

## 混合モデルによる熟達パターンの探索 - プロ野球選手の熟達の軌跡を例として -

著者	清水 和秋
雑誌名	関西大学社会学部紀要
巻	40
号	1
ページ	17-37
発行年	2008-12-20
その他のタイトル	Exploring the Patterns of Skill Development by Mixture Growth Modeling: Using the Batting Average Data on Professional Baseball Players
URL	<a href="http://hdl.handle.net/10112/2725">http://hdl.handle.net/10112/2725</a>

## 混合成長モデルによる熟達パターンの探索 — プロ野球選手の熟達の軌跡を例として —

清水 和 秋

### Exploring the Patterns of Skill Development by Mixture Growth Modeling: Using the Batting Average Data on Professional Baseball Players

Kazuaki SHIMIZU

#### Abstract

Reviewing latent growth modeling for longitudinal data and some results using this methodology on career development, mixture modeling methodologies were introduced for identifying clusters of individuals following similar developmental trajectories. For the latent growth model analysis by Amos and the group-based trajectory model analysis using SAS Traj procedure, the batting average records of Japanese professional baseball players over ten years were selected from the published official records. Results of latent growth modeling demonstrated that the quadratic form trajectory model fit the data well. Six sub-groups were also clustered by the same quadratic form using the Traj. Findings of these analyses were discussed with particular reference to the utility of the group-based trajectory modeling of mixture model methodology for analyzing career development processes.

Keywords: latent growth model, mixture modeling, group-based trajectory model, baseball, career development, trajectory

#### 抄 録

縦断的データへの潜在成長モデルとこの方法論を使ったキャリア発達についての結果を概観しながら、混合モデリングの方法論を、類似した発達軌跡に従う個人のクラスタを特定するために、紹介した。Amosによる潜在成長モデル分析とSAS Trajプロシジャを使った集団ベースの軌跡モデル分析のために、10年間を越える記録を持つ日本のプロ野球選手の打撃成績記録を公開されている公式記録から取り出した。潜在成長モデルの結果は、2次形式軌跡モデルがデータにうまく適合することを示した。6集団が、また、TRAJを使って、同じ2次形式によってクラスタ化された。これらの分析からの見いだしたことを、混合モデル方法論の集団ベースの軌跡モデル化の有用性をキャリア発達過程の解析と関連づけて議論した。

キーワード：潜在成長モデル、混合モデル、集団ベース軌跡モデル、野球、キャリア発達、軌跡

## 1. はじめに

個人間差には2つの種類がある。1つはよく知られているように、伝統的な心理測定の範疇に入ると考えられてきた個人間差である。因子分析法や項目応答理論は、個人間差の原因の源を潜在的な因子に求める方法論として展開してきた。共分散構造と平均構造を包含する構造方程式モデル（SEM: Structural Equation Modeling）は、古典的テスト理論と因子分析法をベースとして理論的にもそして応用的にも発展してきた。

時間経過の中に表れる変化パターンは、個人によって異なる。これがもう1つの個人間差である。個人内変化は、時間軸と共変する（Nesselrode, 1991）。このような発達過程での変化の形を、集合的統計量としての平均の姿においてSEMの応用モデル群では、取り扱うことに成功している。個々人が示す変化の形あるいは変化パターンの個別性は、日常的な観察からも推測できるところではある。

この2つめの個人間差は、縦断的に調査したデータでしか追求することができない。調査対象者全体の中に潜在する複数の変化パターンを見つけ出す方法について、解析例を示しながら、検討してみることにする。

### 1-1. LGM

時間経過の中で起きる個人内変化をモデル化するための方法論は、Bollen & Curran (2006) が詳しく紹介しているように、潜在成長モデル（LGM: Latent Growth Model, McArdle, 1986, 1988）あるいは潜在曲線モデル（LCM: Latent Curve Model, Meredith & Tisak, 1990）と呼ばれるモデル群として、確立されてきた。この方法論の欧米での展開については、キャリア発達の解析結果を報告する中でこれまでも言及してきた。例えば、清水（1999a）では、キャリア発達についての縦断データを対象としてコーホート分析を複数集団の同時モデル化でおこない、測定機会に欠損があっても一貫したキャリア成熟の姿をLGMで捉えることができることをAmosのスク립トとともに報告している。1つの変数の発達の軌跡をLGMでモデル化しながら、線型モデルでの切片や傾きに影響を与える外的要因を組み込んだモデルや多変数LGMを清水（1999b）で紹介している。これらは、いずれも最大で4回までの繰り返し測定を対象とした線形のLGMに焦点を当てたものであった。一般的にキャリア発達の過程は、成長・探索・確立・維持・離脱の5つの段階としてモデル化され、この変化の軌跡は、上昇からピークへそして下降へという2次曲線としてイメージされてきた。生涯にわたる軌跡を描くためのデータを収集するには長期間に

わたる大規模な研究計画を現実のものとしなければならない。キャリア発達の1つの例として、清水（2003）では、プロ野球選手の10年間の打撃成績に求め、潜在成長モデルにより、その軌跡を一般的なキャリア発達のように2次曲線として描き出した。

LGMは、McArdle（1986, 1988）とMeredith & Tisak（1990）によって、時間経過（あるいは時間関数）を固定パラメータとし、このパラメータによって変化に関する因子をモデル化（切片や増分や非線形項など）し、これらを平均構造において推定するSEMの応用モデル群として展開されたものである。切片の推定を組み込んだSEM解析のソフトでは、縦断データの解析手法として、LGMを解析に使用することができる（狩野・三浦, 2002）。最近では、岡林（2006）と服部（2006）によるLGMの紹介がある。

対象とした集団の発達（あるいは変化）の軌跡は、その集団にもっとも適合する関数として表現されることになる。ある意味では、LGMは対象母集団内の平均的な変化の姿を描き出すことになる。集団全体の変化の様相やこれに影響するあるいは影響を受ける要因を特定することを目的とする場合にはLGMは有効な方法である。集団の中に異質な下位集団がいくつか存在する場合には、SEMでいうところの適合度はそれほどよいものとならないかもしれない。異質な集団があらかじめ特定されているなら、これらを多集団同時分析の対象とすることもできる。通常の解析場面では、対象の全体を1つの標本として括っているわけであるから、内部の下位集団を特定する先見的情報の存在を前提とすることはできない。

個々の発達の様相が多様であることは容易に想像できることである。個のデータから、介入や教育プログラムの効果について法則性のある結論を引き出すことは易しいことではない。全体を1つの集団として解析した結果から、被験者を何らかの基準で分類することも考えられる。再現性を確保した方法でなければ、恣意的という批判を浴びることになる。

## 1-2. 成長混合モデル

混合モデル（Mixture Modeling: 混合分布モデルとも呼ばれる）は、未知のメンバーシップを推定することにより、傾向性が似通った下位集団を特定することを目的とした方法論として、近年注目を浴びている（服部, 2006）。元々のアイディアはP.F.ラザースフェルドによる潜在クラス分析（Clogg, 1995）にある。

