることを指摘した。

このような、いわゆる“ 規模の効果 ”と言われてきた二変量間の非線形性の度合いは、両対数変換を行って得た回帰直線の傾きの１からのずれで表される。この回帰直線の１からのずれの統計的有意性は通常のｔ検定で判断できる。表１に示したように都道府県の様々な指標値の場合には、この回帰直線の傾きの標準誤差は　おおまかに 0.04~ 0.06 である。都道府県の相対評価を行う場合は標本数が 47 であるから検定の自由度は45である。自由度45の 5% 有意水準を示すtの値は 2.02 であるから傾きが 1.1 か 0.9 ほどであれば統計的に有意な非線形性が認められることになる。このような非線形性は統計指標ではよく見られる現象で、本論で詳細に検討したように、従来、そして今でも慣用的・慣習的に使用されている人口当たりの値でもって順位付けすることは全く間違った順位付けを与える可能性が強いので地域相対評価

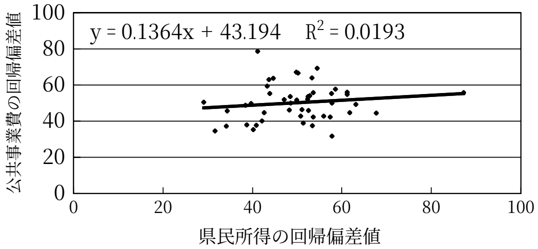

図 20　県民所得と公共事業の回帰偏差値間の相関

表８　県民総所得の回帰偏差値と一人当たりの県民所得の順位差

	回帰偏差値	順位 (a)	あたり県民所得	順位 (b)	(a) - (b)	人口順位
北海道	32.92	45	27.38	29	16	7
青　森	37.61	41	24.18	41	0	28
岩　手	45.37	35	25.91	35	0	30
宮　城	48.75	29	28.47	24	5	15
秋　田	46.06	33	25.56	37	-4	35
山　形	47.90	31	26.12	34	-3	33
福　島	45.67	34	27.33	30	4	17
茨　城	52.97	18	30.60	11	7	11
栃　木	62.81	7	32.03	8	- 1	20
群　馬	57.60	10	30.47	12	-2	19
埼　玉	50.74	26	33.18	5	21	5
千　葉	50.77	25	32.56	7	18	6
東　京	70.54	1	42.49	1	0	1
神奈川	49.94	27	33.75	4	23	3
新　潟	51.94	23	29.61	17	6	14
富　山	62.88	6	29.82	16	-10	38
石　川	63.34	3	30.11	14	-11	36
福　井	62.98	5	28.68	23	-18	43
山　梨	63.27	4	29.00	20	-16	41
長　野	54.84	14	29.99	15	-1	16
岐　阜	52.55	19	29.17	19	0	18
静　岡	52.30	21	31.27	9	12	10
愛　知	61.19	8	36.65	2	6	4
三　重	53.10	17	28.87	21	-4	23
滋　賀	70.22	2	32.63	6	- 4	31
京　都	56.29	12	31.01	10	2	13
大　阪	51.57	24	34.44	3	21	2
兵　庫	40.62	37	29.32	18	19	8
奈　良	54.33	15	28.32	26	-11	29
和歌山	40.54	38	23.92	42	-4	39
鳥　取	56.08	13	25.87	36	-23	47
島　根	47.31	32	24.45	40	-8	46
岡　山	48.03	30	27.66	28	2	21
広　島	52.21	22	30.23	13	9	12
山　口	53.93	16	28.44	25	-9	25
徳　島	56.42	11	26.97	31	-20	44
香　川	60.13	9	28.72	22	-13	40
愛　媛	39.76	39	24.72	39	0	27
高　知	41.74	36	23.37	43	-7	45
福　岡	36.48	42	27.81	27	15	9
佐　賀	52.45	20	26.14	33	-13	42
長　崎	33.44	44	23.29	44	0	26
熊　本	39.45	40	25.29	38	2	22
大　分	48.97	28	26.37	32	-4	34
宮　崎	35.14	43	22.93	46	-3	37
鹿児島	30.75	46	23.11	45	1	24
沖　縄	27.69	47	21.57	47	0	32

として順位付けを行う場合は十分注意しなければならない。

このような規模の効果に依存しないで相対評価を行うことが出来る指数として“ 回帰偏差値 ”を考案した。これは、二変量を両対数変換して実数で見られる規模の効果による非線形性をのぞき分散を安定化させ、対数変換値にひいた回帰直線からの標準化偏差 (Studentized residuals) でもって相対評価を行い、分布の中の順位付けを行う方法である。統計的に有意な、負や（一般旅券発行数）正の（郵便局数）規模の効果の場合に、従来の人口当たりの数値を用いてつけた順位と回帰偏差値による順位との間に、人口の大小に依存した変動が起こることが明白に示され、回帰偏差値の有効性がはっきりと示された。

この回帰偏差値は異質な属性（例えば、人口がちがうとか、GDP がちがうとか、体重がちがうなど）を持つ対象標本からの統計指標を相対評価することが出来るので大きな有効性・有用性を持っている。

