新潟大学教育学部附属教育実践研究
指導センター研究紀要，7 （1988）

S－P表における注意係数の判定基準の改善の試み

南風原　朝　和*　青　木　和　彦**

問題と目的

生徒は、返却されたテストの得点を見て一喜一憂する。教師は、どの問題ができてどの問題ができなかったかを見て復習するよう生徒に指導するが、済んでしまったテストの問題には生徒はほとんど関心を示さない。こうした状況は、小学校から大学まで至る所で見られる。テストが、生徒の今後の学習のために生かされることなく、単なる成績づけに終わってしまうことが多いのである。

しかし、教師の側においても、自分が実施したテストで、どの生徒が何点とったか、誰がどれだけ伸びたか、クラスの平均は何点ぐらいだったかというように、結局はテストの得点だけにしか注目していないというケースが多いのではないだろうか。その場合、たとえばクラスの平均が期待したより低くても、次回はもっと頑張るようにとか、きちんと復習しておくようにといったような、曖昧な指導しかできないであろう。

このようなときに、テストの問題ごとの正答者数まで調べてあれば、復習させる場合や補習授業を行う場合に焦点を絞ることができる。さらに、個々の生徒について、誰がどの問題で正答し、どの問題で誤答したかを把握していれば、個別に宿題を与えたり指導を加えたりすることによって、不十分だった学習を補うことが可能になる。

このように、個々の生徒の、個々のテスト項目に対する正答・誤答のパターン(項目反応パターン)を一覧表の形にしておくことは、学習診断においてきわめて有用な手続きである。こうした項目反応パターンには、1つの得点（合計点）にはとても集約できないほど豊かな情報が含まれている。そのような、いわば質的な情報を取り出すには、一覧表をそのまま眺めるだけでなく、目的に応じた指標を各生徒の項目反応パターンから算出して利用すると便利なことがある。

項目反応パターンから得られる指標の中で、かなり実用化されているものに、佐藤（1975）のS－P表における注意係数がある。この指標は、各生徒の正答・誤答のパターンが、項目の困難度とどの程度一貫しているかを示すもので、正答者数の多い項目に正答し、正答者数の少ない項目に誤答しているというパターンなら0に近い値をとり、そのパターンからずれるほど1に近い値をとる。そして、ある一定の値を超える注意係数を示す生徒は、項目反応パターンが集団全体から見て異質であり、注意して調べてみる必要があるとしている。

その一定の値、すなわち注意係数の判定基準としては、しばしば0.50という値が用いられる。この基準は、〝教育現場の多数の実例によって検討した結果〟として得られたものであり、〝経験上推奨すべき基準である〟（佐藤、1975、p.67）とされているので、ある程度妥当なものと考えてよいかも

*　新潟大学教育学部
**　新潟大学大学院教育学研究科

しれない。しかし、厳密に考えると、項目数や生徒の能力レベルに関係なく一定の判定基準を用いてよいのかといった点など、この基準に対する疑問も残る。

本論文の目的は、注意係数の判定のための改善された方法を提案することである。この方法は、統計的仮説検定の考えを取り入れたもので、同じ値の注意係数でも、生徒の能力レベルによって異なる判定がなされる可能性があるという特徴がある。具体的には、ある理想的な条件のもとでの、テストの繰り返し試行における注意係数の変動の大きさを推定し、実際に得られた注意係数の値をこの変動の大きさに照らし合わせて判定するというものである。

S－P表と注意係数

注意係数を判定するための改善された方法を提案する前に、S－P表と注意係数について簡単に説明しておく。

S－P表は、図1に示したように、生徒（Student：S）×問題項目（Problem：P）の項目反応データを、得点の高い生徒ほど上にくるように、また正答者数の多い項目ほど左にくるように並べ替え、その表中にS曲線とP曲線を記入したものである。S曲線は、各生徒について、左から順に得点の分だけ数えて区切りのマーク（図1では、/）を入れ、それを結ぶことによって得られる。またP曲線は、各項目について、上から順に正答者数の分だけ数えて区切りのマーク（図1では、_）を入れ、それを結ぶことによって得られる。S曲線とP曲線は、それぞれ得点と正答者数の累積度数分布となっている。

