論　文

学力の二極化モデル－全国学力・学習状況調査を中心にして

Bipolarization Model of Academic Ability : Based on the National Assessment Test

藤田　尚文（教育科学コース）

FUJITA Naofumi

Abstract

This paper proposed a model which dealt with the expansion of the difference and the bipolarization of academic ability in the same way, and obtained the followings: (i) Almost all of the distributions of the National Assessment Test were described by the superposition of two Gaussians, named the H-and L-groups respectively, which had a different average and a standard deviation. (ii) The ratio of the two distributions was 0.35 : 0.65. (iii) When the model was applied to prefectural data, the H-group was found to be related to the educational background of the father, economic affluence, and subscription of newspapers, whereas the L-group was found to be related to the unemployment rate, the divorce rate, and the crime rate. Both groups were closely related to the social stratifications. When the model was applied to other data, the bipolarization of academic ability was found as early as the 4th grade of elementary school, and it had occurred for more than thirty years.

1．はじめに

本論文ではまず，以下の3点を指摘する．(1)過去4年間行われてきた全国学力調査の学力分布が見かけ上，二山の分布をしている場合も，一山の分布をしている場合も，その分布を2つの正規分布の重ね合わせ（正規混合分布モデル）として理解できること，(2)その2つの分布の面積比は，学力下位の群（L群）が65%，学力上位の群（H群）が35%という一定の面積比をもつものとして理解できること（このことを主張するために5つのモデルを立て，BICの比較からすべての分布の面積比が一定というモデルが選択された），(3)2つの正規分布の面積比が一定という制約条件を課した正規混合分布モデルを平成19年度の算数B，数学Bの都道府県別学力分布データに適用し，すべての都道府県のH群比率，H群の平均と標準偏差，L群の平均と標準偏差を算出し，これらの値と都道府県別のさまざまな統計指標との相関係数を調べたところ，H群は各県の親世代の大学・短大進学率や経済的豊かさに関わる指標と相関が高く，L群は失業率，離婚率，経済的困難に関わる指標との相関が高かった．

さらに全国学力調査に関する分析結果を詳細に報告するとともに，このモデルを国立教育政策研究所が行った，学年別学力調査の結果に適用し，二極化がどの学年から始まっているかを調べる。また二極化といわれる現象が年代的にいつ頃から始まったのかについても高知県の高校入試データをもとに分析する．さらになぜ二極化が生じるのかについての素描的なモデルを示す．最後にここで展開した分析結果の意義を考察する．

2．学力の二極化について

山田（2004，p52）は二極化，格差拡大，ディバイドという言葉について，以下のように説明している．まず正規分布の状態があり，この分布が左右に引き伸ばされ，扁平になったものが格差拡大，二山になったのものが二極化，さらにその二山に重なりがなくなったものがディバイドである．この説明は直感的に分かりやすいが，筆者はこれらを統一的に扱いたいと考えている．

その第一歩として，次のような「常識的な定義」を取り上げよう．「横軸に学力検査の得点を取り，縦軸に人

数または人数の比率を取ったとき，学力の分布が単峰（一山，unimodal）にならず，双峰（二山，bimodal）になっているとき，学力の二極化が起こっていると言う.」筆者はこのように学力の二極化を定義するのは不適切であると考えている．ここで筆者が主張するのは，学力調査の結果が双峰分布をしているということと，学力の二極化と言われている現象を区別することである．

上の「常識的な定義」では，右の山に属する者が学力上位の群で，左の山に属する者が学力下位の群とみなしているのだろうか．山と山の間の谷よりも右側が学力上位群で左側が学力下位群と単純に分類するのは学力の二極化の正しい理解であるとは思えない．

二極化を問題にするときは，その背後になんらかの社会的階層を想定しているはずである（例えば苅谷，1995；苅谷，2001；苅谷・志水・清水・諸田，2002；苅谷・志水，2004；志水，2005；耳塚，2007，苅谷，2008a,b)．学力の階層差とはまさにそういったことを指している言葉である．つまり2つの階層が，それぞれピークの値を異にする分布をしていると想定しているはずである．しかしながら従来の社会的階層と学力の関係を調査した研究においては，階層間の学力や学習意欲の平均に差があることを指摘しているが，階層ごとの分布については語られてこなかった．言うまでもないことだが，ある同一の階層に属している子どもたちにも学力差は存在するのであって，学力分布は二つの分布が重なり合ったものとして考えるべきである．例えば，塾に通っている子どもたちの群と塾に通っていない子どもたちの群を考えるとわかりやすい（志水，2007)．塾に通っている子どもたちは，概して成績がよいかもしれないが，なかには学力があまり身についていない子もいるだろうし，塾に通っていない子にも優秀な子は当然いるはずである．

したがって「学力の二極化とは，子どもたちの属する社会的・文化的・経済的階層が2つあり，それぞれが異なる平均と標準偏差をもつ学力分布をし，その2つの分布が重なり合っている状態」と定義すべきであろう．双峰分布であることは二極化の必要条件ではないのである．実際，学力分布には台形型のものがかなり多く見受けられる．こういったデータも二極化の証拠とする方が，はるかに学力問題に関する見通しがよくなるものと考えられる．つまり山田（2004）の言う格差拡大，二極化，ディバイドの状態はいずれも2つの分布が重なっている状態ととらえ，これらは2つの分布の平均値の隔たりが小さい場合，大きい場合，極端に大きい場合ととらえれば，統一的に理解できる．

3．学力の二極化モデル

ここで展開する学力の二極化モデルは次の（Ⅰ）（Ⅱ）からなる．さらにH群，L群の意味づけを行うためには（Ⅲ）のプロセスも必要となる．

（Ⅰ）単純正規混合分布モデル

ある学力分布があり，各得点が階級幅Δxで，x_1からx_nまで分布しているとする．このとき，それぞれの得点における相対度数を，$g(x_1)$から$g(x_n)$とすると，

$$\Sigma g(x_k)\cdot\Delta x=1. \tag{1}$$

ここで2つの正規分布の重ね合わせから成る分布を考え，得点x_kにおける確率密度$f(x_k)$を以下のように定義する．

$$f(x_k)=p\cdot \mathrm{normdist}(x_k,\ m_h,\ \sigma_h)\Delta x+q\cdot \mathrm{normdist}(x_k,\ m_l,\ \sigma_l)\Delta x. \tag{2}$$

$\mathrm{normdist}(x, m, \sigma)$は平均が$m$，標準偏差が$\sigma$，得点が$x$の正規分布関数の値を表す．これはGaussian mixture model（正規分布の重ね合わせモデルあるいは正規混合分布モデル）のもっとも単純な場合である．$p+q$は基本的には1である．ここでは平均の低い方の群をL群，高い方の群をH群と呼ぶことにする．これに対応して平均，標準偏差をm_l，σ_l，m_h，σ_hと表記する．

ここで以下のような残差平方和Sを最小にするm_l，σ_l，m_h，σ_h，pを求める問題を考える．

$$S=\Sigma\{g(x_k)-f(x_k)\}^2 \tag{3}$$

ある学力分布に対して（2）式が（3）式を最小にするパラメタをもつとき，それを第一段階における学力の二極化モデルの解と呼ぶことにする（最終的な二極化モデルの解は後述)．本論文では残差平方和の最小値問題を解くために，エクセルのソルバーを用いた．

（Ⅱ）さまざまな制約条件をもつモデル間の比較

ここでは（Ⅰ）で展開したモデルにさまざまな制約条件を課して，より適合度の高いモデルを探っていく．Table 1参照．平成19年，20年，21年に実施された全国学力調査は悉皆調査であり，平成22年に実施されたものは標本調査であった．こういった異なる性質のため，ここでは最初の3年間に実施された学力調査をまず，分析し，そこで得られた特徴が平成22年においてもあてはまるかを検討した．

最初の3年間の学力分布24個のうち，（Ⅰ）段階の分析の結果，3つは単調増加あるいはそれに準じる形をしており，2群を抽出することができなかった．したがってここで比較するモデルは残りの21個の学力分布全体に

Table 1　モデルの説明．H群比率の個数以外のパラメタの個数は，すべてのモデルで共通であり，小学校で4 x 10＝40個，中学校で4 x 11＝44個．

モデル名	モデルの説明	H群比率の個数
個別比率モデル	それぞれの学力分布が，それぞれ異なるH群比率をもち，その比率のもとで，2つの正規分布の形をそれぞれに推定．	小学校の分布で10個，中学校の分布で11個，計21個．
科目内一定比率モデル	小中合わせて8科目がそれぞれの科目内で3年間共通のH群比率，それぞれの比率のもとで2つの正規分布の形を推定	小学校の分布で4個，中学校の分布で4個，計8個
小中学校別年度内一定比率モデル	小学校のある年度の学力分布が4科目間で一定のH群比率，3年間で3つの比率，中学校のある年度の学力分布が4科目間で一定のH群比率，3年間で3つの比率をもち，それぞれの比率のもとで正規分布の形を推定	小学校の分布で3個，中学校の分布で3個，計6個
小中学校別単一比率モデル	小学校の学力分布共通の1つのH群比率，中学校の学力分布共通の1つのH群比率をもち，それぞれの比率のもとで2つの正規分布の形をそれぞれに推定．	小学校の分布で1個，中学校の分布で1個，計2個．
小中学校統一比率モデル	小中学校間ですべて共通の統一された1つのH群比率をもち，そのもとで2つの正規分布の形をそれぞれ推定．	小中学校の分布共通の1個．