ある個人 $i$ の測定機会 $t$ での $y$ についての得点を $y_{it}$ とする。この得点について $G$ 個の集団にからなる有限混合分布に従うとして、次のように表す（Jones & Nagin, 2007）。

$$y_{it} = \sum_{g=1}^G p_i^{(g)} \hat{y}_{it}^{(g)} \quad (1)$$

ここで、 $p_i^{(g)}$  は、個人*i*が集団*g*に属す確率であり、 $\sum_{g=1}^G p_i^{(g)} = 1$ とする。 $\hat{y}_{it}^{(g)}$ は研究対象の関数であり、混合LGMの場合には、ここへSEMベースのLGMの式が当てはめられることになる（Muthén, 2002）。実際の推定値の計算では、正規分布を仮定したEMアルゴリズムによっておこなわれる。Bollen & Curran（2006）は、この種の推定では、初期値が重要な役割を果たし、場合によると局小値に陥ることがあること、多変量正規分布の仮定が通常のSEMでの最尤推定よりも重要となることを指摘している。

SEMの代表的なソフトであるMplus（Muthén & Muthén, 1998-2007）には、この混合モデルがオプションではあるが組み込まれている。変化の軌跡の不変性を確保した上で、対象者をいくつかのグループに分けることができる。なお、Amosの最新版（Ver.16）でもベイズ推定を組み込んだ混合モデルが提供されている。

混合モデルは、LGMだけではなく、横断的なデータを対象とするSEMでも、有力な方法といえるが、先にも紹介したように分布に関する条件を満たすことが難しいことが予想される。1つの母集団から互いに異質な*G*個の集団を特定しようとする際に、その各集団に属する人数が分布の条件を満たすほどに十分に大きなものとなるのかどうか、という点である。1標本を対象とするLGMでも、縦断調査でのデータ管理が適切に行われていない場合には、切片や傾きそして2次の項に相当する因子の分散や共分散の推定において、不適解に遭遇することもある。洗練された調査計画の下で大規模に実施されて縦断データの収集に積極的な欧米では、Mplusによる解析結果の報告が多い。残念ながら、200～400程度のデータに潜む*G*個の集団をクラスタ化するには、分布の条件は厳しすぎる感がある。

Nagin（1999）は、時間経過の変化パターンのクラスタ化に混合モデルを応用している。(1)式の $\hat{y}_{it}^{(g)}$ に、心理的な変数の場合の軌跡（trajectory）として、次の4次関数による成長混合モデルとする集団ベースの方法を提案している。

$$\hat{y}_{it}^{(g)} = \beta_0^{(g)} + \beta_1^{(g)}T_{it} + \beta_2^{(g)}T_{it}^2 + \beta_3^{(g)}T_{it}^3 + \beta_4^{(g)}T_{it}^4 + e_{it} \quad (2)$$

ここで、 $\hat{y}_{it}^{(g)}$ は個人*i*の測定機会*t*についての推定値であり、最尤法で推定される重み係数 $\beta_0^{(g)}$ 、 $\beta_1^{(g)}$ 、 $\beta_2^{(g)}$ 、 $\beta_3^{(g)}$ そして $\beta_4^{(g)}$ と測定機会（あるいは年齢） $T_{it}$ の関数として計算される。*g*は、調査対象者が所属するクラスタ（メンバシップ）を表している。時間経過の中の変化のパターンは、(2)式として表現され、これを(1)式に代入し、最尤推定がおこなわれ、 $p_i^{(g)}$ の確率によりクラスタ化がおこなわれると解することができる。この成長

混合分析については、Jones, Nagin, & Roeder (2001) が<sup>3</sup>、SASのプロシジャTrajを提供している。Jones & Nagin (2007) は、Trajの推定値とメンバシップ確率  $p_i^{(g)}$  についての信頼区間の統計量と複数の変数の軌跡をモデル化する方向での拡張をおこなっている。なお、Trajの使い方については、Arrandale, Koehoom, MacNab, & Kennedy (2006) にも詳しい説明がある。

## 2. 方法

### 解析データ

清水 (2003) で使用したデータをそのまま本稿でも解析の対象とする。このデータの出所は日本野球機構 (1998) である。掲載されている膨大なデータの中から、打率のデータをOB選手と現役選手について抽出したものをここでは使用する (上野, 2000)。野球の成績は、出場した試合におけるものであり、出場試合の数が成績に影響している。キャリア発達の一般的な傾向性を、ある程度の年数以上にわたって活躍してきた選手から得るために、実働年数が10年以上で、各年の平均出場試合数が90試合以上の156名の選手を対象とした。なお、戦前からの選手やデータの一部が欠損している選手は、対象から除外している。現役選手のデータでは、同様に実働が7年以上を基準として、セリーグ・パリーグの139名の選手を対象とした。表1は、これらの選手 (N=295) の打率についての15年目までの平均である。

表1 プロ野球OB選手と現役選手の打率の平均 (日本野球機構 (1998) より)

年	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
OB選手	平均	0.238	0.251	0.259	0.261	0.266	0.267	0.269	0.268	0.266	0.262	0.262	0.260	0.268	0.253	0.250
	人数	115	136	149	151	155	156	155	156	153	148	141	128	115	105	85
現役選手	平均	0.236	0.237	0.262	0.253	0.254	0.260	0.258	0.255	0.255	0.250	0.243	0.263	0.249	0.255	
	人数	66	80	92	113	118	120	115	96	89	76	54	39	29	24	15

注：「現役」は1998年の日本野球機構資料に従っている。清水 (2003) の表 (p.68) を引用。

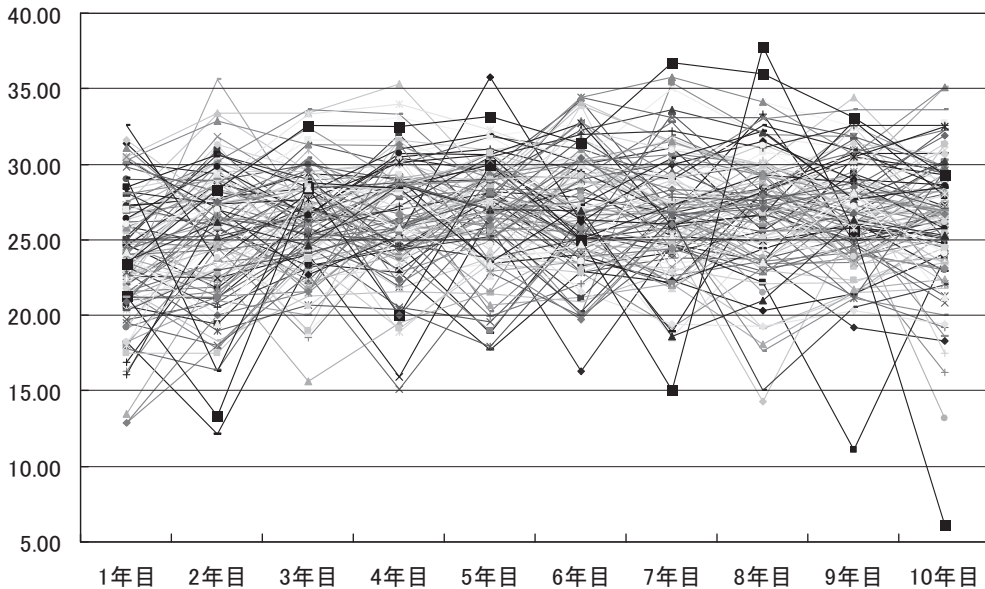


図1 分析対象のプロ野球選手121名の打撃成績の10年間のグラフ  
 注：図では、打率を100倍して表示している。

打率は7年目をピークとしているようである。そこで、打率の上昇・ピーク・下降という熟達の過程をある程度までカバーすることのできる範囲として、10年を対象期間とした。この10年間において、すべての年度において欠損のないデータは、OB選手100名、そして現役選手21名であった。清水（2003）では、完全10年間縦断データとして、このデータ（N = 121）を分析対象としている。図1は分析対象選手の個人別データを折れ線グラフで表示したものである。