問題番号

生徒番号	2	3	13	10	1	6	4	11	12	5	8	7	15	14	9	得点	注意係数
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1 /	15	0.00
20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	/0	14	0.00
22	1	1	1	1	1	1	1	1	1	1	1	1	1	1	/0_	14	0.00
9	1	1	1	1	1	1	1	1	1	1	1	1	1	/0	0	13	0.00
24	1	1	1	1	1	1	1	1	1	0	1	1	1	/1	0	13	0.37
12	1	1	1	1	1	1	1	1	0	1	1	1	1	/1	0	13	0.41
19	1	1	1	1	1	1	1	1	1	1	0	1	1	/0	1	13	0.61
8	1	1	1	1	1	1	1	1	1	1	1	1	/0	0	0	12	0.00
3	1	1	1	1	1	1	1	1	1	1	1	0	/1	0	0	12	0.03
14	1	1	1	1	1	1	1	1	1	0	1	1	/1	0_	0	12	0.17
7	1	1	1	1	1	1	1	1	0	1	1	1	/1	0	0	12	0.21
25	1	1	1	1	1	1	0	1	1	1	1	1	/0	1	0	12	0.38
16	1	1	1	1	1	1	1	1	1	1	1	0	/0	0	1	12	0.41
21	1	1	1	1	1	1	1	1	1	1	0	/0	1_	0	0	11	0.12
2	1	1	1	1	1	0	1	1	1	1	1	/0_	1	0	0	11	0.21
17	1	1	1	1	1	0	1	1	1	1	0	/1	1	0	0	11	0.30
23	1	1	1	0	1	1	1	1	1	1	0	/1	0	1	0	11	0.49
10	1	1	1	1	1	1	1	0	1	1	/1_	0	0	0	0	10	0.06
6	1	1	1	1	1	1	1	1	0	1_	/0	0	0	1	0	10	0.30
5	1	1	1	1	1	1	1	1_	0_	0	/1	0	0	1	0	10	0.33
4	1	1	1	1	1	1_	0_	0	1	0	/1	0	1	1	0	10	0.53
13	1	1	1	1	0_	1	0	0	1	/1	1	1	0	0	0	9	0.33
18	1	1	1	1_	1	1	1	0	/1	0	0	0	0	0	0	8	0.00
11	1	1	1	1	0	0	1	/0	1	1	0	0	0	0	0	7	0.13
15	1_	1_	1_	0	0	0	/0	1	0	0	1	1	0	0	0	6	0.45
正答者数	25	25	25	23	22	21	21	20	20	19	18	15	14	10	3		
注意係数	0.00	0.00	0.00	0.42	0.08	0.33	0.47	0.25	0.68	0.52	0.59	0.44	0.19	0.48	0.32		

図1　S－P表の例

前節で述べた生徒の注意係数は、次の式で算出される。

生徒の注意係数＝（A－B）／（C－D×E） (1)

この式における各項は、以下のような値である。

A＝その生徒のS曲線から左の0に対応する正答者数の和

B＝その生徒のS曲線から右の1に対応する正答者数の和

C＝その生徒のS曲線から左の項目の正答者数の和

D＝その生徒の得点

E＝平均正答者数

もし、ある生徒がかなり易しい項目に誤答し、逆にかなり難しい項目に正答していれば、(1)式の中のAの値は大きくなり、Bの値は小さく抑えられる。その結果として(1)式の分子が大きくなり、注意係数の値が大きくなるということになる。(1)式の分母のほうは、注意係数の最大値がほぼ1となるように規準化するためのものと考えればよい。図1には、各生徒について計算した注意係数の値も示してある。

S－P表においては、生徒の注意係数だけでなく、項目のほうの注意係数も同じような考え方で定義できる。その計算式は次式で与えられる。

項目の注意係数＝（a－b）／（c－d×e） (2)

この式における各項は、以下のような値である。

a＝その項目のP曲線から上の0に対応する得点の和

b＝その項目のP曲線から下の1に対応する得点の和

c＝その項目のP曲線から上の生徒の得点の和

d＝その項目の正答者数

e＝平均得点

注意係数の値の大きな項目は、得点の低い生徒が正答し、得点の高い生徒が誤答しているような、全体から見て異質な項目である。図1には、各項目の注意係数の値も示されている。なお、佐藤（1975）は、項目の注意係数についても、0.50という判定基準を推奨している。

方法の概要

まず、生徒iの項目jにおける得点を、確定的なものではなく、試行ごとに変動する可能性のある確率変数と考え、U_{ij}で表す。U_{ij}は正答のときは1、誤答のときは0という値をとり、

$$\pi_{ij} = \mathrm{Prob}\,(U_{ij} = 1) \tag{3}$$

が、繰り返し試行における正答の確率を表すものとする。このように試行ごとに得点が変動する可能性を認めることは、測定にランダムな誤差が含まれている可能性を認めることであり、きわめて自然な仮定である。

ここで、正答者数の多い項目に対するπ_{ij}は高く、正答者数の少ない項目に対するπ_{ij}は低いというように、項目の困難度と正答確率の順位が完全に対応している生徒を想定する。そのような〝理想的な〟生徒でも、1回の試行で実際に与える項目反応パターンにおいて、注意係数が0になるとは限