関する制約を置くものである．また最少にすべき残差平方和は1つの分布形にたいするものではなく，21個の残差平方和の合計である．

個々のモデルの記述はTable 1にあるので，ここでは1点のみ指摘しておく．ここで検討するモデルにおいては，いずれかの群に属する特定の個人の分布内の位置が科目間で同じと仮定する必要はない．H群，L群のいずれかに属する子どもがいて，その子がそれぞれの分布内である値（同一の相対的位置ではない）を取ることがあるとすれば，科目間でH群の比率が一定というモデルを立てることが可能である．

これらのモデルを比較するさいにはBIC（ベイズ型モデル評価基準）を用いた．残差平方和を用いたBICの定義は以下のとおりである（小西，2010）．

$$\mathrm{BIC} = n \quad \ln[S/n] + t \quad \ln[n]. \tag{4}$$

ここでnは分布の個数（棒グラフの本数），lnは自然対数，Sは上で定義した残差平方和，tは推定に用いたパラメタの個数を表す．モデルを比較するさいには，BICの小さいものが，よりよいモデルとなる．一般にたくさんのパラメタを用いれば，残差平方和を小さくでき，BICの前項を小さくできるが，tの値が大きくなるという後項のペナルティが科されている．BICと類似した指標にAIC（赤池情報量基準）があるが，サンプル数が多い場合，BICはパラメタ数に対するペナルティがAICよりも大きくなり，パラメタ数の少ないモデルを選択する傾向がある．この性質は，本論文におけるモデル比較においては好ましい性質である．

共分散構造分析のソフトウェアAMOSのマニュアル(2003)によると，BICの差が0から2のとき，小さい値をもつモデルの方がよいとする弱い（weak）証拠，BICの差が2から6のときは明白な（positive）証拠，BICの差が6以上のときは強い（strong）証拠があるとされている．

（Ⅲ）都道府県データへの適用

前段で得られた最良の制約条件をもつモデルを都道府県別データに適用する．過去に行われたすべてのデータに適用すると膨大な量になるので，平成19年度の算数B,数学Bに制約条件を課して，分析する．このようにして都道府県別のH群比率，H群平均と標準偏差，L群の平均と標準偏差を求める．得られたH群比率等はかなりの程度ばらつくと思われる．このばらつきが都道府県別のどのような統計指標と関連するのかを調べる．

4．過去4年間に実施された全国学力調査への二極化モデルの適用

平成19年，20年，21年度の全国学力調査，21個の学力分布データに前節（Ⅱ）の制約条件を課して分析した結果，小中学校のすべての学力分布における2つの正規分布の面積比がL群65％，H群35％という小中学校統一比率モデル（面積比一定のモデル）が最良モデルであることがわかった．モデル比較の結果をTable 2に示す．さらにこのモデルを抽出調査である平成22年度の全国学力調査の学力分布に当てはめた結果，平成22年度においても，L群65％，H群35％という同じ面積比が得られた（用いた分布は8個のうち6個であった．2個は分布が単調増加だったので，分析からは除外）．面積比一定モデルをすべての学力分布データに適用した結果をFigure 1に示す．横軸は得点，縦軸は平成19，20，21年度が比率，平成22年度はパーセントとなっている．これは前者が悉皆調査，後者が抽出調査で，実際の文部科学省のデータにおいても，前者は各得点の人数が示されているが，後者はパーセントで表示されていることに対応している．図中の棒グラフは実測値，▲がL群，■がH群，無印の折れ線がL群とH群を足し合わせたモデル値である．すべてのグラフは実測値の分布形が一山のものであれ，二山のものであれ，二極化モデルによってきれいに近似されているのが見てとれる．

Table 2　モデル比較の結果

モデル名	パラメタ数	BIC
個別比率モデル	105	-4500
科目内一定比率モデル	92	-4554
小中学校別年度内一定比率モデル	90	-4561
小中学校別単一比率モデル	86	-4560
小中学校統一比率モデル	85	-4565

これらのグラフはL群，H群を構成する正規分布の面積比が一定であるが，L群，H群の平均や標準偏差はそれぞれに異なっている．グラフの全体的な特徴としては，L群がだいたいにおいて分布の全域に亘っているのに対し，H群は分布の右側半分にある．言い換えるとL群に属する子どもの中で，最優秀のものは最高得点をとっているが，H群に属する子どもの中で，もっとも成績のふるわないものは，L群の子どもの平均点と同程度である．

小学校においては二山分布をしているものはなく，中学校においては数学において二山分布が認められる．このことは小学校では，学力が二極化しておらず，中学校の数学においてのみ二極化していることを意味しない．単にL群とH群の平均値の間隔がやや小さめであるか，二山になるほど平均値間の距離が標準偏差の大きさと比べて大きめであるかにすぎない．すべての学力分布が2つの正規分布の重ね合わせとして統一的に理解できることが重要である．

ここで得られた知見のうち，面積比一定モデルが最良モデルとして採択されたことが最も重要である．日本人の小学生も中学生も全人口のうち，35％がH群に属し，65％がL群に属すということは，この比率の背後に実体としての社会階層が存在することを強く示唆するだろう．

次に面積比一定という制約条件のついた二極化モデルを，都道府県別データに適用した結果をTable 3に示す．全国学力調査においては，秋田県，福井県，富山県などが学力上位を占め，沖縄県，高知県，大阪府，北海道などが学力下位を占めている．H群比率を見ると，富山県，石川県，静岡県，愛知県，福井県などが高く，沖縄県，高知県，北海道などが低くなっていて，学力上位県，学力下位県と重複している場合もあれば，異なっている場合も見られる．たとえば秋田県の場合，全国学力調査においては，小学校で全国一となっているが，H群比率は28％と必ずしも最上位ではない．しかしながら小学校のH群平均，L群平均とも最上位にあり，結果として非常に高い成績を修めていることがわかる．