### 3. 分析・結果

#### 3-1. LGMの概要（清水（2003）より）

##### 3-1-1. 潜在成長モデル

観測変数の素点のベクトルを  $y$  とすると、LGMは、(3)式のように素点からの共通因子分析モデルとして定義される。

$$y = Af + u \tag{3}$$

ここで、 $f$  は因子得点ベクトル、 $u$  は独自性の得点ベクトルである。LGMのオリジナリティは、因子パターン行列  $A$  に、観測機会の時間関数に関する固定パラメータを設定し



たところにある。(3)式の観測変数  $y$  の次数は観測の繰り返し回数である  $T$  とする。この次数は  $u$  も同様であるが、 $f$  次数は、定義する関数によって変わるので、ここでは確定させていない。このため  $A$  の列の次数も同様となる。切片のみの推定の場合には、この因子パターン行列の列の次数は 1 となり、線形式では切片と傾斜の 2 つとなる。

ここではもう少し説明を追加するために、(3)式を 2 次の発達曲線で、観測機会を 4 回として、個人  $i$  の素点から表してみることにする。

$$\begin{aligned}
 y_{1i} &= 1f_{1i} + 0f_{2i} + 0^2 f_{3i} + u_{1i} \\
 y_{2i} &= 1f_{1i} + 1f_{2i} + 1^2 f_{3i} + u_{2i} \\
 y_{3i} &= 1f_{1i} + 2f_{2i} + 2^2 f_{3i} + u_{3i} \\
 y_{4i} &= 1f_{1i} + 3f_{2i} + 3^2 f_{3i} + u_{4i}
 \end{aligned} \tag{4}$$

SEMソフトでの推定の対象は、(4)式でいえば、軌跡の切片に相当する  $f_1$ 、1 次の傾斜に相当する  $f_2$ 、そして、2 次の項に相当する  $f_3$  の平均と分散とこれらの間の共分散である。共通因子分析モデルとして、 $u_1 \sim u_4$  の独自性の分散も推定することになる。通常の因子分析を適用した場合に推定の対象となる因子パターン行列  $A$  に、LGMでは、軌跡を適切に表す方程式を固定パラメータとして組み込んでいるわけである。

AmosによるLGMの推定については、清水(1999a)にスクリプトも掲載している。狩野・三浦(2002)にもLGMの解説がある。服部(2006)はRでのプログラム例も掲載している。この方法論の最も包括的なテキストはBollen & Curran(2006)である。パラメータの推定方法や適合度の評価は、一般的なSEMと同じである。ここでは、詳細は省略する。

### 3-1-2. 全体データ (N=121) の解析結果 (清水(2003)より)

SEMの解析では、完全10年間の縦断データ (N=121) に、まず、(3)式のパラメータ行列を  $T=10$  にまで拡大し、切片と傾斜の 2 つの因子についての線形モデルについて、最尤法によって推定をおこなった。モデルの適合度は、カイ 2 乗統計量は ( $\chi^2=118.957$ ,  $df=50$ ,  $P=0.000$ ) であり、RMSEA=0.107とかなり悪いものであった。すなわち、打率の変化は、線形のモデルには、当てはまらなかったわけである。

次に、(3)式のパラメータ行列を  $T=10$  にまで拡大し、切片因子と傾斜因子に、2 乗の項の因子を設定した非線形モデルで、推定をおこなってみた。結果は、カイ 2 乗統計量は ( $\chi^2=50.426$ ,  $df=46$ ,  $P=0.303$ ) であり、RMSEA=0.028となり、このモデルの適合度は非常によいと判断することができた。なお、線形のモデルとこの 2 次関数モデルの AIC を比較すると、それぞれ 148.957 と 88.426 であり、2 次の非線形モデルのほうが適合度がよい



と、この指標からもいうことができる。

表2 2次関数モデルの推定値 (清水 (2003, p.69) より)

	因子の平均			因子の分散			因子の共分散			
	切片	傾斜	2乗の項	切片	傾斜	2乗の項	切片と傾斜	切片と2乗	傾斜と2乗	
推定値	24.097	1.104	-0.098	9.253	0.997	0.008	-2.157	0.175	-0.083	
標準誤差	0.342	0.138	0.014	1.981	0.331	0.003	0.725	0.068	0.032	
	独自性分散									
	1年目	2年目	3年目	4年目	5年目	6年目	7年目	8年目	9年目	10年目
推定値	7.055	10.390	6.975	8.608	7.237	7.333	8.639	8.658	5.671	10.024
標準誤差	1.732	1.569	1.041	1.233	1.076	1.100	1.267	1.278	0.974	1.752

表2が2次関数をプロ野球選手の打撃成績の軌跡のモデルとして推定された値である。この解析に入る前に、結果を見やすくするために、個々の選手の打撃成績を100倍している。表3が、標本の平均値と推定した結果から得られたモデル式の10年間の値である。図1では、これを折れ線グラフで表した。

表3 標本 (N=121) の平均値と2次関数モデルの推定から得られる平均 (清水 (2003, p.69) を一部改編)

	1年目	2年目	3年目	4年目	5年目	6年目	7年目	8年目	9年目	10年目
観測値	23.788	25.303	26.269	26.626	26.940	27.085	26.793	26.912	26.727	26.230
モデル値	24.097	25.103	25.912	26.525	26.942	27.162	27.186	27.014	26.645	26.079

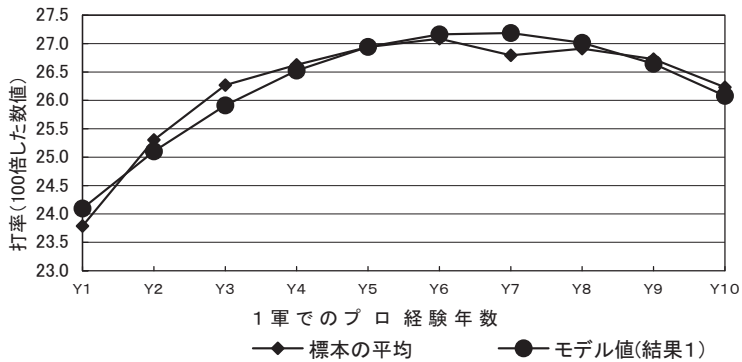


図2 観測値とLGM推定値のグラフ (清水 (2003, p.70) より)

標本の値では、7年目に打撃成績の平均が低下し、8年目に少々ではあるが回復している。ピークは観測値からでは6年目ということになったが、2次関数のモデルでは7年目である。観測値とモデル値とでは、このような食い違いがおきた。