らない。このとき、注意係数の値は、その生徒の正答確率π_{ij}のパターンによって規定される確率分布に従って変動する。すなわち、その確率分布は、項目の困難度と個人の正答確率が完全に対応しているという意味で理想的な生徒において、注意係数がどんな確率でどんな値をとるかを示すものである。

もし、そのような理想的な条件のもとでの個人ごとの注意係数の確率分布が求められれば、実際に得られた注意係数の値が、その分布においてどの程度得られにくいものであるかを評価することができる。そして、その分布において、実際に得られた値以上の注意係数が得られる確率が非常に小さい（たとえば５％以下）ならば、その生徒の正答確率のパターンは理想的なものではないと判断することができるのである。言い換えれば、正答確率のパターンが理想的なものであるという仮説を、データに基づいて検定するという手続きである。

しかしながら、注意係数は、項目得点U_{ij}のかなり複雑な関数であるため、その確率分布を数学的に導くのは容易でない。そこで、上記のような理想的な条件を設定した上で、コンピュータによってU_{ij}のパターンを多数発生させ、注意係数の変動を実験的にとらえるというシミュレーションを採用することにする。このとき、U_{ij}を発生させるためには、各生徒の各項目に対する正答確率π_{ij}を推定しておくことが必要であり、さらにそのπ_{ij}は理想的なパターンに従ったものでなければならない。そのような推定のためには、項目反応に関する何らかのモデルを仮定する必要がある。ここでは、そのモデルとして、項目反応モデルの中で数学的に最も単純なラッシュモデルを用いることにする。

ラッシュモデル

ラッシュモデルにおいては、同一の尺度上に生徒の能力を表すパラメータθと項目の困難度を表すパラメータｂを定義し、その差（$\theta-b$）の関数として、項目への正答確率(項目反応関数という)を

$$P(\theta)=exp(\theta-b)/\{1+exp(\theta-b)\} \tag{4}$$

で与える。

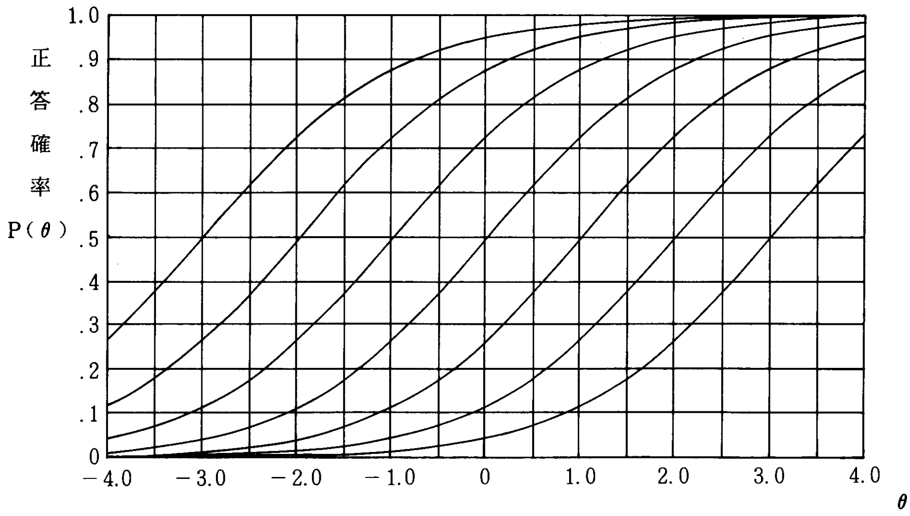

図２　ラッシュモデルによる項目反応関数
（項目の困難度は左から順にｂ＝－３、－２、－１、０、１、２、３）

図2には、左から順にb＝－3、－2、－1、0、1、2、3の7個の項目の項目反応関数を$-4<\theta<4$の範囲で描いてある。この図で明らかなように、項目困難度bが大きい項目ほど、その項目反応関数のグラフが尺度の右側に位置し、能力θがかなり高くないと正答できないことになる。

ラッシュモデルの特徴のうち、本論文の目的との関連で最も重要なものは、能力および困難度のパラメータと正答確率との関係の単調性である。すなわち、任意の生徒iについては、項目jへの正答確率$P_j(\theta_i)$が、その項目の困難度b_jの単調減少関数となり、また任意の項目jについては、生徒iの正答確率$P_j(\theta_i)$が、その生徒の能力θ_iの単調増加関数となるのである。したがって、ラッシュモデルにおける正答確率$P_j(\theta_i)$をデータから推定し、それを(3)式の正答確率π_{ij}の推定値として用いれば、すべての生徒について、項目の困難度と完全に対応した理想的な正答確率の推定値が得られることになる。