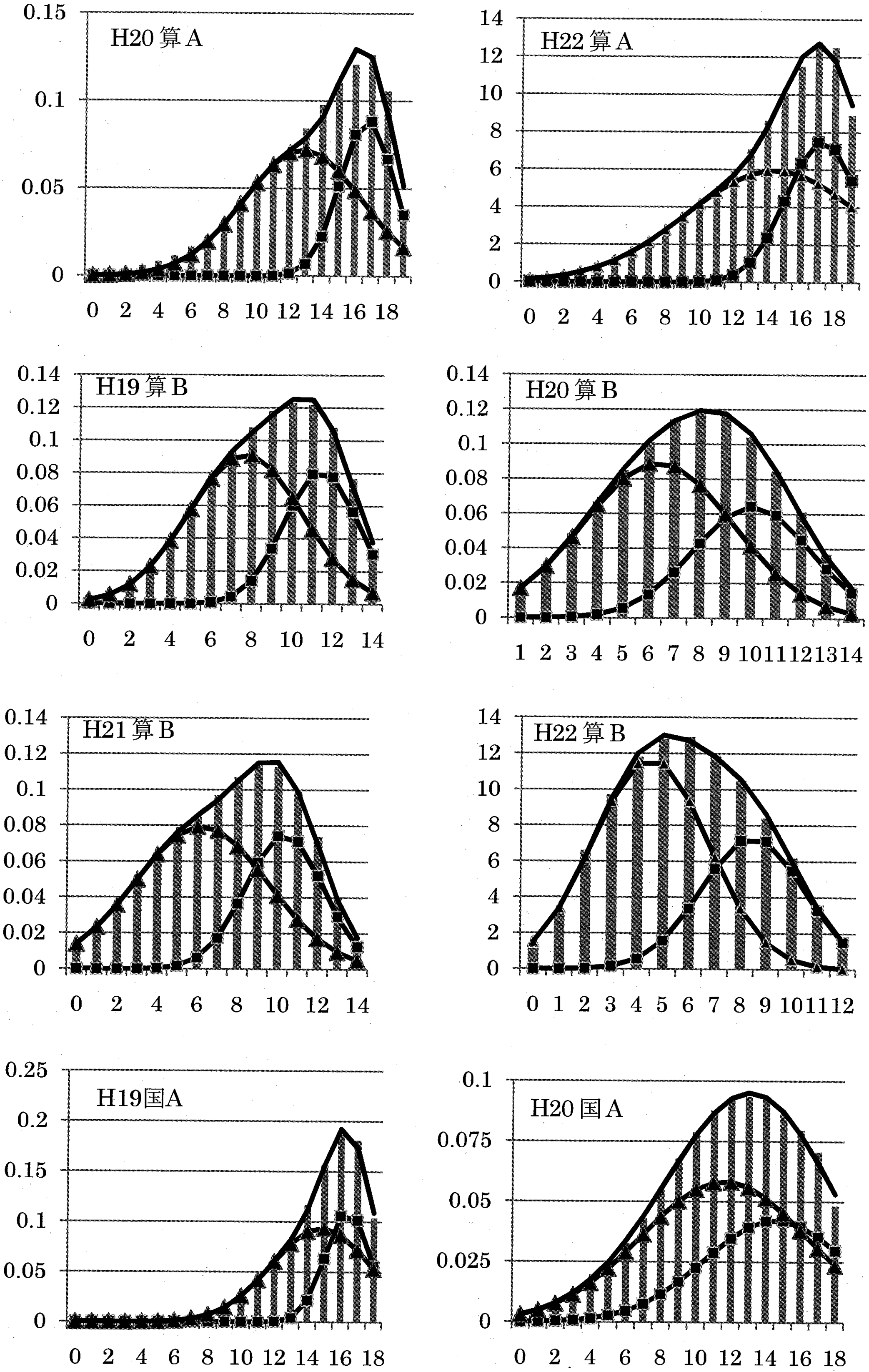

Figure 1　全国学力調査の学力分布とモデルの値

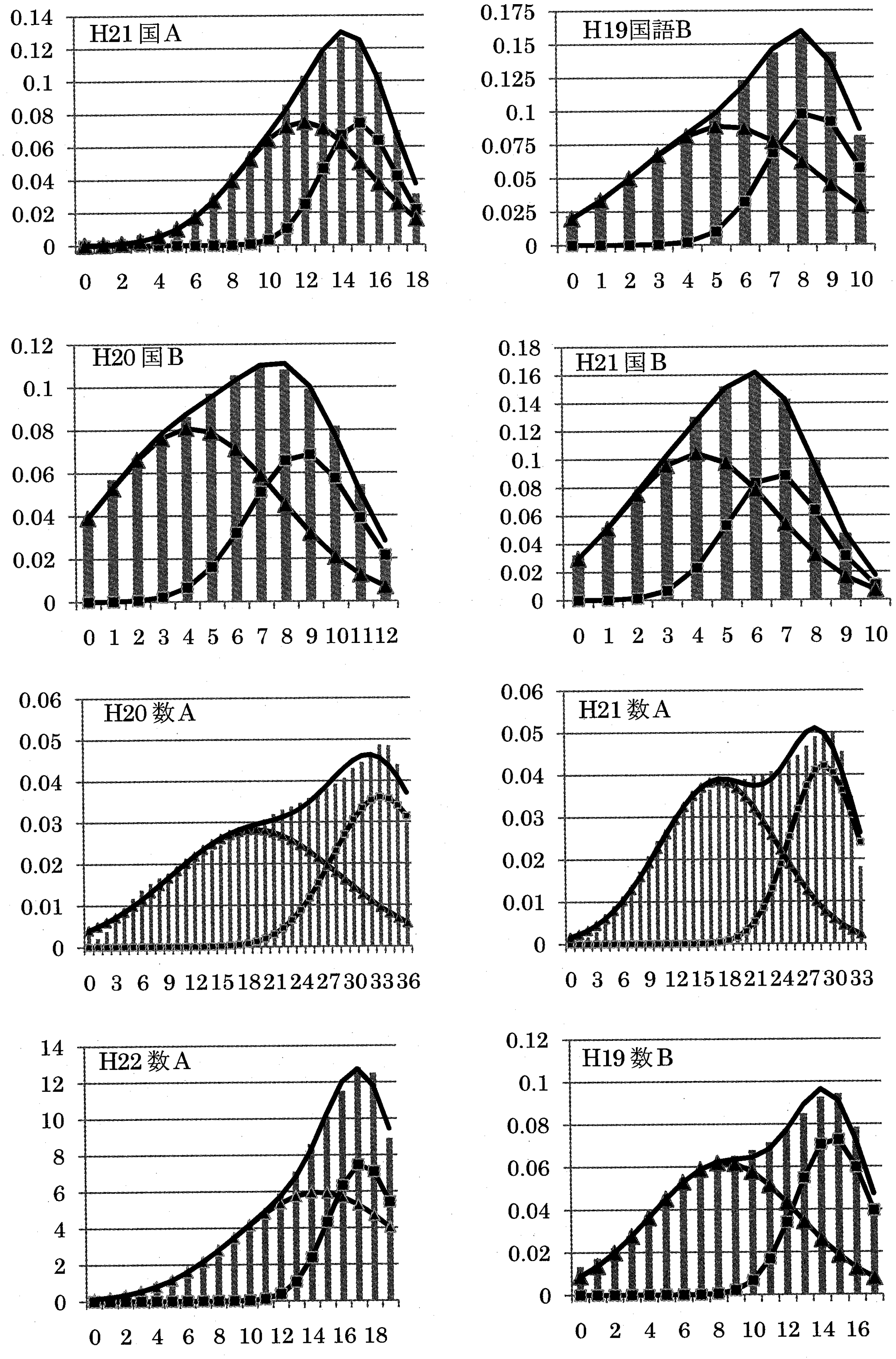

Figure 1　全国学力調査の学力分布とモデルの値（続き）

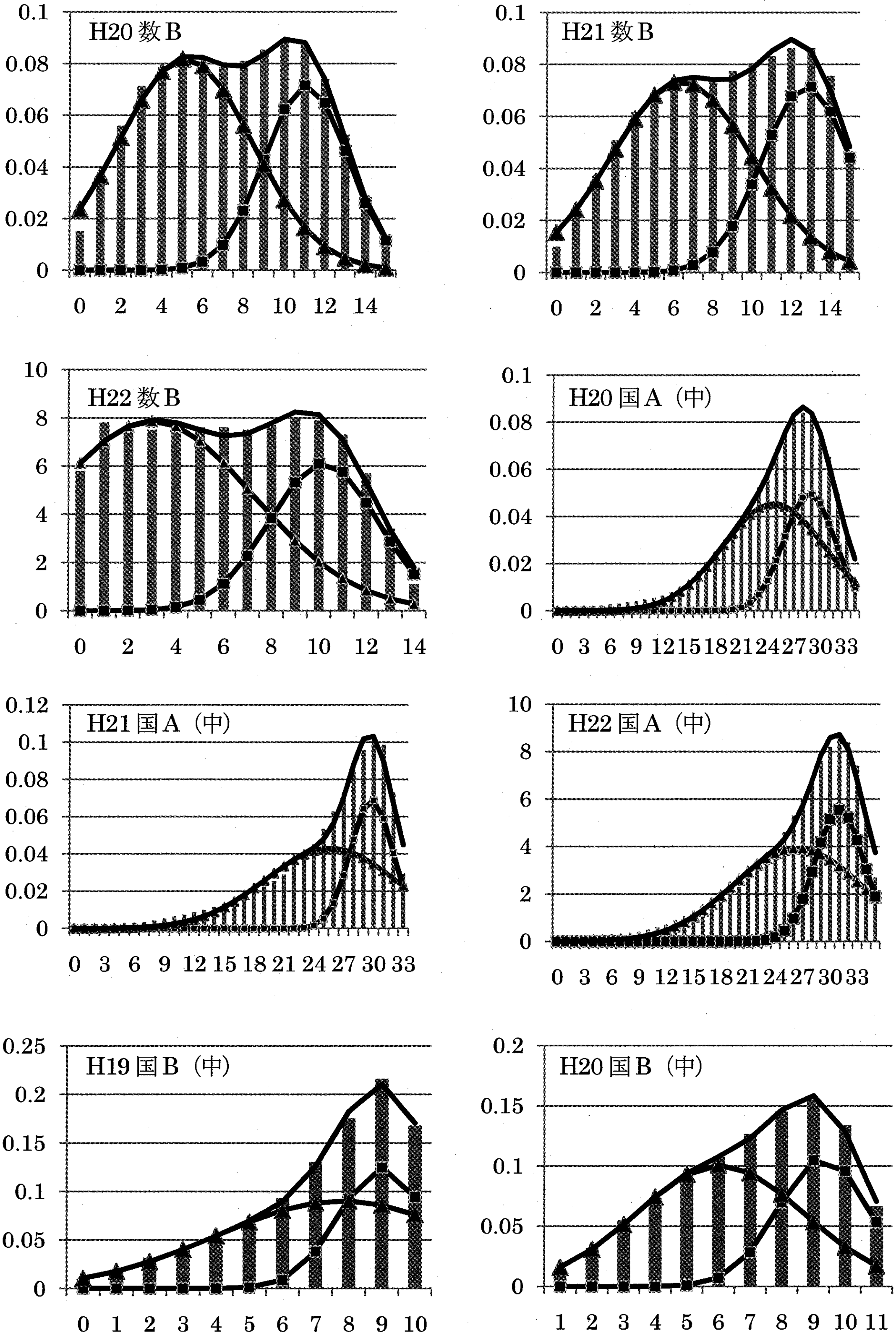

Figure 1　全国学力調査の学力分布とモデルの値（続き）

Table 3 都道府県別統一比率二極化モデル

都道府県	H群比率	小H群 平均	小H群 SD	小L群 平均	小L群 SD	中H群 平均	中H群 SD	中L群 平均	中L群 SD
北海道	0.22	11.23	1.65	7.54	2.83	14.61	1.72	8.64	4.53
青森	0.31	11.58	1.58	8.58	2.70	14.63	1.93	8.95	4.10
岩手	0.22	11.38	1.52	8.37	2.67	14.50	1.63	8.74	3.96
宮城	0.24	11.30	1.61	7.95	2.77	14.48	1.72	8.97	4.22
秋田	0.28	12.12	1.44	8.93	2.59	14.71	1.68	10.26	4.00
山形	0.32	11.24	1.71	8.23	2.71	14.60	1.81	9.52	3.90
福島	0.25	11.59	1.30	8.22	2.84	14.65	1.72	9.04	4.23
茨城	0.27	11.54	1.61	8.03	2.96	14.50	1.84	8.75	4.21
栃木	0.23	11.46	1.47	8.04	2.85	14.69	1.58	9.34	4.27
群馬	0.27	11.42	1.59	8.03	2.75	14.82	1.63	9.78	4.31
埼玉	0.27	11.44	1.68	8.20	2.92	14.61	1.81	8.91	4.45
千葉	0.29	11.75	1.63	8.40	2.96	14.60	1.85	8.84	4.32
東京	0.29	11.90	1.60	8.53	3.04	14.66	1.75	8.96	4.46
神奈川	0.27	11.75	1.65	8.12	3.05	14.77	1.69	9.05	4.48
新潟	0.25	11.51	1.47	8.17	2.68	14.44	1.80	9.35	4.29
富山	0.35	11.65	1.45	8.37	2.75	14.88	1.60	10.00	4.36
石川	0.34	11.30	1.77	8.19	2.74	14.77	1.80	9.75	4.13
福井	0.31	11.70	1.57	8.85	2.76	15.16	1.53	10.71	3.95
山梨	0.32	11.21	1.71	7.70	2.69	14.60	1.84	8.92	4.02
長野	0.26	11.56	1.56	8.27	2.72	14.57	1.68	9.37	4.20
岐阜	0.32	11.53	1.70	8.21	2.82	14.88	1.69	10.05	4.22
静岡	0.32	11.37	1.62	7.94	2.77	14.68	1.80	9.57	4.22
愛知	0.32	11.66	1.63	8.19	3.00	14.92	1.69	9.64	4.31
三重	0.23	11.36	1.58	7.96	2.92	14.78	1.59	9.45	4.58
滋賀	0.25	11.68	1.51	8.02	2.84	14.67	1.78	8.73	4.57
京都	0.29	11.89	1.48	8.49	2.89	14.71	1.87	9.08	4.42
大阪	0.22	11.64	1.55	7.90	3.06	14.53	1.74	7.91	4.92
兵庫	0.30	11.68	1.62	8.09	2.94	14.72	1.78	9.01	4.57
奈良	0.29	11.77	1.58	8.18	2.92	14.76	1.75	9.25	4.56
和歌山	0.24	11.64	1.70	8.09	2.99	14.67	1.72	8.61	4.57
鳥取	0.33	11.49	1.64	8.14	2.68	14.60	1.88	8.73	4.20
島根	0.24	11.24	1.69	8.22	2.94	14.45	1.82	9.54	3.94
岡山	0.27	11.21	1.78	7.95	2.97	14.66	1.83	8.65	4.48
広島	0.29	11.73	1.61	8.33	2.85	14.51	1.87	8.95	4.40
山口	0.31	11.18	1.61	7.75	2.84	14.57	1.77	9.22	4.34
徳島	0.27	11.72	1.59	7.67	2.75	14.72	1.71	9.51	4.65
香川	0.29	11.92	1.36	8.77	2.84	14.83	1.68	9.60	4.61
愛媛	0.27	11.49	1.53	8.22	3.09	14.93	1.64	9.66	4.33
高知	0.16	11.63	1.26	8.17	2.98	13.86	1.88	7.50	4.75
福岡	0.26	11.34	1.62	7.89	2.86	14.42	1.78	8.71	4.45
佐賀	0.25	11.47	1.60	7.94	2.76	14.44	1.83	9.25	4.24
長崎	0.31	11.24	1.73	7.75	2.80	14.52	1.83	9.26	3.98
熊本	0.25	11.54	1.53	8.20	2.76	14.84	1.63	9.81	4.17
大分	0.25	11.34	1.61	7.76	2.79	14.35	1.89	8.92	4.25
宮崎	0.27	11.52	1.55	8.07	2.75	14.59	1.77	9.65	4.02
