

入試データについての一考察

金 田 峯*
今 田 孝 保**

A Study on the Data of the Entrance
Examination of Tomakomai Technical College

Takashi KANETA

Takayasu IMADA

要 旨

本校の入試データに対して、多変量解析の一手法である主成分分析法による処理を行ない、地域別、年度別の観点から考察を試みた。

Synopsis

This article deals with the data of the entrance examination of our college, using the method of Principal Component Analysis, which is one of the methods in Multivariate Analysis, and also tries to give a consideration to the result of the examination by regional and annual groups.

1 ま え が き

本校開設以来 13 年を経過したわけであるが、全学挙げての最大の年中行事は入学者選抜試験であり、その都度志願者の状況や試験成績に多くの関心が寄せられている。年度毎に志願者数が異なり試験問題の質や難易の程度が異なる中で、年度毎の入学者の質的差異を入試成績のみをもって知ることはむずかしい。志願者数の多少や辞退者数の多少、更には学校差などの問題点を含んでいるとはいえ、中学校からの内申書による成績の変化等をみることによっていくらかは把握できる。実際的には日常学生に接している教授者自身が、それらに最も敏感であり、授業時の反応や各種のテスト成績、経験を通じて肌で感じているところであろう。

近年全国的な傾向として高専入学者の質的低下が叫ばれている。大学志向の社会的背景、それに伴う中学校の進路指導の徹底、高専制度の認識不

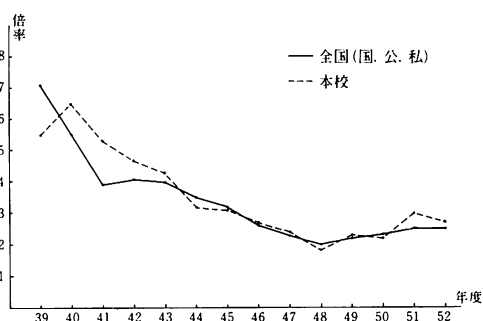
足、義務教育での教育課程の問題等々が微妙に絡み合っただけの現象と考えられるが、この傾向は本校も例外ではないと思われる。日常の教育活動が効果的に行なわれるためには、受け入れた学生について、時宜に応じてその実態や質的变化を的確に捕え、出来るだけ速やかに対応策を講じることが必要であろう。この辺の傾向把握に寄与できそうな資料を提供出来ないかと考え着手したのが、本稿での調査分析の動機である。共通項目として不足のあった初年度を除く、13 年間にわたる入試データについて、多変量解析の一手法である主成分分析法による分析を行ない、全受験者については地域別の年度による変化を、入学者については全受験者集団における相対的位置とその年度による変化を示し、考察を加えるものである。使用した分析手法については、その数理の概要や取扱に関して、本稿筆者の一人である金田⁽¹⁾が述べているので、それに従うことにする。