正答確率の推定のアルゴリズムと精度

ラッシュモデルにおける正答確率$P_j(\theta_i)$を推定するには、まず実際に得られた項目反応パターンから、生徒iの能力θ_iと項目jの困難度b_jを推定し、それらの推定値θ_i、b_jを(4)式に代入することになる。ラッシュモデルにおけるパラメータを正確に推定するには、最尤法などを用いることができる（Swaminathan、1983）。また、複雑な数値計算を含まない近似法としては、Cohen（Wright、1977）による以下のようなアルゴリズムがある。

1. はじめに、満点または零点の生徒と、全員正答または全員誤答の項目を削除する。
2. 残ったM項目のそれぞれについて正答者数s_jを求める。
3. 得点がrとなる生徒の人数n_rを求める。（人数の総和をNとする。）
4. 以下の値を計算する。

$$x_j=\log_e[(N-s_j)/s_j] \quad (5)$$

$$x.=\Sigma x_j/M \quad (6)$$

$$U=\Sigma(x_j-x.)^2/(M-1) \quad (7)$$

$$y_r=\log_e[r/(M-r)] \quad (8)$$

$$y.=\Sigma n_r y_r/N \quad (9)$$

$$V=\Sigma n_r(y_r-y.)^2/(N-1) \quad (10)$$

$$X=[(1+U/2.89)/(1-UV/8.35)]^{\frac{1}{2}} \quad (11)$$

$$Y=[(1+V/2.89)/(1-UV/8.35)]^{\frac{1}{2}} \quad (12)$$

5. 最後に、項目jの困難度の推定値b_jと、得点がrの生徒の能力の推定値$\theta_{(r)}$を次の式で求める。

$$b_j=Y(x_j-x.) \quad (13)$$

$$\theta_{(r)}=X y_r \quad (14)$$

次節に示す数値例では、計算上の負担を軽くするために、この近似法を用いている。そこで、このアルゴリズムによる正答確率$P_j(\theta_i)$の推定がどの程度正確なものであるかを評価するために、以下のようなシミュレーションを行った。

まず生徒数を20、30、40、60、80の5通り設定し、項目数を10、20、30、40の4通り設定した。そ

して、それらの生徒数と項目数の組合せのそれぞれにおいて、生徒の能力を標準正規分布に従ってランダムに発生させ、同様に項目の困難度も標準正規分布に従ってランダムに発生させた。こうして得られる能力と困難度を(4)式に代入して、各生徒の各項目への〝真の〟正答確率$P_j(\theta_i)$を計算し、その確率に基づいて項目得点（正答なら1、誤答なら0）を発生させた。そして、それをデータとしてあらためて各生徒の能力の推定値$\hat{\theta}_i$と各項目の困難度の推定値$\hat{b}_j$を計算し、(4)式によって正答確率の推定値$\hat{P}_j(\hat{\theta}_i)$を求めた。

この推定値が、データ発生に用いられた真の正答確率$P_j(\theta_i)$にどれだけ近いかを評価するために、まず推定値と真値との差の絶対値を求め、それをすべての生徒とすべての項目の組合せにわたって平均した。そして、以上の手続きを10回繰り返し、10回分の平均絶対誤差の平均を求めた。その結果を示したのが図3である。