鹿児島	0.22	11.49	1.50	8.00	2.75	14.38	1.84	8.93	4.08
沖縄	0.12	10.98	1.75	7.22	3.04	13.86	1.69	6.99	4.12

5．H群，L群の意味づけ―都道府県別統計指標との相関

本節においては，H群，L群とはどのようなものなのかについての社会学的，教育学的，経済学的意味づけについて述べる．そのために県別のさまざまな統計指標を収集し，統計指標とH群，L群の各種値との相関係数を求め，H群，L群の意味づけを行った．ここで用いた統計は，文部科学省の学校基本調査と内閣府統計局の「統計でみる都道府県のすがた」(2010）である．

ここでは学校基本調査の報告書のデータを元にして，親世代の大学進学率（短大進学を含む）とH群比率等との間にどんな関係があるかを探る．ここで用いたのは昭和55年度，昭和60年度，平成2年度の高校卒業者が大学・短大に進学した比率である．これらの年度の男女込の進学率，男子のみの進学率，女子のみの進学率，9種類（=3年分x 3種）を分析に用いた．Table 4参照．

表をみるとH群比率と親世代の男子の進学率との相関が高いことに気づく．これ対して親世代の女子の進学率との関連はあまり高くない．男女込の進学率はこの中間の関連を示している．L群比率との相関は表中に示されていないが，H群比率との相関の符号を負にしたものになっている．

親世代の男子の進学率は中学校のH群平均とも密接な関連があるが，小学校のH群比率との関連は薄い．これに対して親世代の女子の進学率は小学校のH群平均との関連が高く，中学校のH群平均との関連は薄い．

中学校L群の平均は親世代の男子の進学率と関連がある．これはどういうことだろうか．父親の進学率，H群比率，中学校H群平均，中学校L群平均の間には，パス解析的に考えて，この順に単調な直列モデル（シリアルモデル）が成立している．つまり親の進学率がH群比率に正の影響を与え（パス係数0.46)，H群比率が中学校H群平均に正の影響を与え（同0.68)，中学校H群平均が中学校L群平均に正の影響を与えている（同0.80)．つまり学級内に一定の割合のH群に属する生徒がおり，その生徒たちが学級内に学力形成に対して前向きな影響をあたえることによってL群の生徒たちの学習への意欲を喚起し学力を高めていると解釈できる．

吉川（2006，2009）は，日本人の階層帰属意識には，親が大卒（短大を含まない）であるかどうかということが成熟型学歴社会となった日本において決定的であると主張する．現在の小学6年生，中学3年生の父親の大学進学率は36%である．

さまざまな経済指標との関連について結果のみを要約

Table 4　都道府県別二極化モデルの値と親世代の大学進学率との相関

			小H群平均	小H群SD	小L群平均	小L群SD	中H群平均	中H群SD	中L群平均	中L群SD	H群比率
昭和55年	男子	大学短大進学率	0.26	0.17	0.11	0.27	**0.47**	−0.01	**0.26**	0.35	**0.51**
		大学進学率	0.27	0.17	0.11	0.36	**0.47**	−0.04	**0.27**	0.37	0.46
		短大進学率	−0.01	0.04	0.02	−0.30	0.04	0.10	−0.02	−0.04	0.23
	女子	大学短大進学率	**0.36**	−0.04	0.06	**0.49**	0.19	0.09	−0.07	**0.61**	0.19
		大学進学率	**0.39**	−0.02	0.18	**0.47**	0.29	0.02	0.02	**0.58**	0.27
		短大進学率	0.29	−0.05	−0.04	0.43	0.09	0.13	−0.11	0.52	0.10
昭和60年	男子	大学短大進学率	0.14	0.20	0.02	0.25	**0.43**	−0.06	**0.24**	0.27	**0.44**
		大学進学率	0.15	0.18	0.04	0.36	**0.43**	−0.11	**0.28**	0.26	0.36
		短大進学率	0.00	0.05	−0.05	−0.26	0.02	0.13	−0.08	0.03	0.22
	女子	大学短大進学率	**0.35**	0.01	0.02	**0.56**	0.21	0.05	−0.07	**0.60**	0.18
		大学進学率	**0.44**	−0.04	0.23	**0.49**	0.35	−0.05	0.09	**0.54**	0.29
		短大進学率	0.22	0.04	−0.13	0.50	0.06	0.12	−0.17	0.52	0.06
平成2年	男子	大学短大進学率	0.11	0.15	0.08	0.03	**0.56**	−0.16	**0.47**	0.10	**0.51**
		大学進学率	0.12	0.13	0.07	0.15	**0.53**	−0.20	**0.46**	0.13	0.41
		短大進学率	−0.03	0.08	0.04	−0.38	0.12	0.10	0.05	−0.09	0.33
	女子	大学短大進学率	**0.31**	−0.01	0.06	**0.37**	0.34	−0.01	0.13	**0.50**	0.30
		大学進学率	**0.38**	−0.03	0.21	**0.38**	0.42	−0.11	0.22	**0.43**	0.36
		短大進学率	0.19	0.00	−0.06	0.28	0.21	0.06	0.04	0.44	0.20

しておくと，H群の一般的特徴は，親が大卒であり，経済的に余裕があり，新聞を定期購読するなど一定の文化的環境が整っており，核家族傾向があり，給与所得者の階層が主である．L群は経済的，家庭的に困難な階層，同居親族のいる家庭，離婚，失業と関連が深い．Table 5 参照．表中の相関は0.29以上が 5 %水準で有意である．