清水（2003）では、このズレの原因をキャリア開始時の年齢から説明することを試みている。ここで取り扱ったデータでは、プロ野球選手としての1年目をキャリア初年として、キャリアを開始する前の経歴である高卒・大卒・社会人別での取り扱いをしていない。そこで、キャリア開始時の年齢をモデルに組み込むことを試みている。その結果、モデルの適合度は、カイ2乗統計量では（ $\chi^2=60.732$ ,  $df=53$ ,  $P=0.217$ ）となり、RMSEA = 0.035となった。そして、年齢変数からのパス係数は、切片が12.41（3.61）、傾斜が4.91（1.48）となり、2乗の項が-0.39（0.15）となった。なお、括弧内は標準誤差である。この結果から、打撃の熟達の形の切片は、年齢からの影響を最も大きく受け、曲線のカーブの上がり方も年齢が高いほど大きい、と解釈しているが、観測値の7年目の平均値の減少を説明することには成功していない。

全体として見えるものの中に存在する何かを追求する可能性が、多集団の同時分析にあることを示唆して、清水（2003）では報告を終えた。本稿は、この問題の追及を目的としているわけではない。N = 121に潜在するいくつかの下位集団を特定することによって、新しい切り口からの視点を提供してみたい。

### 3-2. 混合モデル：Trajによる推定

パソコン版SASにProc Traj（注にURLを記載）を導入し、上で紹介してきたプロ野球選手の打撃成績データ（N = 121）において、軌跡のパターンを探索してみることにする。ここで探索としたのは、(1)式の集団の数 $G$ の決定と変化に関する関数式（(2)式）の適切な項の決定に関して確定した方法がないため、結果を比較しながら適切どころを探さなければならぬからである。

最尤推定から提供される統計量としては、モデル全体の適合度に関するものと関数の重み係数に関するものがある。前者では、よく知られている赤池情報基準量（AIC）とベイズ情報基準量（BIC）が他の混合モデル解析と同じように提供される。Trajでは、このBICの値はマイナスで出力され、負の数がより小さなモデルがより適合度のよいモデルと判断することができる。Jones et al.（2001）は、さらに、SEMのような相対的な比較指標を導入している。単純なヌル・モデルを設定し、複雑なモデルとの差（(5)式）をとって、解釈に値するモデルを選び出すという提案である。

$$\Delta BIC = BIC_{(complex)} - BIC_{(null)} \quad (5)$$

ヌル・モデルとしては単純なモデル、すなわち $G$ の数が小さく、方程式の項の数が低いも

を推奨している。*BIC*の差の値が2よりも小さい場合には、検討の対象とする必要はなく、10よりも大きい場合には複雑なモデルがヌル・モデルと異なる十分な根拠があるとしている。この他にも、ヌル・モデルを相対的に位置させる方法など、いくつかの指標が提案されている。

(2)式の間数式の切片や重み係数 ( $\beta_0^{(s)}$  から  $\beta_4^{(s)}$ ) については、推定値の標準誤差の計算から、係数をゼロと置く帰無仮説の棄却についての出力がおこなわれる。この結果を利用して、有意な重み係数の項だけから (2)式を再構成することになる。

### 3-3. 下位集団の探索と結果

#### 3-3-1. 集団の数を1としたモデル

探索の前にデータの全体を1つの集団としてTrajにより4次関数から1次関数までのパラメータと適合度の指標を計算してみた(表4)。

表4 集団の数を1としたTrajでの推定値

モデル	切片	1次	2次	3次	4次	BIC
4次関数	21.287 ***	3.068 *	-0.635	0.061	-0.002	-3337.90
3次関数	22.072 ***	2.061 ***	-0.267 **	0.010	—	-3335.87
2次関数	22.959 ***	1.274 ***	-0.096 ***	—	—	-3334.87
1次関数	25.074 ***	0.217 ***	—	—	—	-3352.81

注：推定値の有意性については、\*\*\*は0.1%水準、\*\*は1%水準、\*は5%水準を表す。  
推定に該当しない欄は—と表示した。

これらの4つの*BIC*を比較すると1次関数の線形モデルの適合度が最も悪く、2次関数のモデルの適合度がよいという結果になった。これらの推定値は、表2のLGMの2次関数のパラメータの推定値とは微妙に異なる値となった。各モデルの特徴を確認するために、各モデルの10年間のモデル値を算出してみた(表5と図3)。

表5 集団の数を1とした推定値から計算したモデル値

モデル	1年目	2年目	3年目	4年目	5年目	6年目	7年目	8年目	9年目	10年目
4次関数	23.779	25.339	26.261	26.791	27.127	27.419	27.769	28.231	28.811	<b>29.467</b>
3次関数	23.876	25.206	26.122	26.684	26.952	<b>26.986</b>	26.846	26.592	26.284	25.982
2次関数	24.137	25.123	25.917	26.519	26.929	27.147	<b>27.173</b>	27.007	26.649	26.099
1次関数	25.291	25.508	25.725	25.942	26.159	26.376	26.593	26.81	27.027	<b>27.244</b>

注：各ピークは太字で表示した

2次関数の軌跡の様相は、表3のLGMと微妙に違ったものとなった。この違いは、(2)式と(3)あるいは(4)式が仮定するモデルの違いによるものと推測される。前者は、観測変数への重み係数を4次関数において求めるものであり、一般的な非線形回帰モデルに相当する。後者では、観測変数の独自性を推定しており、軌跡のモデルは、共通因子空間において推定される。因子分析の流れからみれば、Trajが採用しているモデルよりは、後者のLGMのほうが個人差を捉える方法としては、より適切であることは明らかである。1つの集団としての軌跡は、表3の「LGMのモデル値」のほうが誤差が少ないと考えられるが、Trajとの間の理論的な比較は別な機会とする。

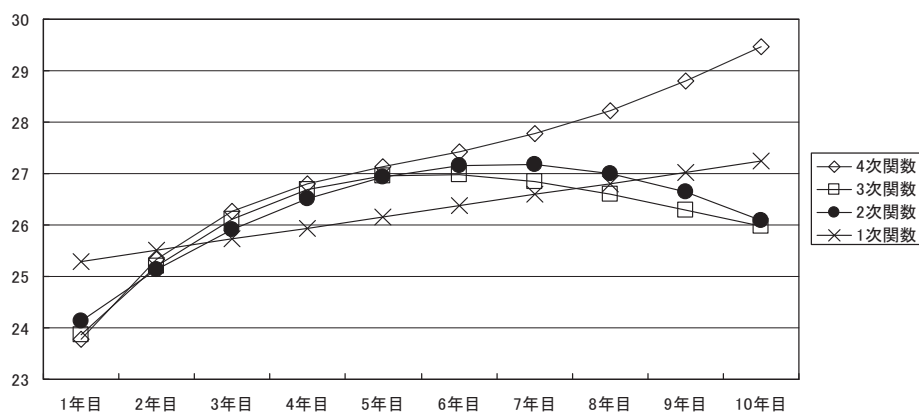


図3 4つのモデルの軌跡（表5のグラフ）

2次関数のピーク（太字で表示）はLGMと同じように7年目となった。これに近い軌跡を示したのが、3次関数であった。この場合には、ピークは1年下へ移動して、6年目となった。この3次の推定値は表4にあるように有意ではなく0.010というように非常に低い値であった。このモデルの軌跡は、3次の項がゼロに近くなるので、形は2次関数に近く、ピークが前に移動し、後半では下がる割合が大きくなっている。