2 受験者、入学者の概観

入学志願者は学区制度がないこともあって地元を中心に全道全域に渡っている。初年度からの志

* 助教授 一般教科数学

** 助教授 電気工学科



図－1 39年度以降志願者倍率

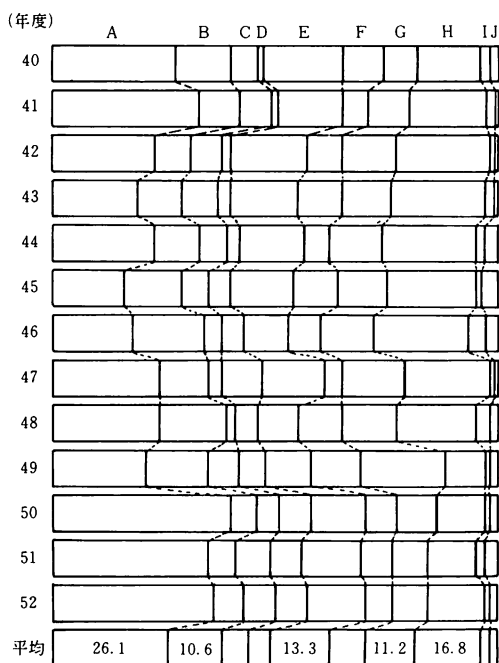
願状況は図1⁽²⁾に見られる通りであり全国高専の平均的傾向とはほぼ一致し、下降安定の様子がうかがえる。高等学校や大学への志願の実態をみると、倍率の高低が入学者の学力の高低に対応するとは限らないが、各地の高専の現場からの声や本校での実感から残念ながら或る程度は容認しなければならない。昭和40年度から昭和52年度までの入試データは入試5科目の成績(100点法)と中学校の調査書による第3学年の9科目の評定の和の2倍(5段階評定による)を内申成績とし、それらはすべて標準化したものを基礎データとした。各特性は次の通りである。

x_1 : 国語, x_2 : 社会, x_3 : 数学
 x_4 : 理科, x_5 : 英語, x_6 : 内申

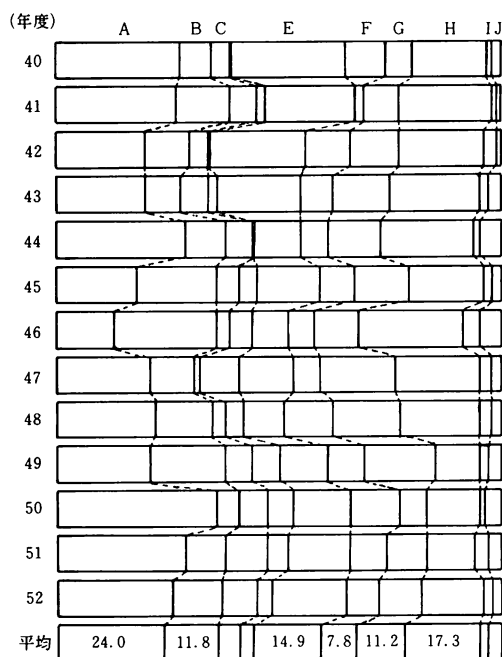
中学校の内申成績についての学校格差の問題について、芳谷⁽³⁾、山本・井関^{(4),(5)}の調査研究があるが、ここでは加工修正は行わずそのままとした。受験者の出身地は次のような地域に分けてある。

- A: 中規模の市1
- B: 大規模の市1
- C: 中規模の市2
- D: 大規模の市2
- E: 小市町村部1
- F: 小市町村部2
- G: 小市町村部3
- H: 小市町村部4
- I: 小市町村部5
- J: その他の市町村部

この10の分類は便宜的なもので、確たる根拠によって分けられたものではないが、地元を中心に据えてほぼ放射状に分けたものである。これによる年度毎の受験状況は図2、入学者の状況は図3である。地域を具体的に明示すれば、受験状況、入学状況の現実的一面の把握が可能と思われる



図－2 地域別受験者の割合



図－3 地域別入学者の割合

が、ここでは触れない。更にはA, B, E, G, Hの5地域が受験者、入学者ともに平均(13年間の)10%を越えていてこれらの地域だけでどちらもほぼ80%をしめている。後述の地域別検討では

表－1 受験者の年度別固有値

λ \ 年度	40	41	42	43	44	45	46	47	48	49	50	51	52	平均
λ_1	3.253	2.904	3.066	2.948	3.114	2.827	2.571	2.659	2.705	2.976	2.791	3.015	3.033	2.915
λ_2	0.714	0.851	0.792	0.763	0.725	0.851	0.819	0.872	0.853	0.783	0.832	0.810	0.730	0.800