このような要約に加えて，L群標準偏差について補足説明をしておこう．公民館の数や図書館の数が多い県では，L群標準偏差が小さくなる傾向が見て取れる．また県の教育費負担を増やしたり社会教育費を増やしたりすると，L群の標準偏差が小さくなる傾向も見て取れる．社会教育を充実させるとL群の下方への広がりを押さえる働きがあるのであろう．保健師数の効果も同様と思われる．つまり学校教育や社会教育，保健衛生を充実させると，L群のばらつきを小さくし，社会の刑法犯の発生を抑制する効果が期待できることが示唆される．

このことを共分散構造分析を用いて確かめておこう．Figure 2 参照．図には 2 つの潜在変数「L群の背景的要因」と「L群のばらつき」があり，前者の観測変数としては，公民館数，図書館数等の行政的措置によって改善できる指標が置かれており，後者には小学校と中学校のL群標準偏差が置かれている．さらに「刑法犯認知件数」に対して「L群のばらつき」からパスが引かれており，パス係数は0.97，また「L群の背景的要因」から「L群

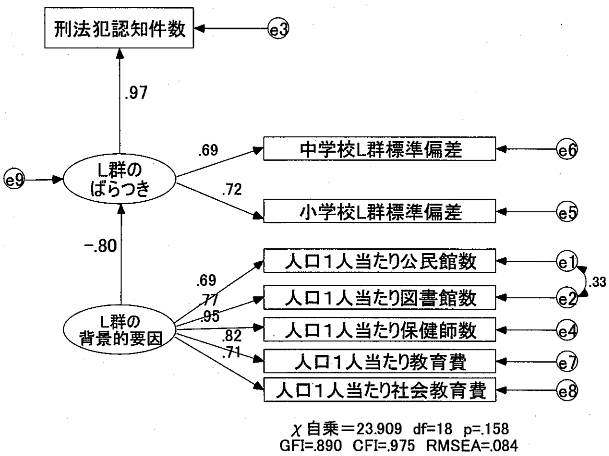

Figure 2　刑法犯認知件数とL群

Table 5　都道府県別二極化モデルの各種値と経済指標等との相関

	小 H群 平均	小 H群 SD	小 L群 平均	小 L群 SD	中 H群 平均	中 H群 SD	中 L群 平均	中 L群 SD	H群 比率
県民所得	**0.35**	0.02	0.21	0.30	**0.42**	−0.17	0.18	0.32	0.33
個人預貯金残高	**0.46**	−0.07	0.25	0.32	**0.43**	−0.21	0.20	0.58	0.31
郵便貯金残高	**0.36**	−0.03	0.18	0.24	**0.47**	−0.12	0.27	0.52	0.35
核家族世帯割合	**0.32**	−0.10	0.15	0.10	**0.41**	0.11	0.26	0.40	0.33
一人あたり新聞	**0.26**	0.10	0.10	0.32	**0.32**	−0.03	0.04	0.45	0.32
第2次産業就業者比率	0.06	−0.07	0.21	−0.23	0.56	**−0.34**	0.55	−0.07	0.45
第3次産業就業者比率	0.04	**0.22**	−0.21	0.53	−0.35	**0.23**	−0.51	0.33	−0.28
共働き世帯割合	−0.17	−0.05	**0.23**	−0.61	0.20	−0.23	**0.50**	−0.58	0.26
65歳以上親族世帯割合	0.00	−0.24	**0.33**	−0.64	0.13	−0.03	**0.47**	−0.43	0.22
離婚率	−0.13	0.07	**−0.45**	0.45	−0.43	0.11	**−0.68**	0.32	−0.51
完全失業率	−0.14	−0.04	**−0.39**	0.27	−0.60	0.13	**−0.69**	0.20	−0.60
借家比率	−0.12	0.20	**−0.37**	0.54	−0.35	0.17	**−0.55**	0.30	−0.31
刑法犯認知件数	0.24	0.04	−0.08	**0.69**	0.06	0.00	−0.32	**0.67**	−0.08
企業倒産件数	0.30	0.06	0.08	**0.46**	0.05	0.00	−0.21	**0.39**	0.01
生活保護被保護実世帯数	0.07	−0.10	−0.25	**0.31**	−0.46	0.31	−0.60	**0.42**	−0.45
小学校長期欠席児童比率	0.11	0.06	−0.32	**0.52**	−0.22	0.14	−0.56	**0.61**	−0.31
中学校長期欠席児童比率	0.24	−0.03	−0.11	**0.50**	−0.06	0.15	−0.39	**0.64**	−0.14
公民館数（人口あたり）	−0.10	−0.03	0.12	**−0.46**	0.02	−0.04	0.23	**−0.33**	0.11
図書館数（人口あたり）	−0.02	−0.11	0.20	**−0.45**	0.03	−0.04	0.21	**−0.30**	0.19
教育費（人口あたり）	−0.19	−0.10	0.01	**−0.46**	−0.28	0.06	0.03	**−0.48**	−0.21
社会教育費（人口あたり）	−0.14	−0.03	0.22	**−0.38**	0.10	−0.07	0.37	**−0.51**	0.20
保健師数（人口あたり）	−0.28	−0.12	0.03	**−0.54**	−0.13	0.01	0.18	**−0.46**	−0.03

のばらつき」には－0.80のパスが引かれている．つまりL群の形成を改善できるような予算措置（言い換えると学校教育を下支えし学業不振児の発生を未然に防ぐ社会的な環境改善）を講ずれば，かなり大幅にL群のばらつきを小さくすることができ，これを改善すれば刑法犯の発生を劇的に改善できると予想される．

ただしL群発生を防ぐ行政的予算措置だけでは限界があり，小中学校における学力のばらつきを小さくするような教育的努力を介在させることによって刑法犯の発生を減少させることができるのである．

県別の二極化モデルのH群，L群の比率，標準偏差と算数B，数学Bの成績との相関およびこれまで触れてこなかったいくつかの統計指標との相関について簡単に触れておこう．L群標準偏差の指標は算数B，数学Bの標準偏差とも同じような相関がある．L群標準偏差と算数B，数学Bの標準偏差の相関は0.8以上あるので，これはある意味で当然である．分布の左側への広がりが，刑法犯認知件数等と関連があるという指摘は，どちらの標準偏差を用いても指摘可能である．しかしながら，その意味づけはL群との関連で論じた方が説得的であろう（算数B，数学Bの標準偏差への言及は苅谷（2009，p118）を参照のこと）．

これら以外に男女の平均身長は学力と正の相関がある．学力には生物学的基盤があることが示唆される．死産率，2500g未満の出生率は学力と負の相関，一般病院外来患者数（常勤医師１人１日当たり）は学力と正の相関がある．これらは学力の高い県，あるいは教育力の高い県では死産や未熟児の出生を低め，病気のさいの通院率を高める傾向を生むと言える．学力調査を受けている子どもの学力の背景にある教育環境が，健康に関する「学習の社会的成果」（OECD教育研究革新センター，2008，p149）をもたらしているのだろう．また学力は年平均気温とも関連があり，暖かい県ほど学力が低い傾向がある．これらは単に疑似相関である可能性もあり，今後の詳細な分析が必要である．

６．二極化はどの学年から始まっているか－特定の課題に関する調査（国語，算数・数学）への適用

全国学力調査のデータ分析によって，小学校６年生と中学３年生の学力データには二極化の特徴が顕著であり，H群比率は両者で一定であることを見てきた．ここでは二極化の傾向が学年的にいつ頃から始まるのかという問題を分析する．分析の対象としたのは国立教育政策研究所が2005年に公表した「特定の課題に関する調査（国語，算数・数学）」のデータである．これは調査対象学年が小学校４年生から中学３年生で，科目としては国語の漢字の読みと書き，算数・数学の計算に関する力，数学的に考える力の４科目の各学年の得点分布である．国語の問題は易しすぎて，本来は二極化モデルの適用がかなり困難な分布をしているものが多かったが，各学年内において４科目間でH群比率一定の制約条件を置き，学年ごとのH群，L群比率を求めてみた．

結果をFigure 3に示す．横軸は学年，縦軸はH群の比率を示している．これによるとH群比率は小４から中３までほぼ一定で約４割である．小学校の４年生からすでに二極化を確認でき，その得られた比率も先のデータとおおよそ同じである．H群比率がやや高め，L群比率がやや低めではあるが，問題の難易度の影響も多少あるのかもしれない．

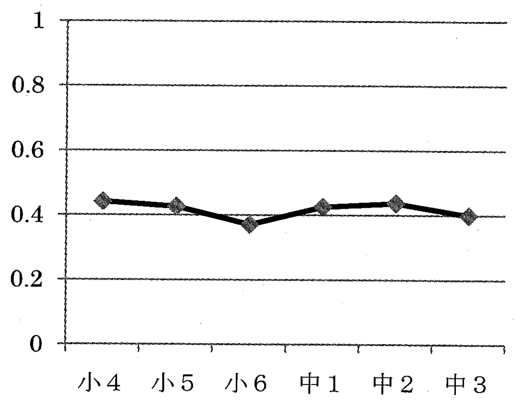

Figure 3　特定の課題に関する調査におけるＨ群比率

７．二極化は年代的にいつから始まっているか－高知県の高校入試データの経年変化

ここでは二極化がいつ頃の年代から始まったのかを調べる．分析対象としたのは，高知県公立高校の入試データで，昭和53年から平成11年までの英語，数学の２科目の得点分布である（高知県教育委員会から提供を受けた）．分析のモデルとしては，これまでと同じく，２科目間でH群比率が一定という制約をもつ二極化モデルである．