まず、回帰偏差値は、回帰分析における残差分析として統計学的理論付けが十分行われている。通常の回帰分析（Draper and Smith 1966, Sokal and Rohlf 1973, Chatterjee and Price 1977, Dixon and Massey 1983, 広津　1992, 大友　1997）では、残差の分析・評価は回帰からはずれた外れ値を検出するのに用いられてきただけで、筆者が考案したような偏差値化して相対評価に用いる手法はいまだ例を見ない。標準化した偏差は標本数が多いと正規分布（０,１）にほぼ従う統計量とみなして良いので (Chatterjee and Price 1977)、原理としては、さまざまな試験などで従来用いられてきた正規分布に基づくＺスコア (Z score) (eg Dixon and Massey 1983) を平均値 50 とした通常の偏差値と異なることはない。特に本論で考察したような、地域間の相対評価を考える際には有効かつ有力である。したがって、筆者が本論で考案した回帰偏差値の手法は、すべての回帰分析において適用することが可能である。本論では、一次回帰だけを考えたが、二重回帰でも回帰偏差値の考え方は有効である。

回帰直線を利用した方法は、回帰直線からはずれた2~３ の異常点に大きく左右される傾向にあるから危険性が大きいという批判もあり得る。しかし、そのための残差を用いた残差分析であり、この残差を影響力係数を用いて有効利用した回帰偏差値なのである。従来の“ 人口当たり ”という慣用的手法にしたがっで“ 暗黙の内に ”非線形効果をふくんだ相対評価を行って過大評価をしたり過小評価をして意味のない順位付けを行い都市対地域という歪んだ印象を与えているいることの方が、回帰偏差値の危険性よりももっと危険である。

また、回帰偏差値は偏差値であるから、尺度フリーであり、したがって、回帰偏差値の平均値や総和として異質な指標を混みにした総合的な相対評価を行うことも可能である。これは、一般の学力についての正規分布に基づく偏差値と同様のことで、ある学生の個別科目の偏差値の総和をとって全体の学力とみなす方法と同じことが可能である。しかし、従来の“ 人口当たりの数値 ”では、異質な指標間では規模の効果がそれぞれに異なるであろうから、回帰偏差値で可能であるような総和をもって総合評価とするようなことは出来ないのである。

本論では、従来の慣用的に用いられている“ 人口当たりの数値 ”は、人口の小さい地域や大きい地域を過大評価したり過小評価したりする危険性があることを指摘した。そしてこの危険性をのぞき、相対評価を可能にする回帰偏差値を提案したのである。今まで、慣用的に用いられてきた“ 一人当たり ”の指標値に替え、ここで詳述してきた“ 回帰偏差値 ”を用いることによって、より客観的な相対評価が可能となることが望ましい。そして、“ 都市対地域 ”というわかりやすい図式に十分気をつけて正当な相対評価を行わねばならない。

参考文献

Chatterjee,S., Price,B. (1977) *Regression Analysis by Example.* John Wiley & Sons, Inc. (佐和隆光・加納悟訳（1981）『回帰分析の実際』新曜社)

Cook,R.D. and Weisberg,S. (1994) *An introduction to regression graphics.* John Wiley & Sons, Inc. pp. 203-216.

Cook,R.D. and Weisberg,S. (1999) *Applied regression including computing and graphics.* John Wiley & Sons, Inc. pp. 354-372.

Dixon, W.J., Massey,F.J. (1983) *Introduction to Statistical Analysis,* 4th ed. McGraw-Hill.

Draper,N.R., Smith,H. (1998) *Applied Regression Analysis,* 3rd ed. John Wiley ＆ Sons, Inc.（中村慶一訳（1968）『応用回帰分析』　森北出版）

広津千尋（1992）『実験データの解析　－分散分析を超えて－』共立出版.

河野龍太郎（2001）90 年代の公共投資は景気対策ではなかった！、『金融ビジネス』JUN. 2001、76-80.

日下部眞一（2002）NPO の規模をはかる回帰偏差値、“ NPO 指数 ” の考案：NPO 指数を通して見えてきた地域格差、日本 NPO 学会機関誌『ノン・プロフィットレビュー』（投稿中）

Myers, R.H. (1990) *Classical and modern regression with applications,* 2nd ed. PWS-KENT Publishing Company, Boston.

大友篤（1997）『地域分析入門（改訂版）』東洋経済新報社.

総務省統計局（2002）『統計でみる県のすがた』（財）日本統計協会.

Sokal,R.R. , Rohlf,F.J. (1973) *Introduction to Biostatistics.* Freeman, New York. (藤井宏一訳（1983）『生物統計学』共立出版)

Snedecor,G.W., Cochran,W.G. (1980) *Statistical Methods,* 7th ed. The Iowa State University Press, Iowa. (奥野ら共訳（1972）『統計的方法原書第６版』岩波書店)

Weisberg,S. (1985) *Applied Linear Regression,* 2nd ed. John Wiley & Sons, Inc.