表－2 累積寄与率

λ \ 年度	40	41	42	43	44	45	46	47	48	49	50	51	52	平均
λ_1	54.22	48.40	51.11	49.14	51.90	47.11	42.85	44.31	45.08	49.59	46.52	50.25	50.54	48.54
λ_2	66.12	62.57	64.30	61.85	63.99	61.30	56.51	58.84	59.30	62.65	60.39	63.74	62.71	61.87

表－3 Z_1 の固有ベクトル

ℓ \ 年度	40	41	42	43	44	45	46	47	48	49	50	51	52	平均
ℓ_1	0.373	0.320	0.389	0.383	0.406	0.357	0.356	0.374	0.424	0.368	0.406	0.389	0.372	0.378
ℓ_2	0.389	0.429	0.419	0.399	0.436	0.424	0.424	0.405	0.380	0.418	0.427	0.409	0.389	0.411
ℓ_3	0.416	0.376	0.386	0.408	0.375	0.359	0.429	0.393	0.371	0.411	0.406	0.364	0.412	0.393
ℓ_4	0.398	0.433	0.412	0.427	0.408	0.457	0.406	0.435	0.393	0.386	0.381	0.433	0.423	0.415
ℓ_5	0.434	0.451	0.407	0.422	0.428	0.432	0.426	0.418	0.442	0.459	0.402	0.425	0.407	0.427
ℓ_6	0.435	0.426	0.434	0.408	0.393	0.410	0.404	0.421	0.434	0.401	0.427	0.425	0.442	0.420

表－4 Z_2 の固有ベクトル

ℓ \ 年度	40	41	42	43	44	45	46	47	48	49	50	51	52	平均
ℓ_1	-0.509	0.785	-0.468	0.641	-0.461	-0.391	0.782	-0.335	-0.168	0.719	-0.238	-0.514	0.717	0.518
ℓ_2	-0.365	0.069	-0.143	-0.496	-0.204	-0.271	0.205	-0.431	-0.630	0.252	0.212	-0.255	-0.043	0.275
ℓ_3	0.531	-0.589	0.610	0.125	0.711	0.769	-0.273	0.576	0.724	-0.424	0.272	0.762	-0.363	0.518
ℓ_4	0.505	-0.172	0.444	-0.529	0.378	0.196	0.099	0.487	0.135	-0.455	0.665	0.155	-0.385	0.354
ℓ_5	0.053	-0.014	-0.437	0.159	-0.054	0.092	0.256	-0.364	0.114	0.093	-0.491	0.127	0.393	0.204
ℓ_6	-0.261	0.050	0.004	0.146	-0.309	-0.368	-0.444	0.032	-0.142	-0.157	-0.376	-0.222	-0.221	0.210

表－5 Z_1 の因子負荷量 ($\sqrt{\lambda_1}\ell_1$)

x \ 年度	40	41	42	43	44	45	46	47	48	49	50	51	52	平均
x_1	0.672	0.545	0.681	0.658	0.716	0.601	0.571	0.611	0.697	0.634	0.678	0.675	0.647	0.645
x_2	0.702	0.731	0.734	0.685	0.770	0.713	0.680	0.660	0.625	0.722	0.713	0.711	0.678	0.702
x_3	0.750	0.640	0.676	0.701	0.661	0.604	0.687	0.642	0.611	0.709	0.678	0.632	0.717	0.670
x_4	0.719	0.738	0.721	0.734	0.720	0.768	0.651	0.709	0.646	0.665	0.636	0.752	0.737	0.708
x_5	0.783	0.769	0.713	0.725	0.756	0.727	0.683	0.682	0.727	0.792	0.672	0.738	0.709	0.729
x_6	0.785	0.726	0.760	0.701	0.694	0.690	0.648	0.687	0.714	0.691	0.713	0.738	0.770	0.717