結果をFigure 4に示す．横軸が年度，縦軸がH群比率である．これを見ると，単年度ごとにはH群の比率が上下しているが，全体としては昭和53年度入試からすでに二極化が始まっており，H群の比率は５割程度であることが見て取れる．つまり学力の二極化は，本論文のように定義すれば，最近起こったことではなく，少なくとも30年くらい前から始まっていたことが示唆される（苅谷，1995）．ここで取り上げたデータは悉皆のデータではなく，入試データであり，高校へ進学しない子どもが一定の率で存在すること，また高知県の特殊事情として，中学段階で私立へ抜ける子どもが多いので，比率そのもの

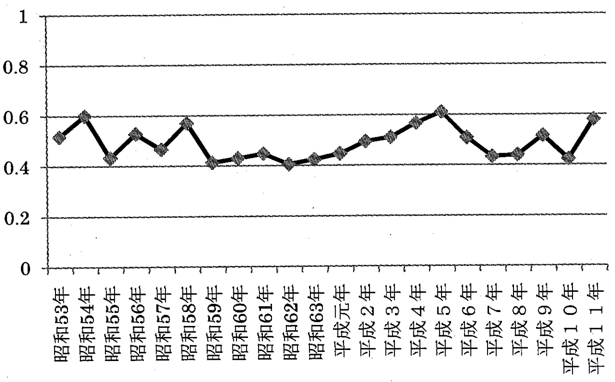

Figure 4　高知県高校入試におけるＨ群の経年変化

は，高知県全体の傾向を反映していない可能性がある．

8．総合的考察

本論文で提案したモデルの枠組みを確認しておく．

（Ⅰ）正規分布していないと思われる学力分布を二つの正規分布の混合によって表現した．

（Ⅱ）複数の学力分布にこれを適用するときに，分布間にかかわる制約条件の分析を行った．

（Ⅲ）県別の学力分布データに（Ⅱ）で得られたモデルを適用し，県別データから得られたモデルの各値（Ｈ群比率，Ｈ群，Ｌ群の平均，標準偏差）と他のさまざまな県別統計指標との相関を調べることにより，モデルの各値の意味づけを行った．

まず（Ⅰ）においては正規分布の形を決める２組の平均と標準偏差および２つの分布の面積比を求めた．この値を求めるために，(2)（3）式を用いた．（Ⅱ）における制約条件分析において，本モデルの場合はすべての分布に共通な面積比が得られたが，他のデータに適用する場合，これとは異なる制約条件が採用されることはありうるだろう．（Ⅲ）においては県別の学力分布データが入手可能であったが，このようなデータが容易に入手可能でない場合，（Ⅲ）の分析は困難となる．

さて本論文では学力の二極化を定義し，その定義にもとづいてH群比率一定という制約条件付き二極化モデルを構築した．単に学力分布が双峰のときに二極化というのではないという主張は十分に裏付けられたのではないだろうか．

H群・L群の意味づけ　なぜこのような分析が可能だったのだろう．まず全国学力調査の全体データにモデルを適用し，いろいろな制約条件の中で，どれが最適かを探り，面積比が一定の２つの正規分布の混合モデルを選び出した．モデル比較の手法により，分布の背後に，なにか「硬い」実体があることを抉りだした．

次に全体データで開発されたモデルを全国学力調査の各県データに適用，各県の「硬い」実体，すなわちH群，Ｌ群の比率，平均，標準偏差を求め，これら各県に関するモデルの値と各県の大学等への進学率や経済指標との関連を探った．その結果，各県のＨ群の比率等は親の進学率や経済的豊かさと密接な関連をもつことが明らかになった．つまり学力分布から求めた，分布を決める値が学歴・経済力によって決定されることが見えてきた．

通常の社会学的な手法だと，個人（子ども）に着目し，その親の階層（学歴，年収，社会的地位）の指標を得，子どもの学力，学習時間等との関連を平均値をもとに分析する．これは個人に着目する極めてオーソドックスな手法であり，いわばミクロな手法である．それに対してここで行った手法は学力分布から出発するというマクロな手法である．これはミクロ経済学的手法vs. マクロ経済学的手法に類比的かもしれない．本手法は個人には一切注目しておらず，誰がH群に属し，誰がL群に属するかも分からない．分かるのは，例えば平成21年度の数学Bである得点を取った比率のうち，何%がH群の子どもたちのもので何%がL群の子どもたちのものであるかということだけである．これはその得点のところで縦線を引き，H群とL群の折れ線の高さを測ることによって可能である．

このような手法が可能となったのは，現象の背後に扁平な分布，二山の分布を作り出す実体があったからであろう．男女込の身長分布の背後に男子と女子という平均と標準偏差を異にする実体がある．回遊魚の身体計測をすると二山の分布となることが知られている．それは回遊歴が１周目の回遊魚と２周目の回遊魚が混じり合っており，それぞれが異なる平均と標準偏差をもっているので，ここで展開したモデルと同じ分析をすると，１周目と２周目の魚の大きさの平均と標準偏差が分かるのである．学力もこれらと同じように，背後に２つの集団があったので，このような分析が可能となったのである．

１次元のデータに正規混合分布モデルを適用するさいの１つの弱点は，いろいろと学力データを分析してきた筆者個人の経験によると，１つの分布に対して最適な近似を与える解はたしかに一意に定まるが，その最適解以外にもよい近似を与える解はいくらでもあるということである．全国学力調査のデータ数は膨大であるので，非常になめらかな分布をしているが，それでも，ひとつひとつの分布に対して，Ｈ群の面積比が0.1くらいから0.4くらいまでの値に対して，それなりによい近似を与えてしまう．この難点を克服し，また背後の実体に迫る糸口を与えたのが制約条件分析であった．分布間に制約を与えて正規混合分布モデルを適用すれば，モデル適合の観点から最良モデルを導くことが可能となった．これによって誤差にとらわれず，分布の背後にある真相にうまく迫れたのではないだろうか．

このようにして開発されたモデルを県別に適用することによって，相関分析に持ち込むことができた．学力問題の特徴の１つは，地域間格差が大きいことだといわれる（苅谷，2008b）．むしろこれを逆手にとって，そのような学力の地域間格差と地域の経済力，進学データとの関連から，地域間格差として生じる学力格差，それの背後にあるＨ群，Ｌ群の実体が明らかになった．

学力はなぜ二極化するのか　次に学力分布はなぜ二極化するのかについて，Figure 5 を用いて仮説的に素描してみたい．学力形成に必要な生得的能力は正規分布していると考えられる．つまり生まれた時の能力は二極化していないが，生まれた後の家庭環境に関しては家庭教育あるいは学力形成に熱心な群と熱心ではない群があると考えるのである．前者は結果としてＨ群となり，後者は結果としてＬ群になる．前者は学力形成に有利な文化的環境にある群と言い換えてもよい．

家庭教育に熱心な群の生得的能力の分布は正規分布をしており，当然ながら全体の正規分布の中の部分集合をなしている．Figure 5 （a）参照．ここでは全体の分布とＨ群の分布を描いている．このとき全体分布からＨ群の分布を差し引いた分布がほぼ正規分布となることがある．Figure 5 （b）参照．これは厳密には正規分布ではない．どういうときに正規分布に近い分布をするかという問題は詳細な分析を必要とするが，ここではFigure 5 （b）のような場合が多くあることを指摘しておく．

家庭教育に熱心な群の子どもたちも，熱心ではない群の子どもたちも学校教育によって学力が形成されていき，横軸に，ある学年における，ある教科の得点をとり，学力分布を描くと，家庭教育に熱心な群の子どもたちは，家庭教育に熱心ではない群の子どもたちよりも家庭教育力が大きく作用するので，両群の分布は，生得的能力の分布が両群を合わせたときに正規分布をしていたのとは異なり，家庭教育に熱心な群の分布が熱心ではない群の分布に比べて右側（学力上位の方）に平行移動し，結果として二山のグラフが描かれることとなる．Figure 5 （c）参照．もちろん平行移動量が小さいときは，一山の扁平なグラフとなる．

ただし家庭教育に熱心な群の子どもたち全員が，熱心ではない群の子どもたちに比して単純に右側に平行移動しているのではなく，家庭教育に熱心な群の平均が熱心ではない群の平均と比べてより右側に移動してはいるが，個々の子どもたちの移動量はさまざまであるはずである．同様にＬ群の子どもたちにおいても生得的能力を横軸にとったときと学力を横軸にとったときとでは，順位が大幅に入れ替わっていてもなんら差し支えない．家庭教育力による差が加えられることによって，学力の二極化が生じるのである．