線形の1次関数は、切片の25.074から1年ごとに傾き分(0.217)の単調増加を示し、10年目に27.244の値でピークとなった。4次関数は複雑な傾向を示している。表4の2次、3次、4次ともに推定値は有意ではなく、BICでも適合度は良くない。ピークは、線形と同様に10年目という異常な結果となった。熟達の現象を説明するモデルとしては、この1次と4次の結果は、不適切なものといわざるを得ない。キャリア生活からの離脱の頃に、ピークを迎えるということだけではなく、キャリアを継続すればするほどに打率が上昇す

るというモデルとなってしまうからである。

Trajのヌル・モデルとして推奨されるのは集団の数が少なく、より単純なモデルであった。表4の値からみると4次関数と3次関数はゼロと判断すべき値が推定されていることになる。2次と1次では、いずれの推定値とも有意であるので、解釈可能かどうかは別にして、変化パターンとしての統計的な意味はあることになる。無に帰すことのできる仮説的な軌跡としてみると1次関数となる。そこで、ここでは、これをヌル・モデルとすることにする。

表6 下位集団の数を2～8とした2次関数モデルによる推定値

集団数		人数	切片	1次	2次	BIC
2	下位集団1	52	21.348 ***	1.226 ***	-0.101 ***	-3240.60
	下位集団2	69	24.275 ***	1.314 ***	-0.092 ***	
3	下位集団1	17	18.344 ***	1.742 ***	-0.142 ***	-3225.94
	下位集団2	67	22.861 ***	1.15 ***	-0.088 ***	
	下位集団3	37	25.239 ***	1.279 ***	-0.089 ***	
4	下位集団1	12	15.774 ***	2.41 ***	-0.179 ***	-3220.17
	下位集団2	44	23.330 ***	0.79 **	-0.071 **	
	下位集団3	61	23.814 ***	1.389 ***	-0.097 ***	
	下位集団4	4	27.343 ***	1.525 *	-0.109	
5	下位集団1	9	15.716 ***	2.389 ***	-0.180 ***	-3214.19
	下位集団2	48	21.323 ***	1.483 ***	-0.104 ***	
	下位集団3	12	26.008 ***	0.137	-0.067	
	下位集団4	48	25.120 ***	1.058 ***	-0.074 **	
	下位集団5	4	27.404 ***	1.595 *	-0.116	
6	下位集団1	10	15.635 ***	2.416 ***	-0.182 ***	-3213.41
	下位集団2	38	22.778 ***	0.896 **	-0.065	
	下位集団3	9	26.290 ***	0.114	-0.072	
	下位集団4	14	16.655 ***	3.477 ***	-0.243 ***	
	下位集団5	46	25.706 ***	0.855 ***	-0.060 **	
	下位集団6	4	27.371 ***	1.597 *	-0.116	
7	下位集団1	10	15.558 ***	2.446 ***	-0.186 ***	-3219.23
	下位集団2	16	16.896 ***	3.339 ***	-0.234 ***	
	下位集団3	9	26.320 ***	0.127	-0.074	
	下位集団4	16	23.883 ***	1.786 **	-0.126 **	
	下位集団5	33	26.545 ***	0.332	-0.022	
	下位集団6	33	22.567 ***	0.897 *	-0.065 *	
	下位集団7	4	27.572 ***	1.572 *	-0.115	
8	下位集団1	8	16.465 ***	1.621 **	-0.085	-3221.05
	下位集団2	15	17.053 ***	3.318 ***	-0.232 ***	
	下位集団3	6	15.864 ***	3.477 ***	-0.332 ***	
	下位集団4	17	23.975 ***	1.763 **	-0.125 **	
	下位集団5	34	22.713 ***	0.9 *	-0.067 *	
	下位集団6	30	26.666 ***	0.289	-0.019	
	下位集団7	7	28.208 ***	-0.523	-0.024	
	下位集団8	4	27.564 ***	1.579 *	-0.115	

注：有意水準は、0.1%水準に\*\*\*、1%水準に\*\*、5%水準に\*をつけている。

BICはBayesian Information Criterionである。

### 3-3-2. 下位集団の数の探索

下位集団の探索では、集団の数を2から8まで変化させながら、線形の1次関数と2次関数の2つのモデルについて検討をおこなった。全体での結果と同じように、軌跡のモデルとしては、集団数を変化させても、2次関数のほうが、*BIC*の値からも適合度が良いと言えた。表6は、2次関数の結果である。Trajによる軌跡の推定では、一貫して、各下位集団の切片の推定値が、いずれの方法でも有意となった。すなわち、変化パターンのはじまりを異にする下位集団を特定することができたわけである。

ここでは計算結果を提示していないが、表6の推定値からそれぞれの下位集団について10年間の軌跡を描くための計算をおこなった。集団の数ごとに、それぞれの全体の平均を2次関数の加重平均から計算してみた。情報量が膨大となるので、ここでは、これらの数値や図は、省略することにして、まず、結果の概要を検討してみることにする。

集団数を1として得られた2次関数の軌跡が、集団の数を増やすごとに、曲線の形はそのままに打撃成績のレベルに応じて平行するように軌跡が分割していった。集団の数を2とした場合には、下位集団1（キャリアのはじまりが2割5分5厘、ピークが2割9分、10年目が2割8分2厘）と下位集団2（同じく、2割2分5厘～2割5分1厘～2割3分5厘）の2つの軌跡となった。ピークの位置が、下のレベルの軌跡では6年目で、下位集団1は7年目であった。集団数を3とした場合にもレベルの低い下位集団のピークは6年目で、中間と上位のピークは7年目となった。LGMの解析のところで紹介した素点での平均が7年目で減少する現象を報告したが、2つの集団のピークの違いにこの原因を求めることができそうである。

表7 下位集団数を6とした推定から計算した各集団の軌跡

	1年目	2年目	3年目	4年目	5年目	6年目	7年目	8年目	9年目	10年目
下位集団1	17.869	19.739	21.245	22.387	23.165	23.579	<b>23.629</b>	23.315	22.637	21.595
下位集団2	23.609	24.310	24.881	25.322	25.633	25.814	<b>25.865</b>	25.786	25.577	25.238
下位集団3	<b>26.332</b>	26.230	25.984	25.594	25.060	24.382	23.560	22.594	21.484	20.230
下位集団4	19.889	22.637	24.899	26.675	27.965	28.769	<b>29.087</b>	28.919	28.265	27.125
下位集団5	26.501	27.176	27.731	28.166	28.481	28.676	<b>28.751</b>	28.706	28.541	28.256
下位集団6	28.852	30.101	31.118	31.903	32.456	32.777	<b>32.866</b>	32.723	32.348	31.741