表－6 Z_2 の因子負荷量 ($\sqrt{\lambda_2}\ell_2$)

x \ 年度	40	41	42	43	44	45	46	47	48	49	50	51	52	平均
x_1	-0.430	0.724	-0.417	0.560	-0.392	-0.361	0.708	-0.313	-0.155	0.636	-0.217	-0.462	0.613	0.461
x_2	-0.308	0.064	-0.127	-0.433	-0.174	-0.250	0.186	-0.402	-0.582	0.223	0.193	-0.230	-0.037	0.247
x_3	0.449	-0.543	0.543	0.109	0.605	0.710	-0.248	0.538	0.668	-0.375	0.248	0.686	-0.310	0.464
x_4	0.427	-0.159	-0.395	0.462	0.322	0.181	0.090	0.455	0.124	-0.403	0.607	0.140	-0.329	0.315
x_5	0.045	-0.013	-0.388	0.139	-0.046	0.084	-0.232	-0.339	0.105	0.082	-0.448	0.114	0.336	0.182
x_6	-0.220	0.046	0.003	0.128	-0.263	-0.340	-0.402	0.029	-0.131	-0.139	-0.343	-0.200	-0.189	0.187

この5地域をとりあげる。

3 主成分分析の結果

前述の特性について、各年度毎別に受験者全員

を対象に主成分分析した結果が表1から表6までである。

ここに Z 、 λ 、 ℓ 、 $\sqrt{\lambda}\ell$ 等について略述する。

Z は次式で表わされる特性値である。

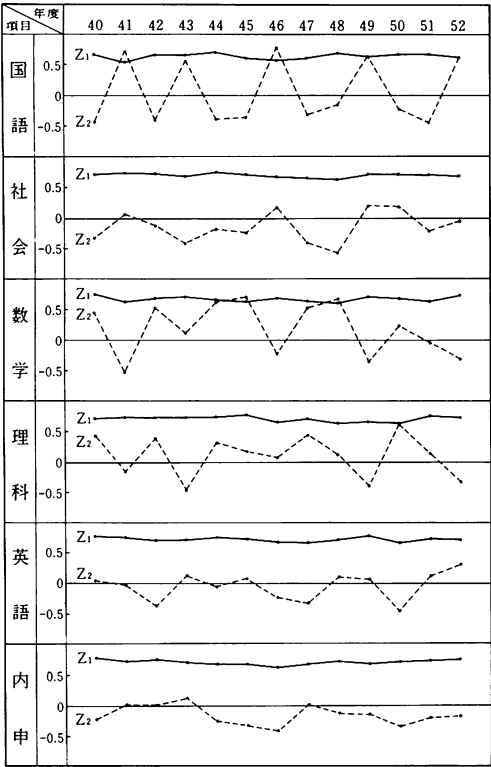


図-4 項目別にみた因子負荷量の変化

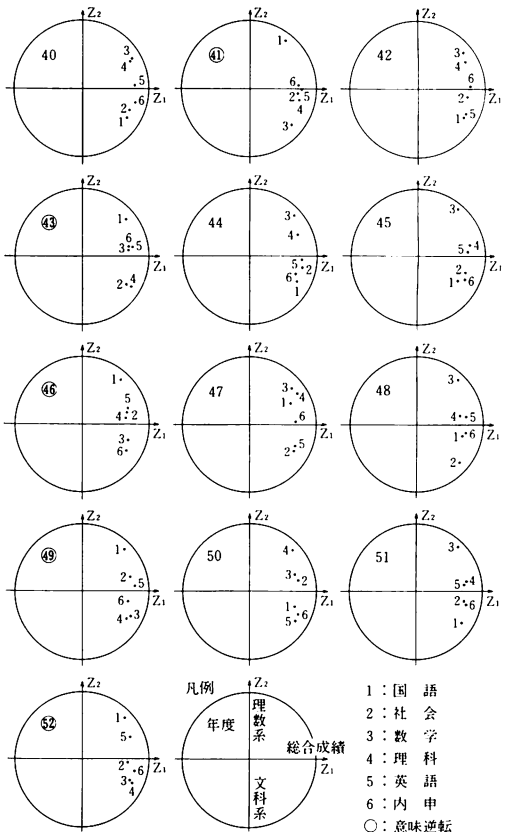


図-5 各科目の因子負荷量

$$Z = l_1x_1 + l_2x_2 + \dots + l_6x_6 \quad (*)$$

(ただし $l_1^2 + l_2^2 + \dots + l_6^2 = 1$)