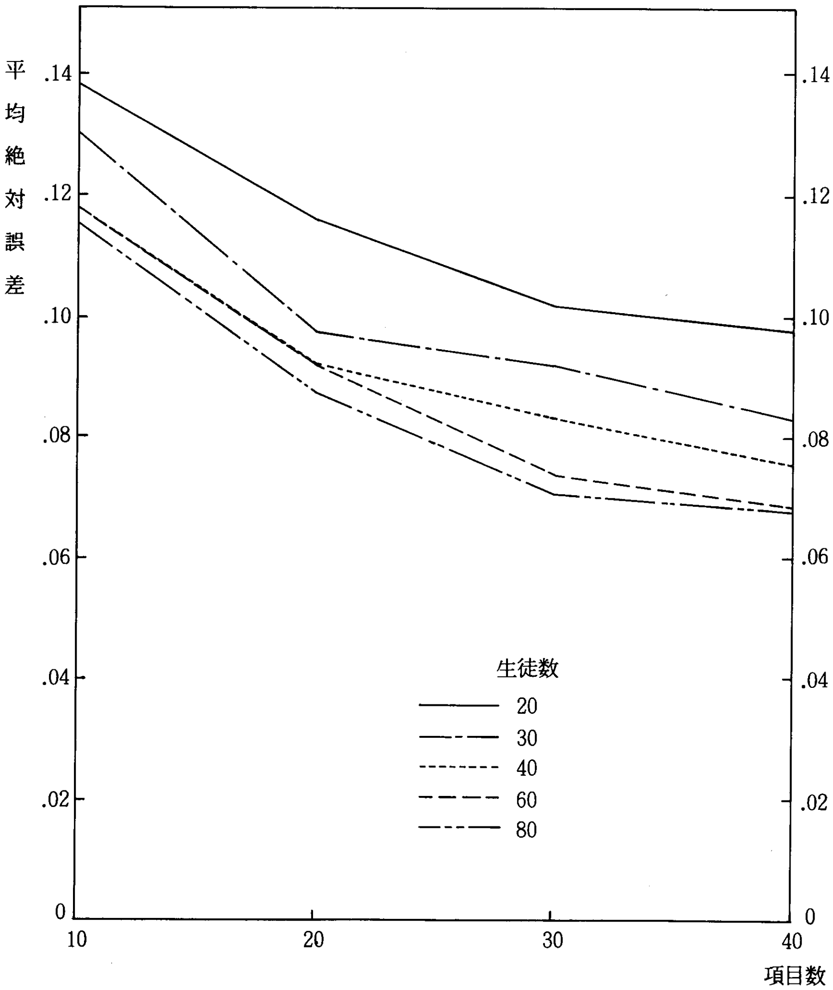

図3　正答確率$P_j(\theta_i)$の推定値の平均絶対誤差

この図を見ると、生徒数が多いほど、また項目数が多いほど、正答確率の推定の精度が高くなるということがわかる。ここで、仮に平均絶対誤差が 0.1 以下であれば実用的に有用であると考えるとすると、その条件を満たすには、項目数が20であれば30人以上の生徒が必要であり、項目数が30以上であれば生徒数が約20人以上あればよいということになる。この結果は、限定された条件のもとで得られた結果に過ぎず、あまり一般化するのは危険であるが、現実に注意係数を算出するような状況において、ラッシュモデルを仮定した方法を導入することが、それほど無理なことではないということは、少なくとも言えるであろう。

数 値 例 (1)

理想的な条件のもとでの注意係数の変動を調べ、実際に得られた注意係数の値をその分布に照らして判定するという、ここで提案した方法を、実際のデータに適用してみた。使用したテストは30個の多肢選択項目からなる統計学のテストで、〝生徒〞は49人の大学生である。このテストにおける平均得点は 20.1 点で、平均正答者数は 32.8 人、アルファ係数は 0.68 であった。

はじめに前節で示したアルゴリズムによって、各生徒の能力および各項目の困難度を推定し、それを(4)式に代入して各生徒の各項目に対する正答確率を推定した。そして、それに基づいて、すべての生徒のすべての項目における得点を発生させることを 100 回繰り返した。各回の繰り返しにおいて、それぞれの生徒の注意係数が算出されることになるが、その 100 回のうち何回まで実際に得られた注意係数の値以上の値が得られたかを数え、それを 100 で割って注意係数の有意水準を推定した。この推定された有意水準が小さいほど、理想的な条件のもとでは得られにくい大きさの注意係数が得られたことになり、要注意であるということになる。

表 1　生徒の注意係数と推定された有意水準
（奇数順位の生徒のみ）

順　位	得　点	注意係数	有意水準	平　均	標準偏差	標準化された注意係数
1	28	.66	.03	.215	.242	1.839
3	25	.29	.41	.257	.176	0.188
5	24	.36	.35	.275	.169	0.503
7	23	.13	.76	.279	.152	− 0.980
9	23	.27	.57	.279	.152	− 0.059
11	23	.31	.32	.279	.152	0.204
13	23	.45	.13	.279	.152	1.125
15	22	.15	.75	.273	.141	− 0.872
17	22	.24	.55	.273	.141	− 0.234
19	22	.34	.26	.273	.141	0.475
21	21	.26	.37	.267	.139	− 0.050
23	21	.27	.35	.267	.139	0.022
25	20	.13	.89	.278	.145	− 1.021
27	20	.22	.63	.278	.145	− 0.400
29	20	.33	.40	.278	.145	0.359
31	19	.10	.91	.285	.142	− 1.303
33	19	.20	.64	.285	.142	− 0.599
35	19	.33	.40	.285	.142	0.317
37	19	.46	.14	.285	.142	1.232
39	19	.55	.02	.285	.142	1.866
41	17	.27	.45	.267	.124	0.024
43	16	.18	.70	.267	.137	− 0.635
45	14	.23	.61	.270	.122	− 0.328
47	12	.62	.00	.263	.135	2.644
49	11	.36	.17	.253	.132	0.811