注：太線で示した数値はそれぞれの集団のピーク値である。

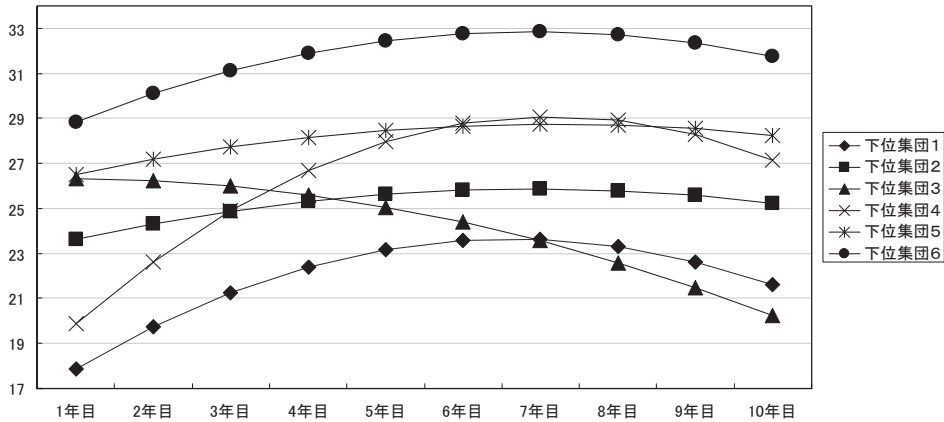


図4 下位集団数を6とした推定から計算した各集団の軌跡

下位集団の数をBICからみてみるとGを6とすることを結論とすることができる。 $\Delta BIC$ について、ヌル・モデルを全体の1次関数の値として、計算しても、結論は同じである。そこで、2次関数で、10年間の軌跡について、表6から軌跡の推定値を計算してみた(表7)。これをグラフ表示したのが、図4である。

表6の集団の数を増やすという試みにおいて、集団数を4としてから以降では、高打率で10年間のキャリアを通した4人からなる集団が浮かびあがってきた。この4名とは下位集団6の「長嶋茂雄、若松 勉、張本 勲、落合博満」である。この集団の打撃成績(表7)は、キャリアのはじまりで2割8分9厘あり、7年目で3割2分9厘のピークを迎え、10年目でも3割1分7厘である。

長嶋の所属する集団とよく似た2次関数の形を示した下位集団が他に3つある。打撃成績のレベルで次に位置するのが、下位集団5である。これに属するのは「高田 繁、掛布雅之、藤田 平、広沢 克など46名」であり、2割6分～2割9分弱～2割8分という変化を示している。次のレベルには、2割4分弱～2割6分弱～2割5分と大きな変動のなかった「富田 勝、山倉和博、大矢明彦、田淵幸一、伊東 勤など38名」からなる下位集団2が位置している。1割8分弱からキャリアがはじまりまった下位集団1は、表6にもあるように1次の項(傾き)が他の集団よりも大きく、キャリア初期において急速に打撃成績を向上させている。この集団のピークは2割3分6厘であり、2次関数の軌跡を示した打撃レベルで上の集団のキャリア始めの値を超えることはなかった。この集団は、10年目を2割1分6厘で終えている。

これらの4つの下位集団では、変化のパターンがほぼ平行して推移しており、平均の軌



跡をレベルに応じて分割したようでもある。この4本の軌跡は互いに接することはなく、打撃の熟達が、レベルに違いはあっても、LGMで全体から描き出した2次関数に近い形で、独立して進行している。レベルという質での違いが、4つの下位集団を分けているともいえる。

下位集団4は低いレベルから急速に打撃成績が向上している。この集団には「大杉勝男、王 貞治、大島康徳、高木守道など14名」が属する。6下位集団の中で2番目に低いポイント（約2割）からキャリアがはじまり、ピークは2割9分と上から2番目に位置し、10年目には2割7分となる。表6のこの集団の推定値は、1次も2次もともに他よりも大きく、図4で示されているように、軌跡の変動は大きい。これまでにみえてきた下位集団と同じ様な熟達の傾向にあるといえそうである。

ユニークな軌跡を示したのは、下位集団3である。約2割6分からキャリアがはじまるわけであるが、このはじまりがピークで、10年目には約2割へとほぼ直線的に打撃成績が低下する。この集団の1次と2次の重みは、他の集団とは違って、有意な推定値ではない。重みをゼロとする帰無仮説を棄却できなかったわけであるから、有意な切片がそのまま推移すると解釈すべきと考えることができる。言い換えれば、高いレベルからキャリアを開始し、これを維持するように10年間を過ごした、ということもできる。この集団には「広岡達朗、本屋敷錦吾、大沢啓二など9名」が属している。

下位集団の数を5としたBICと6のそれとの差はほとんどなかった。メンバーの入れ替わりもほとんどなく、違いは、急激に成績が上昇を示した下位集団4が出現したことである。このように、Trajでは、指定した集団数に対応させながら個々の変化のパターンをクラスタに分けることができた。ここまでに示してきたように、適合度の全体をBICで、軌跡の構成要素である関数の重み係数についての検定統計量を検討の対象としながら、図1のような全体の中に潜在する下位集団を、図4のように特定することができる。

10年間の中から1つの年度に注目すれば、この年度での個人間の差異が見えてくる。このような個人間差が、10年間にわたってほぼ維持されるのは、集団数を6とした結果での下位集団6、5、2、1の4つの集団に属する選手たちである。変化パターンとしては、これらの選手は、それぞれが所属する下位集団の特徴を共有している。すなわち、2つの個人間差が、10年間にわたって一貫しているといえることができる。

ここに下位集団4を加えてみると、この集団に所属する選手は、キャリア初期での低いレベルの打撃成績からピーク時では、2番目に高いレベルへと変動する。伝統的な個人間差という観点からみると変動するわけではあるが、10年間の期間においては、図4にある

ように2次傾向の強い変化ではあるが、一貫した姿を示しており、この形は、他の集団とは違ったものであることもまた一貫している。集団3も同様に、ユニークな変動を一貫して示していると解することができる。ここでは集団として表現したが、ここで検討してきたように、縦断的に捉えた一貫した変動のパターンもまた個人を特徴付けるものであるとすることができる。

## 4. 考察

### 4.1 変化パターンの解析と混合モデル

LGMがキャリア発達研究に有効な道具であることを、清水（1999a,b, 2003）などにおいて強調してきた。LGMは、観測変数の独自性を推定することによって、変化のパターンあるいは軌跡を共通因子として描き出すことに成功しているからである。そして、モデルの記述の自由度が高いことなど解析結果を示しながら、その魅力を紹介してきた。SEM系のソフトであるMplus（Muthén & Muthén, 1998-2007）には、ここで検討してきた混合モデルがLGMの体系の中に組み込まれている。十分に大きな標本を対象にして緻密に管理された縦断調査を実施し、変数の分布が多変量正規分布であるなら、そして、潜在する異質な下位集団においても、この条件が満たされるなら、その能力を遺憾なく発揮するはずである。Muthén（2002）は、カテゴリー変数への展開も合わせて整理しているように、SEMは、ここにおいて新しい時代を迎えたといえよう（Bollen & Curran, 2006）。

混合モデルは、(1)式を基本とするものであって、ここで紹介したTrajでは、(2)式を代入することになる。LGM系では(3)式を取り扱うことになる。この違いは実は大きく、Trajの多項式関数では、各項を独立していると仮定してパラメータの推定をおこなう。Trajには、(2)の多項式の他に2つの関数が提供されているが、ここで示してきたように項の数を増減させることができるだけで、利用者が対象とする問題に適切な関数を取り扱うことはできないこともある。これに対して、LGMは、(3)式あるいは(4)式の軌跡を構成する因子間に共分散を設定し、これを集団間で拘束の対象とすることができるなど、仮説的モデルの記述の自由度は非常に高い。