Z_1 が第1主成分で多変量データの縮約化による総合特性値で最も多くの情報を保有する特性値である。データからの相関行列を $\mathbf{R}$ とすると、 Z の分散最大化に伴う $\mathbf{R}$ の固有方程式の最大固有値 λ_1 が Z_1 の分散であり、 λ_1 に対応する固有ベクトルの要素が重み係数 l_i ($i=1, 2, \dots, 6$)である。
 $\sqrt{\lambda_1} l_{1i}$ は Z_1 と x_i の相関係数で因子負荷量である。 Z_2 が第2主成分で Z_1 と無相関で $\mathbf{R}$ の固有方程式の2番目に大きい固有値 λ_2 が Z_2 の分散であり、 l_{2i} が λ_2 に対応する固有ベクトルの要素である。この場合以下同様にして Z_6 まで求められる。また Z_1 の寄与率が $\lambda_1/6$ であり、第 k 主成分までの累積寄与率は $(\lambda_1 + \lambda_2 + \dots + \lambda_k)/6$ で情報の保有率を表わしている。(詳細は(1), (6), (7)等)
入試成績は試験科目別、年度別に平均値や分散が異なっている筈である。ここで同一の手法で分析した結果を概観してみることにする。

表1, 2からは6次元空間における成績分布の

集中の度合いにいくつかの差異が認められるが平均的にみて第1主成分での情報保有の割合はほぼ50%以上, 第2主成分まででのそれは60%以上である。縮約された資料としてこの程度情報が保有されていれば、この新しい特性値で分布構造の概略がみれると考える。
表3からはどの年度も各変量の係数がほぼ等しい重みをもち0.4位の大きさであるのがみられる。
変量が標準化されていることから Z_1 は単純合計された総合値と大小はほぼ一致する。
表4からは符号を無視すると国語, 数学の重みが大ききようである。符号のつき方は年度によって異なるが逆転してみるとほぼ同じ形態をしているようである。 Z_1 の値が等しいものの中で Z_2 の値を比較すると国語の成績の良いものは Z_2 の値は小さく、逆に数学の成績の良いものは Z_2 の値が大きいという性質がみられる。 Z_2 は国語, 数学の成績のよさに反応する特性値のようである。
表5, 6は各変量と総合特性値 Z_1, Z_2 との相関