学力の二極化が年齢とともに進行することについても同様の考え方で説明できる．つまり年齢とともに家庭教育の力が蓄積され平行移動の量が増加していくのである．

数学的に言えば，生得的能力に関して家庭教育に熱心な群の分布がFigure 5 （a）（b）に描かれているほど，全体の分布の中の上位に位置する必要はない．家庭教育に熱心な群の平行移動量が熱心ではない群の移動量よりも大きければよいのである．

多極化モデルの可能性について　ここまで正規分布を２つ用いることで，全国学力調査のほぼすべての学力分布を近似できることを見てきた．ここで問題を少し一般

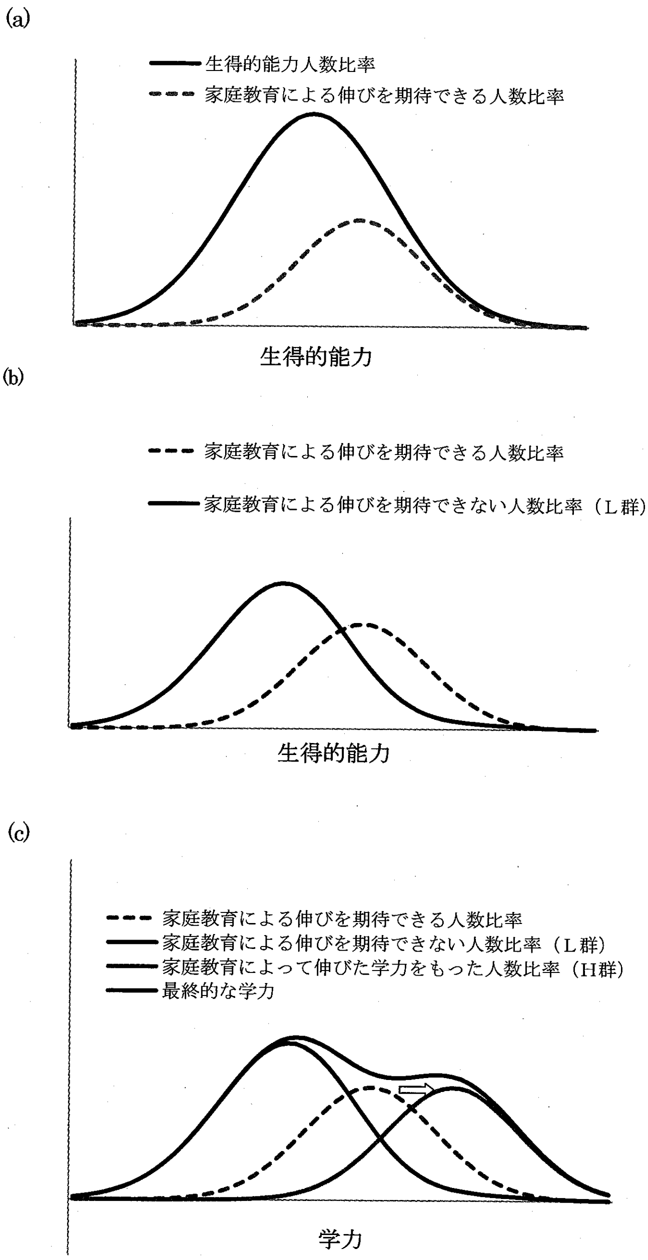

Figure 5　なぜ学力の二極化が起こるかの説明

化し，全国学力調査の学力分布を表現するモデルが3つ以上の正規分布の混合分布である可能性があるかどうかについて議論する．実際のところ，3つの正規分布を用いる三極化モデルを全国学力調査の分布に適用してみると良好な近似が得られる．これは，より多くのパラメタを用いるので当然の結果であり，BICを用いてモデル比較をしても，個別の学力分布に関しては二極化モデルよりも，よい値を示す場合もある．ただ平成19年度，平成21年度の小学校の国語の問題は10点満点の試験であり，これに8個のパラメタを用いて分布の近似を得ようとするのは，あまり意味がない．あてはまって当然だからである．ここではよりよい近似が得られるかどうかというよりも，学力分布の形状から，その背後にいかなる社会階層が存在しうるかという，ある種の数学的な問題について考える．

最初に確認しておくが，以下の議論は数学的なものであるが，全体の分布の成分となっている分布（例えばL群）に言及するとき，それは1つの社会階層に対応していることに留意されたい．

記述されるべき学力分布が今後，無限に産出されるとする．そのときその分布形にあらわれる特徴がこれまでの32個の分布形によってあらわれてきた特徴と基本的に同じであるとする．具体的には二山あるいは扁平な一山であり，その分布形は変曲点が高々4個であるという特徴を持ち続けるとする．ここで変曲点とは，数学的に定義すると曲率が＋から－（あるいは－から＋）に変化する点のことを言い，曲線の2階微分（差分）が0の点で定義される．たとえば，二山の分布の場合，左側から順に記述すると，まず下に凸の曲線が現れ，1つ目の変曲点を通過して左側の頂点に達する前に上に凸となり，頂点を超えると，2つ目の変曲点を通過して下に凸となり，3つ目の変曲点を通過して2つ目の頂点に達する前に再び上に凸となり，4つ目の変曲点を通過し，さらに下に凸となって曲線が終わる．すでに数えたように，この場合は4個の変曲点がある．一山の分布の場合，平成19年度の算数Aを例にとると，曲線は初め下に凸，次に上に凸となり，再び下に凸となって，さらに上に凸となって頂点を超え，さらに下に凸となって（ここはやや曖昧）曲線が終わる．この場合も変曲点は4個である．平成21年度国語Bの一山分布の場合，変曲点は2個である．

ただし分布をよく観察すると1点刻みで凹凸が細かく変化するものもある．例えば平成21年度中学国語A．これを厳密に数学的に記述すると変曲点が4個を超えるものがあることになる．しかしながらこれはグラフの全体的特徴というよりは，細かな得点のゆらぎとみるべきであろう．ここまでみいだされた全国学力調査の分布は基本的に変曲点が高々4個という特徴をもっているとみなせる．

今後5個以上の変曲点をもつ分布形が出現しないとは限らないが，もし変曲点が4個までなら，これを産み出す正規分布（すなわち社会階層）の個数は2個で十分である．言い換えると多極化モデル（正規分布を3個以上混合するモデル）は変曲点を5個以上もつ可能性があり，現実の全国学力調査の特徴とは異なる分布を産み出す可能性が大である．

では多くて4個の変曲点をもつ分布をつねに産み出す

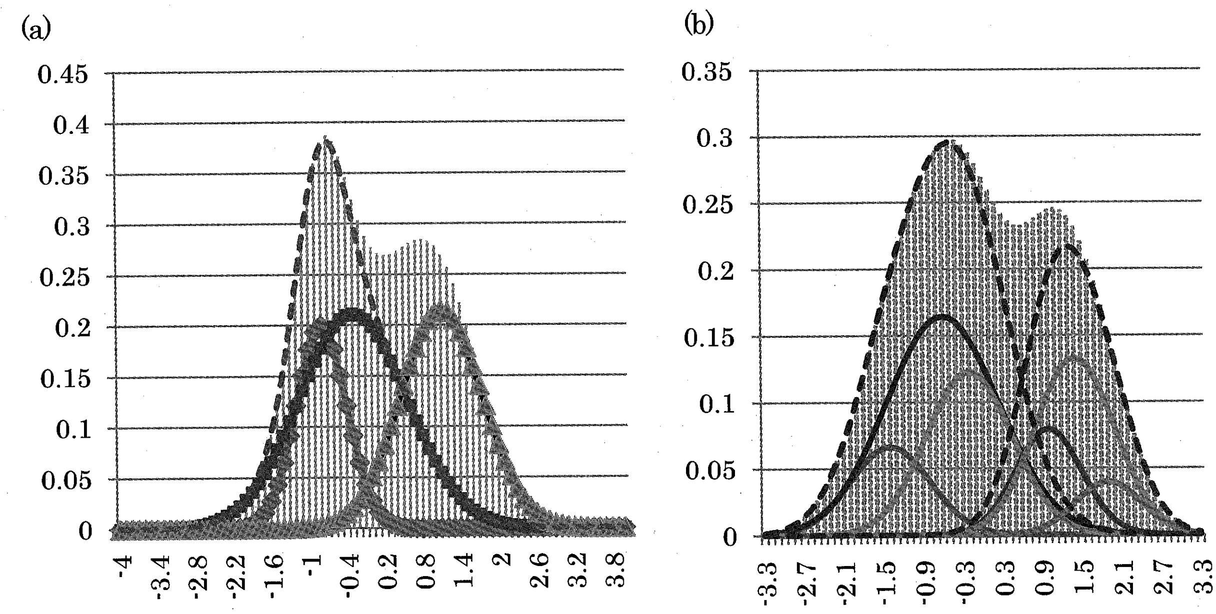

Figure 6　3つの正規分布が混合ししている例（左）と6つの正規分布が混合している例(右)

には，２つの正規分布の混合に限るかと言うと，必ずしも，そうは言えない．仮に３つの正規分布の混合を考えるとして，そのうちの２つの正規分布の平均値が標準偏差と比べて接近していて，その２つの正規分布の混合がつねに一山の分布をもたらすような場合，そのような３つの正規分布の混合は一山または二山の分布となる．４つの正規分布の混合も同様であり，そのうちの２つの正規分布が混合されて一山になるものどうしが混合されたときも，一山または二山の分布となる．Figure 6 (a) 参照．図中の点線の分布は左側にある２つの正規分布の混合である．これと右側の正規分布が混合されて，全体の分布が与えられている．