Thissen & Bock（1990）が紹介しているように発達現象の説明で検討を加えられてきたのは2次関数だけではない。ゴンペルツ関数もその代表であるが、TrajでもMplusでも取り扱うことができない。Browne（1993）は、因子パターン行列 $A$ において、指数・対数・ゴンペルツ関数を定義し、この関数のパラメータを推定する非線形モデルを提案し、構造的潜在曲線（structured latent curve）モデルと呼んでいる。Mx（Neale, 1998; Dolan,

Schmittman, Lubke, & Neale, 2005) や LISREL の最新版では、この理論を実際のデータ解析に応用するが可能であることを指摘しておきたい。

ここで議論してきたように、方法論としての魅力という点では一歩引けをとる Traj をこのように紹介してきたのは、Mplus の分布条件などの敷居が高いことだけが理由ではない。少ない数のデータにおいても潜在する集団を探し出すことができれば、応用的研究の可能性を広げることができると考えたからである。そして、混合モデルの目的そのものが適切な数の潜在する集団を探索することにあると考えたからである。

混合モデルでのメンバシップ確率と所属集団の特定する方法（アルゴリズム）については、本稿では、省略してきた。Yung (1997) が CFA を対象として詳細な検討をおこなっている。Bollen & Curran (2006) も EM 法による計算過程を簡潔に紹介している。基本的には、個人  $i$  が集団  $g$  に属す確率と軌跡の関数式の推定を EM 法でおこなうことで、その個人が属する集団を特定する。Amos の混合モデル分析では、パラメータの推定過程で、個人の属する確率を集団ごとに表示してくれるので、この方法の理解には有効であるかもしれない。混合モデルは、結果として、(1) 式の  $p_i^{(g)}$  の値から個人の属する集団を確定する方法で、ある種のクラスター分析とも考えることができる。

#### 4-2. 混合モデルによるクラスタ化とその応用

本稿の分析では、10 年間にわたって高いレベルでの打撃成績を残した 4 人を Traj で 1 つの集団として特定することができた。この 4 人は、対象とした資料の中では、傑出した打者であったことは、誰もが認めるのではないだろうか。データ解析の観点からみると、この 4 人という数からでは正則な共分散行列を算出することは困難となる。すなわち、下位集団ごとに (3) 式を表現することは不可能となり、Mplus の混合モデル分析では、探し出すことのできない集団といえる。

混合モデルは、対象となる集団全体の中に潜在する異質な下位集団を探し出すことを目的として展開されてきた (Muthén, 2002)。今回の例では、全体を対象とした分析からは、LGM でも Traj の分析でも、適合度の良い 2 次関数モデルを得ることができていた。すなわち、(3) 式あるいは (2) 式での 1 つの方程式で全体を集約することができるという結論を得たことになる。言い方を変えると、内部に適合度を悪くするほどの異質なメンバーが潜んでいない、ということになる。しかしながら、表 6 と図 4 で検討してきたように、関数としては同じでも、この形が 2 次関数において異なる 6 つの集団を、繰り返すようであるが、特定することに Traj は成功した。この 6 集団が、混合モデルでいうところの異質な

集団であるかどうかは、議論が残るところではあるが、下位集団に所属するメンバーをみると、この下位集団間の違いは了解できるところでもある。

集団全体を対象とした介入が効果的なものであるための条件の1つは、この集団がそこで採用した介入法に敏感に反応するという点で等質的なメンバーから構成されていることではないだろうか。横断的な個人間差を手がかりに集団の内部を分類することで、介入の効果をあげる方法も考えられる。混合モデルは1回の測定にも適用できる (Muthén, 2002)。

変化パターンの個人間差がある場合には、それぞれに適した介入が効果を上げることは容易に想像ができる。ここで解析してきたような6つの変化パターンを示す集団には、それぞれの特徴に合わせたトレーニング・プログラムを、この分野の専門家でなくとも、イメージすることができる。服部 (2006) も強調するように、指導・介入をより効果的なものとする方途として、混合モデルの応用利用が期待されるわけである。

変化のパターンは、2回以上の繰り返し測定であれば、ここで紹介してきた方法論を適用することで見つけ出すことができる。清水・花井・宮坂・松下 (2008) は、1日のキャリア教育プログラムの前後に測定したキャリア関連行動の尺度の分析にTrajを適用して、『計画立案』に関する自己効力感で4つの得点レベルの異なる下位集団を、そして、意思決定の『決定不安』でも同じく4つの下位集団を、特定している。Guay, Ratelle, Senécal, Larose, & Deschênes (2006) は、キャリア不決断についての3回の繰り返し測定から、決断状態で推移した集団、不決断状態で推移した集団、そして、不決断から決断へと発達に変化した集団を、Trajを使用して特定している。Mplusの混合モデルによる解析を適用した研究の報告も増えてきている。

ここで取り上げたデータにはイチロー選手は含まれていない。大リーグへ移籍した後の記録は、日本のプロ野球記録では欠損の扱いとなる。10年間に限定したことは、プロ野球選手のキャリアを正確に評価したことにはならないかもしれない。表1では、15年間を示した。キャリアの全体をカバーするには、これでも足りない。対象の期間を長くすれば、表1にあるように欠損値が増えてくる。Trajは、欠損値の処理にも対応している (Dodge, Shen, & Ganguli, 2008)。二値データへの関数も準備されている。そして、複数変数の同時モデル化なども可能である (Jones & Nagin, 2007))。今後の課題の1つとしたい。

#### 4-3. 横断研究から縦断研究へ

最後に強調したいことは、LGM系の分析に1,000名を超えるような大規模な縦断的デー

タの存在を絶対的な条件とはしないしてほしい、ということである。100程度あるいは200程度のデータであっても、ここで示したような結果を解析することが可能なのである。

横断的研究では、図4のキャリア初年をこの姿で捉えることはできない。そして、このデータでの121名を同時に捉えることもできない。ある年にデータを収集すれば、そこには、キャリア初年の選手から離脱直前の選手までの成績情報が混在することになる。そのような横断的データからでも古典的な意味での個人間差は観察できる。それぞれの経歴を踏まえたものとしてだけでなく、同期にキャリアを開始した者の間でも個人差は存在する。しかし、このようにして横断で捉えた個人が、図4のいずれの軌跡（下位集団）に属するのは、下位集団3と5の1年をみれば分かるように、判断することができない。

本稿で検討してきた例は10回の繰り返し測定を1年間隔の管理の下で整理したものに相当する。繰り返しの回数は、軌跡の形を特定することに貢献することは確かである。潜在するクラスタを探索することと繰り返し測定の回数と、どのような関係にあるのかなど、ここまでで指摘した方法論上の課題もあわせて、追求すべきことはたくさん残されている。残されている課題の回答へは、理論的追求だけではなく、実際のデータでの追求からも追えることができるのではないかと考えて、少々古くなったデータではあるが、本稿の試みをおこなってみたいわけである。

2回以上の繰り返し測定は、再検査信頼性を得るために、数多く試みられている。2回の繰り返し測定で得られることの意味合いは、例えば、図4を2時点として、キャリア初年と6年目との間で6下位集団の値を繋ぐことで、このことはイメージできるのではないだろうか。実際のデータ解析でも、先にも紹介したように、清水ほか（2008）は2回の繰り返し測定から4つのレベルでの下位集団を報告しているように、潜在するクラスタを取り出すことは可能である。