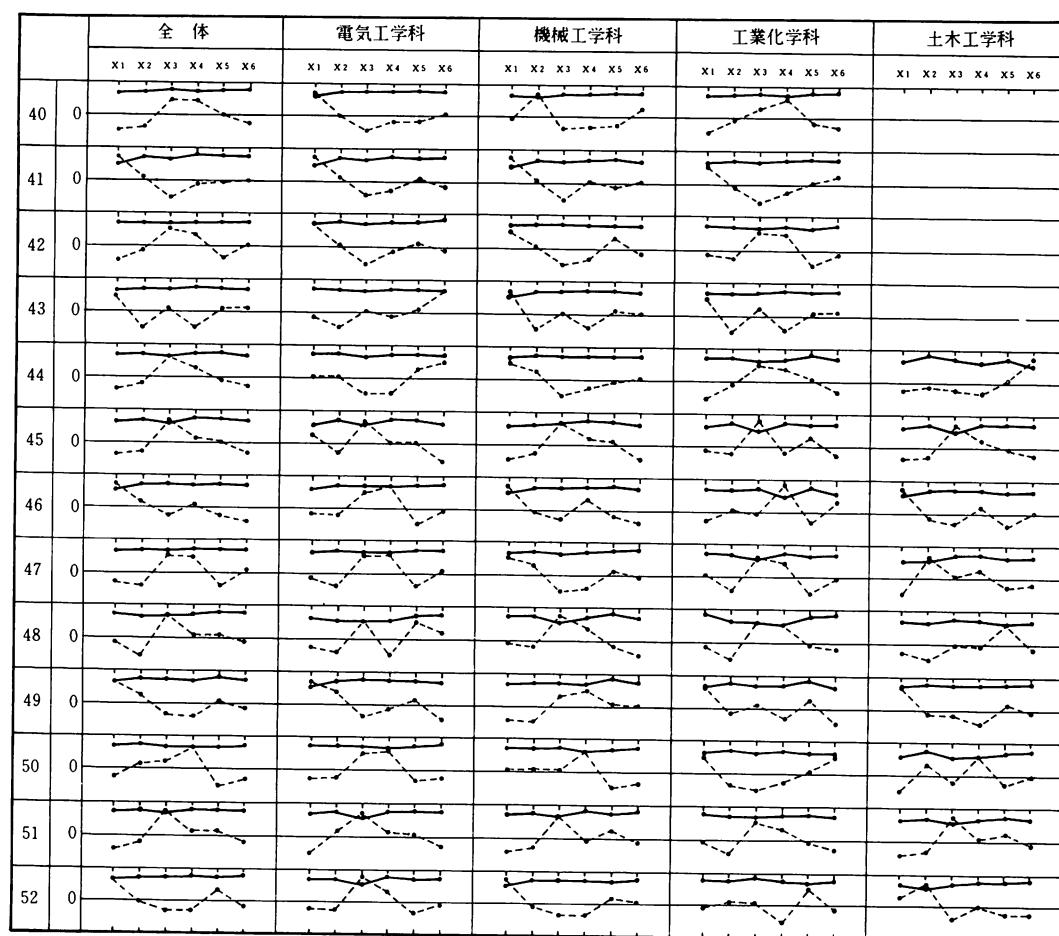


図-6 学科別に分析したときの因子負荷量

係数で表3, 4より寄与の様子がわかりやすい。各年度とも数値にいくらかの差がみられるが年度平均でみると各変量とも0.7前後である。これらを変量別に年度変化を示したのが図4である。第1主成分への各変量の寄与はその数量もほぼ一定で差が小さい様子がみれる。第2主成分について英語、内申が変動が小さく国語、数学の変動が大きい。

更に Z_1 , Z_2 の組にして因子空間上にプロットしたものが図5である。これによると Z_2 の方向での科目の分離がよくわかり、41, 43, 46, 49, 52の年度において Z_2 軸の向きを逆転させると他の年度と科目のもつ位置の方向がほぼ一致する。 Z_2 軸方向において文科系科目と理数系科目が分離されている様子が理解されるところである。ここで各年度とも学科別にそれぞれ主成分分析した結果を図6に示す。

各学科とも年度別は勿論のこと、同年度でも成績分布が異なる筈であるが、分析結果から、第1主成分、第2主成分共に先述と同じ傾向がみられる。金田⁽¹⁾の述べている事柄への実証例が多数追加されたことになる。

4 考 察

3による保証の上に立って具体的検討を加えよう。

受験者個々の第1主成分、第2主成分による評点は、3の(*)の形で決定された Z_1 , Z_2 の式に各人のもつ科目成績、内申成績を代入することによって得られる。 Z_1 の評点は総合成績の順位を決定し、 Z_2 の評点は文科系科目、理数系科目の成績の良さを表わしている。

Z_1 , Z_2 を座標軸とする平面上に評点の組を点(Z_1 , Z_2)としてプロットすると分布の状態が視

表-7 Z_1^* の地域別平均(13年間)