つまり全国学力調査の分布の特徴をもたらす混合分布は，必ずしも２つの正規分布とは限らず，平均値の近い３つまたは４つ，あるいはそれ以上の個数の正規分布の混合によっても，このような性質をもつ学力分布がもたらされる可能性がある．

ここまでの議論は学力分布を記述するモデルが正規分布を基本成分とする混合分布を考えてきた．ではモデルの基本成分が正規分布以外の分布のとき，上述の議論には変更が加えられるだろうか．基本成分となる分布（それに対応する社会階層）が左右両端で０に漸近的に近付き，真中辺（左右対称でなくてよい）に頂点をもつものならば，上記の議論はそのまま成り立つ（ただし三角分布のような不連続点をもつ分布は除く）．つまりこの分布形は２つの変曲点をもち，これを２つ混合することで，４個以下の変曲点をもつことになる．また３つを混合する場合は，そのうちの２つの混合が一山となる場合，４個以下の変曲点をもつことが予想される．じつはこの議論は上述の２つの正規分布が一山となる場合の議論の中ですでに部分的に展開している．つまり３つの正規分布のうち，２つの正規分布が混合されて変曲点を２個もつ扁平な分布となり，それがもう１つの正規分布と混合すると高々４個の変曲点をもつ分布を産み出す．

結論的に言うと，全国学力調査の学力分布に見られるような分布形を産み出す混合分布の基本成分がいくつであるかを，変曲点の個数に基づいて特定することはできないが，その基本成分が仮に多数あったとしても，つまりその基本成分となる分布に対応する社会階層が多数あったとしても，それらをすべて混合する前段階で，必ず２つの分布の混合として表現できる．その分布は正規分布であるかもしれないし，２つの正規分布を混合した扁平な一山分布であるかもしれないし，変曲点を２個もつ，それ以外の分布であるかもしれない．その２つの分布形を混合することで全体の分布を表現できるのである．本モデルにおいては，その２つの分布形がともに正規分布であると仮定したが，理論的には，基本成分となる分布が正規分布以外の可能性はある．

つまり多様な社会階層が日本に存在し，それらが重なり合って学力分布が決定されているとしても，現実の学力分布が二山か一山である場合，基本成分が正規分布であるかそれ以外の分布が適切であるかを決定することはできないが，それを２つの分布の混合モデルで表現できるのである．

ではある社会階層があったとして，その階層がL群とH群の分布を横断するような分布となることがあるだろうか？おそらくない．その社会階層はL群またはH群内に収まるような分布をしているだろう．Figure 6 (b) 参照．この図は仮想的な学力分布である．「全体」が二極化した学力分布で，Ｌ群とＨ群の混合によって「全体」が近似されている．両群は点線で表示されている．さらにこの図ではＬ群の分布が分布１，分布２，分布３の混合，Ｈ群の分布が分布４，分布５，分布６の混合となっている．このように実際には６つの分布の混合となっている場合でも，上で説明したように，二極化グラフとみなすことができる場合がある．

社会階層が３つあって，そのうちの分布１と分布２がL群，分布３がH群を構成しているとしよう．このとき，誤って分布１のみをL群，分布２と分布３をH群とグループ化してしまうことはないのだろうか？おそらくない．本論文で展開しているモデルにおいては，多数の分布間に共通して成り立つ制約条件を求めている．単独の分布において，誤ったグループ化をすることはありうるが，多数の分布間においては，正しいグループ化が採用されるはずである．ただし社会階層が３つあることが既知なら，三極化モデルを構築すべきであろう．

日本には何らかの指標でそれを同定できる，さまざまな社会階層があるとしても，全国学力調査にあらわれている特徴から言えることは，そのような社会階層が他の社会階層と統合されてL群またはH群となっているということである．言い換えれば，それぞれの社会階層は仮にそれが存在するとしてもL群，H群の下位集合とみなすことができる．

橘木（2010）は学力が高校卒業時において三極化していると主張している．つまり高卒あるいは専門学校に進学する群（おおよそL群に対応）と大学に進学する群（おおよそH群に対応）があり，さらに後者が二群に分かれて，難関大学へ進学する群とそうでない一般大学進学者の群になっているとされる．これらは上の議論から明らかなように，H群の下位集合とみなせるだろう．ただし高卒層がL群，大卒層がH群という対応は，おおざっぱなもので，L群の中にも大学に進学する者は一定割合いるはずであるし，H群の中にも大学に進学しない者がいるだろう．橘木の指摘は正しいかもしれないが，それ

は小中学生を対象とした全国学力調査の学力分布からは見てとることができない．

最近，大学入試センター試験のある教科の得点分布が双峰分布していると報道された（2010年10月25日，朝日新聞）．これは，中学時代までH群に属していたものがさらに2つの下位集合に分かれた事態である可能性と，中学時代までL群に属していたものが，センター試験に参入するようになり，試験の成績が双峰分布するようになった可能性とがある．この点については，データを入手できれば検証が可能である．今後の研究テーマとしたい．

二極化モデルの精度　本論文で展開した二極化モデルは基本成分として正規分布を想定している．この想定自体は自明のものではないが，H群，L群の社会階層的な意味づけが適切になされているので，想定は適切だったのではないかと思われる．しかしながら，前段までの議論をふまえ，ここでは二極化モデルの評価をしてみることにする．

もし全国学力調査の分布は3つの正規分布の混合と理解するのが正しく，二極化モデルは，これを粗く近似しているにすぎないとしたらどういうことになるだろうか．これは上で述べたように，3つの分布のうちの隣り合う2つが混合されて正規分布からは多少なりとも逸脱した1つの分布となり，残りもう1つは正規分布のままの分布となり，これらが混合されて全体分布を形成している場合である．ここでは敢えて2つの分布の混合が一山とはなっているが正規分布からはかなり逸脱した形となるものを選んで数値シミュレーションを試みたところ，残念ながら，1つの正規分布の方の面積比はただしく再現されず，もとの3つの分布の面積比からは，かなり離れた値が計算された．これは2つの正規分布の混合が正規分布から逸脱しており，これを正規分布を用いて近似しようとしたことから生じた誤差である．もちろん2つの正規分布の混合がほとんど正規分布となることも多く，その場合はほぼ正しく面積比が計算される．

次に基本成分が正規分布以外の分布となる場合を考えてみよう．この場合も上の議論と同様，それぞれの分布が正規分布から逸脱する程度が大きくなるにつれて，計算される面積比も正しい面積比から逸脱する．

身近に存在するものごとは，身長の分布のようにきれいな正規分布をするものもあるが，必ずしも正規分布をするとは限らない．しかしこういった多くのものの分布は，正規分布に近い形をしているのも事実である（また対数変換をすると正規分布するものもある）．正規分布から多少逸脱している現象は，詳細な分析をすると，ここで展開した正規混合分布をしているものが数多くあると思われる．それらの分布の場合は基本成分として正規分布を想定するものが適切であるということである．全国学力調査の分布も，そのようなものの1つであり，基本成分が正規分布である蓋然性は高いと思われる．

文献

Arbuckle, J. L.（2003）Amos 5.0 update to the Amos user's guide.　SmallWaters Corporation

苅谷剛彦　1995　大衆教育社会のゆくえ―学歴主義と平等神話の戦後史　中央公論社

苅谷剛彦　2001　階層化日本と教育危機―不平等再生産から意欲格差社会　有信堂

苅谷剛彦　2008a　教育改革の迷走　筑摩書房

苅谷剛彦　2008b　学力と階層―学力の綻びをどう修正するか　朝日新聞出版

苅谷剛彦　2009　学力調査と格差問題の時代変化　In 東京大学学校教育高度化センター編　基礎学力を問う―21世紀日本の教育への展望　Pp.81-130.　東京大学出版会

苅谷剛彦・志水宏吉　2004　学力の社会学―調査が示す学力の変化と学習の課題　岩波書店

苅谷剛彦・志水宏吉・清水睦美・諸田裕子　2002「学力低下」の実態　岩波書店

吉川徹　2006　学歴と格差・不平等―成熟する日本型学歴社会　東京大学出版会

吉川徹　2009　学歴分断社会　筑摩書房

小西貞則　2010　多変量解析入門―線形から非線形へ　岩波書店

国立教育政策研究所　2005　特定の課題に関する調査（国語，算数・数学）

耳塚寛明　2007 学力と家庭的背景－保護者調査を用いた小6算数学力の分析『JELS　第10集』1-15.　お茶の水大学

OECD教育研究革新センター編（教育テスト研究センター監訳）2008　学習の社会的成果―健康，市民・社会的関与と社会関係資本　明石書店

志水宏吉　2005　学力を育てる　岩波書店

志水宏吉　2007　教育資本について　教育文化学年報第2号　3-20. 大阪大学大学院人間科学研究科教育文化学研究室

内閣府統計局編　2010　統計でみる都道府県のすがた

橘木俊詔　2010　日本の教育格差　岩波書店

山田昌弘　2004　希望格差社会―負け組の絶望感が日本を引き裂く　筑摩書房