表1は、49人の生徒のうち、得点の高い順に並べたときの奇数順位の生徒について、得点、注意係数、推定された有意水準、そして100回の繰り返しにおける注意係数の平均と標準偏差を示したものである。この表の最右欄の値は、実際に得られた注意係数の値から平均を引いて標準偏差で割ることによって標準化した値である。この値も、理想的な条件のもとでの注意係数の分布からの実測値の逸脱度を示すものである。なお、同得点の生徒がいる場合には、平均も標準偏差も、同得点の者の間で平均することにした。それは、ラッシュモデルにおいては、得点の等しい生徒の能力推定値は等しくなり、したがって、データ発生のために用いられる正答確率も等しくなるからである。

この表を見ると、1番と39番と47番の生徒の注意係数の有意水準が0.05 以下と推定されていることがわかる。この例では、注意係数が0.50を超えているのもこの3人の生徒であるから、ここでは、現在一般的に用いられている判定基準と有意水準0.05という判定基準が同じ結果を与えている。しかし、注意係数の値では1番の生徒のほうが47番の生徒より高いのに、有意水準は47番の生徒のほうが小さくなっていることからもわかるように、この2つの判定基準は一般には異なる結果をもたらすことになる。なお、標準化された注意係数の値が標準正規分布の上側5％の値1.645を超える生徒を取り出してみると、やはり同じ3人の生徒が選ばれることになる。

次に、有意水準や標準化された注意係数の値に大きな影響を与える標準偏差が、生徒の能力レベルとどう関係するかを見るために、注意係数の標準偏差と得点との散布図を描いてみた（図4）。この図を見ると、一般に得点の高い生徒ほど注意係数の標準偏差も大きくなる傾向が読みとれる。ただし、この傾向は低い得点の範囲では見られない。（実際、低い得点の範囲では、逆に得点が低いほど標準偏差が大きくなる傾向があることが次節の例において示される。）なお、この図においては、同得点の生徒の標準偏差のばらつきの程度がわかるように、平均せずにそのまま示した。

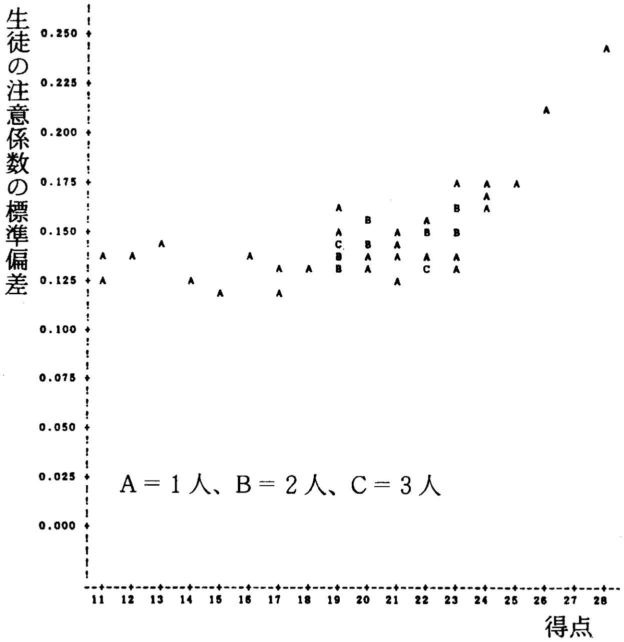

図4　生徒の注意係数の標準偏差と得点の散布図
（実際のデータから得られたもの）

次に、項目の注意係数について、100回の繰り返しに基づく有意水準の推定値、そして平均、標準偏差、標準化された注意係数を表2に示す（奇数順位の項目のみ)。この表を見ると、もとの注意係数では約半数の項目が0.50を超えていて要注意とみなされることになるが、推定された有意水準では29番の項目だけが0.05以下となっている。また、標準化された注意係数が1.645を超えるのもこの項目だけである。全体から見て異質な項目を発見するのが注意係数の役割であるが、全項目のうちの半数が異質であるという判定をもたらす基準は、はたして妥当なものと言えるだろうか。

表2　項目の注意係数と推定された有意水準
（奇数順位の項目のみ）

順　位	正答者数	注意係数	有意水準	平　均	標準偏差	標準化された注意係数
1	48	.66	.36	.448	.397	0.534
3	47	.44	.48	.447	.366	− 0.019
5	45	.69	.34	.566	.259	0.479
7	42	.70	.21	.521	.205	0.873
9	41	.58	.25	.518	.179	0.346
11	41	.77	.09	.518	.179	1.408
13	40	.69	.19	.497	.201	0.960
15	36	.36	.72	.464	.155	− 0.671
17	32	.36	.77	.466	.136	− 0.779
19	30	.55	.24	.453	.145	0.669
21	27	.48	.47	.478	.145	0.014
23	21	.48	.43	.463	.146	0.116
25	19	.47	.40	.458	.151	0.079
27	5	.45	.48	.455	.205	− 0.024
29	5	.83	.04	.455	.205	1.829