欧米には、縦断データの収集の伝統が1920年代の知能研究のはじまりの頃からある。LGM系の分析の対象は心理的変数にかぎらない。Bollen & Curran（2006）は、ニューヨーク州の359の地域の犯罪率を取り上げている。社会科学あるいは行動科学の研究デザインは縦断的データの集積とその解析を中心として展開するようになってきた。このような分野での最近の動向をLGMや混合モデルも含めて整理したものとしてMenard（2008）がある。データの存在は方法論を刺激する。これまでにも言及してきたように、LGMの歴史がそのことを証明している。

わが国では、縦断的データの集積はこれからという感がある。SEMを応用する研究にもその感がある。心理学諸学会の一部には、統計的なevidenceを重要視しないという風潮

もある。古典的な意味での個人間差を批判することにはそれなりの意味があったかもしれない。繰り返しになるが、どれほど大規模なデータであろうとも1回限りのデータから何らかの結論を引き出そうとする研究スタイルから脱却することを主張したい。介入という活動とその評価をおこなうためには、追求の対象を変化パターンに表れる個人間差に求めることを提案したい。

注：本研究の一部は、清水（2003）のデータをそのまま使用している。データの入力では、ゼミ卒業生の上野利恵さん（現、株式会社総合企画センター大阪）の努力に負うところが大きい。卒業後もデータの管理で貢献してくれた上野さんに心から感謝いたします。なお、結果の一部は、清水（2000）でも発表している。1999年以降の記録については、収集・分析の対象とはしていないが、本稿の目的は、このデータで十分に果たせたのではないだろうか。

#### 引用文献

- Arrandale, V., Koehoorn, M., MacNab, Y., & Kennedy, S. M. (2006). *How to use SAS® Proc Traj and SAS® Proc Glimmix in respiratory epidemiology*.
- Bollen, K. A. & Curran, P. J. (2006). *Latent curve models: A structural equation perspective*. Hoboken, NJ: Wiley.
- Browne, M. W. (1993). Structured latent curve models. In C. M. Cuadras & C. R. Rao (Eds.), *Multivariate analysis: Future directions 2*. (pp. 171-197). Amsterdam: North-Holland.
- Clogg, C. C. (1995). Latent class models. In G. Arminger, C. C. Clogg, & M. E. Sobel (Eds.), *Handbook of statistical modeling for the social and behavioral sciences* (pp. 311-359). New York: Plenum.
- Dodge, H. H., Shen, C., & Ganguli, M. (2008). Application of the pattern-mixture latent trajectory model in an epidemiological study with non-ignorable missingness. *Journal of Data Science*, 6, 231-246.
- Dolan, C. V., Schmittmann, V. D., Lubke, G. H., & Neale, M. C. (2005). Regime switching in the latent growth curve mixture model. *Structural Equation Modeling*, 12, 94-119.
- Guay, F., Ratelle, C. F., Sénécal, C., Larose, S., & Deschênes, A. (2006). Distinguishing developmental from chronic career indecision: Self-efficacy, autonomy, and social support. *Journal of Career Assessment*, 14, 235-251.
- 服部 環 (2006) 成長混合モデリングとその分析事例—岡林論文・高比良論文・山形論文に関連して—  
パーソナリティ研究, 15, 129-134.
- Jones, B. L. & Nagin, D. S. (2007). Advances in group-based trajectory modeling and SAS Procedure for estimating them. *Sociological Methods Research*, 35, 254-571.
- Jones, B. L., Nagin, D. S., & Roeder, L. (2001). A SAS Procedure based on mixture models for estimating developmental trajectories. *Sociological Methods Research*, 29, 374-393.
- 狩野 裕・三浦麻子 (2002) グラフィカル多変量解析 (増補版) 現代数学社.
- McArdle, J. J. (1986). Latent variable growth within behavior genetic models. *Behavior Genetics*, 16, 163-200.
- McArdle, J. J. (1988). Dynamic but structural equation modeling of repeated measures data. In J. R.



- Nesselroade & R. B. Cattell (Eds.), *Handbook of multivariate experimental psychology* (2nd ed.) (pp. 561-614). New York, NY: Plenum.
- Menard, S. (Ed.) (2008). *Handbook of longitudinal research: Design, measurement, and analysis*. Burlington, MA: Elsevier.
- Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, *55*, 107-122.
- Muthén, B. O. (2002). Beyond SEM: General latent variable modeling. *Behaviormetrika*, *29*, 81-117.
- Muthén, L. K. & Muthén, B. O. (1998-2007). *Mplus user's guide*. Los Angeles, CA: Muthén & Muthén.
- Nagin, D. S. (1999). Analyzing developmental trajectories : A semiparametric, group-based approach. *Psychological Methods*, *4*, 139-157.
- Neale, M. C. (1998). *Mx statistical modeling*. Unpublished program manual, Department of Human Genetics, Medical College of Virginia, Virginia Commonwealth University, Richmond, VA.
- Nesselroade, J. R. (1991). Interindividual differences in intraindividual change. In L. M. Collins & J. L. Horn (Eds.), *Best methods for analysis of change* (pp.92-105). Washington, DC: American Psychological Association.
- 日本野球機構 (1999) The Official Baseball Encyclopedia '98 ベースボール・マガジン社.
- 岡林秀樹 (2006) 発達研究における問題点と縦断データの解析方法 パーソナリティ研究, *15*, 76-86.
- 清水和秋 (1999a) 潜在成長モデルによる進路成熟の解析—不完全コーホート・データへの適用— 関西大学社会学部紀要, *30* (3), 1-47.
- 清水和秋 (1999b) キャリア発達の構造的解析モデルに関する比較研究 進路指導研究, *19* (2), 1-12.
- 清水和秋 (2000) 熟達の過程—潜在成長モデルによる野球データの解析— 日本行動計量学会第28回大会発表論文抄録集, 379-382.
- 清水和秋 (2003) プロ野球選手の熟達の軌跡—潜在成長モデルの打撃成績への適用— 悠峰職業科学研究所紀要, *11*, 63-72. .
- 清水和秋・花井洋子・宮坂吉有樹・松下眞治 (2008) キャリア教育の効果測定—混合モデルによる変化パターンのクラスタ化の試み—日本キャリア教育学会第30回研究大会発表論文集, 122-123.
- Thissen, D., & Bock, R. D. (1990). Linear and nonlinear curve fitting. In A. von Eye (Ed.), *Statistical methods in longitudinal research (Vol. II): Time series and categorical longitudinal data*. (pp.289-318). New York: Academic Press.
- 上野利恵 (2000) 加齢に伴うプロ野球選手の打撃成績の変化に関する研究—現役選手と一流OB選手との比較— 関西大学社会学部産業心理学専攻卒業論文 (未公刊).
- Yung, Y.-F. (1997). Finite mixtures in confirmatory factor-analysis models. *Psychometrika*, *62*, 297-330.

注 : SAS® Proc Traj の入手先

<http://www.andrew.cmu.edu/user/bjones/index.htm>

— 2008. 8. 21受稿 —