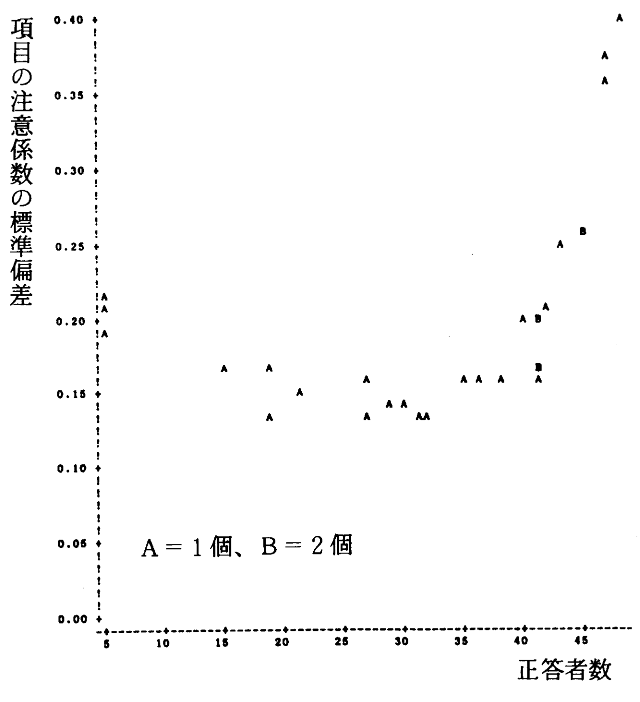

図5　項目の注意係数の標準偏差と正答者数の散布図
（実際のデータから得られたもの）

図5は、項目の注意係数の標準偏差と正答者数との関係を示したものである。この図では、正答者数の尺度の中央部分から左右に隔たるほど注意係数の標準偏差が大きくなる傾向があることが、はっきりと読みとれる。

数　値　例 (2)

先の数値例では、あまり得点の低い生徒がいなかったために、注意係数の変動の大きさと生徒の能力レベルとの関係が、はっきりとしない部分があった。そこで、今度はシミュレーションによって広い範囲の能力値を発生させて検討してみた。

まず生徒数を100、項目数を30として、生徒の能力と項目の困難度を平均0の正規分布を用いて発生させた。ただし、生徒の能力が項目の困難度の分布に対して十分広く分布するように、能力分布の標準偏差を1.5、困難度分布の標準偏差を1とした。

このような設定のもとで項目反応データを発生させることを100回繰り返して得られた結果を示したのが図6である。この図の横軸は、仮想的な生徒のひとりひとりについて、項目の正答確率を全項目にわたって合計したもので、その生徒の期待得点または真の得点とよばれるものである。この図は、期待得点が満点または零点に近い生徒ほど、注意係数の変動の大きさが大きいということを明確に示している。

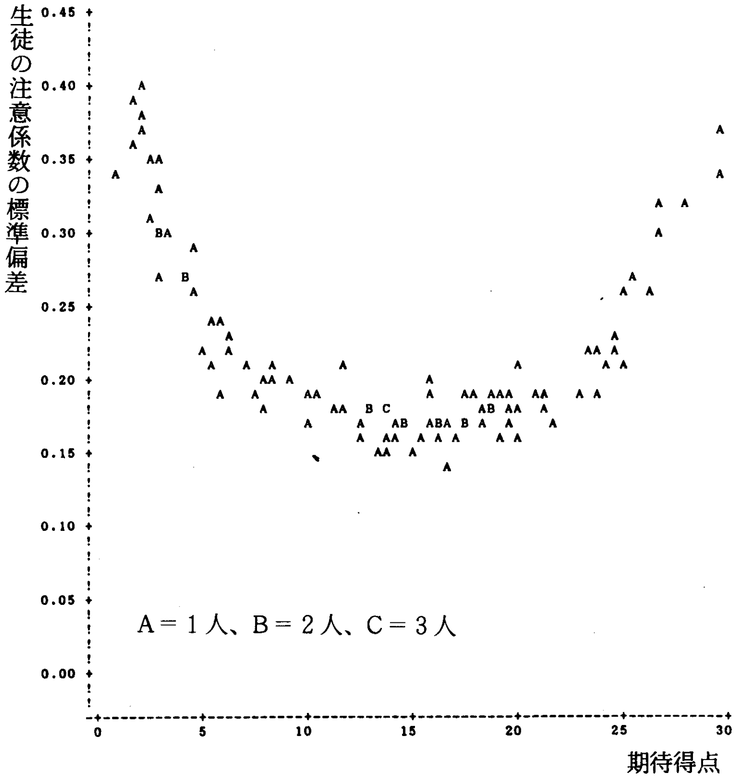

図6　生徒の注意係数の標準偏差と期待得点の散布図
（シミュレーションによるもの）

図7は、項目のほうの注意係数の標準偏差と期待正答者数との関係を示したものである。項目の困難度の分布を能力の分布に比べて狭く設定したため、図6ほどはっきりした関係は見られないが、先の数値例（図5）で見られた関係は、ここでも同様に認められる。

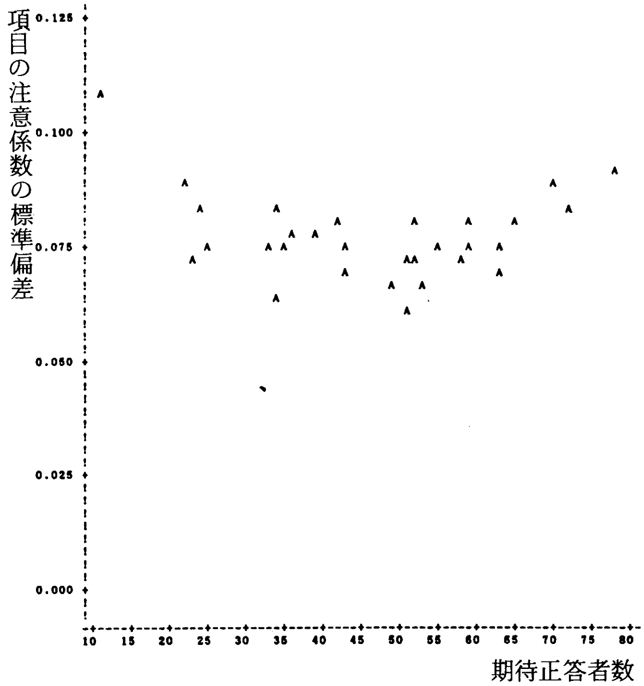

図7　項目の注意係数の標準偏差と期待正答者数の散布図
（シミュレーションによるもの）

次に、生徒数と項目数だけを変えて、同様のシミュレーションを行うことによって、注意係数の変動の大きさが生徒数や項目数によってどのように変わってくるかを調べてみた。生徒数を50および100とし、項目数を15および30としたときの、生徒の注意係数の標準偏差の平均を表3に、項目の注意係数の標準偏差の平均を表4に示す。ここでは、生徒数も項目数も2通りずつしか設定しておらず、シミュレーションの条件もかなり限定されたものではあるが、生徒の注意係数の変動は項目が少ないときほど大きく、項目の注意係数の変動は生徒が少ないときほど大きくなることは、一般的な結論として述べてよいだろう。

表3　生徒の注意係数の標準偏差の平均

生徒数	項目数	
	15	30
50	0.308	0.208
100	0.313	0.220

表4　項目の注意係数の標準偏差の平均

生徒数	項目数	
	15	30
50	0.101	0.126
100	0.075	0.077

結論と問題点

項目への正答確率が項目の困難度および生徒の能力と完全に対応しているという条件のもとでの注意係数の変動の大きさは、生徒の能力レベルによって、また項目の困難度によって、さらには生徒数や項目数によって組織的に変化することがわかった。したがって、データから求められた注意係数によって、そのような理想的な条件が満たされているか否かを判定するには、注意係数の変動の大きさを考慮に入れた方法が望まれる。本論文で提案したのは、まさにそのような方法である。

この方法を実用化する上で問題となるのは、データを発生させて注意係数を計算するという手続きを多数回繰り返すのに要する時間である。数値例(1)で示した49人×30項目のデータでは、パーソナル・コンピュータＰＣ—9801 ＶＸを用いて100回の繰り返し計算を行ったところ、計算時間が1時間を超えた。将来パーソナル・コンピュータの演算速度はさらに速くなると期待されるものの、現時点においては、この方法の実行に時間がかかりすぎるということは否定できない。したがって、本論文で提案した方法は、すぐに実用化できるものというよりも、注意係数の統計的性質に目を向けることによってより適切な解釈を可能にするための1つの方向づけを与えるものと考えるべきだろう。

文　献

佐藤隆博　1975　Ｓ－Ｐ表の作成と解釈　明治図書

Swaminathan, H. 1983 Parameter estimation in item response models. In R. K. Hambleton (Ed.), *Applications of item response theory.* Vancouver : Educational Research Institute of British Columbia.

Wright, B. D. 1977 Solving measurement problems with the Rasch model. *Journal of Educational Measurement, 14,* 97-116.