地区	A	B	C	D	E	F	G	H	I	J
受験者	0.017	-0.010	-0.221	-0.114	0.135	0.014	0.008	0.019	-0.004	0.043
入学者	0.693	0.691	0.588	0.487	0.798	0.835	0.642	0.718	0.832	0.589

表-8 入学者の年度別平均評点の Z^*

年度 Z^*	40	41	42	43	44	45	46	47	48	49	50	51	52	平均 (偏差)
Z_1^*	1.038	0.818	1.184	1.066	0.890	0.829	0.737	0.506	0.311	0.426	0.442	0.475	0.649	0.721
Z_2^*	0.217	-0.010	-0.051	-0.087	0.142	0.202	0.015	0.015	0.049	-0.055	0.095	-0.058	0.034	0.079

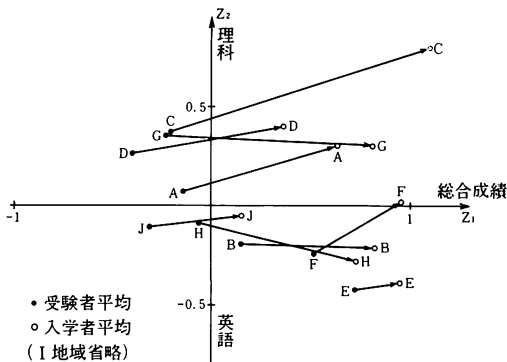
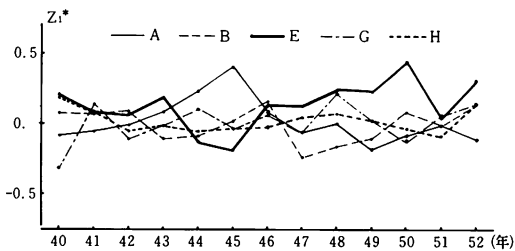
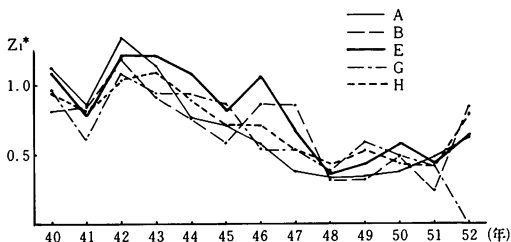
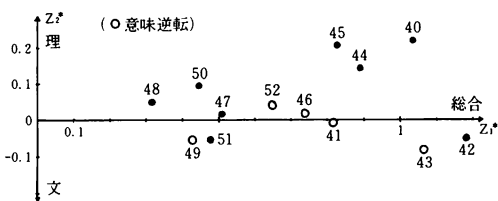


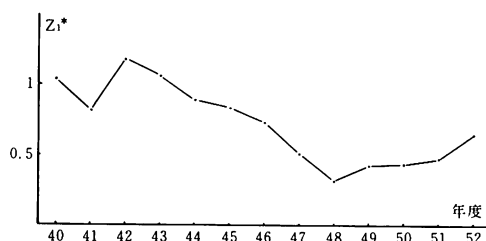
図-7 地域別受験者, 入学者の平均評点(50年度)

図-8 受験者の地域別 Z_1^* 図-9 入学者の地域別 Z_1^* 図-10 入学者平均評点による (Z_1^*, Z_2^*)

覚化される。50年度について地域別に受験者と入学者の平均評点をプロットしたものが図7である。 Z_1 軸方向で原点(平均)から右に離れた地域程平均成績がよくその順位がはつきりでている。 Z_2 については正方向(上)に原点から離れる程理科の成績が、負方向(下)に原点から離れる程英語の成績が良い事を意味する。同一平面で入学者の平均もプロットすると受験者と入学者の平均についての変位がみられ、同地域の点を結ぶ矢線がその方向と大きさを示している。年度によって固有値、固有ベクトルが異なるのでこの形で年度比較は出来ない。 Z_1, Z_2 とも平均が0で分散が λ_1, λ_2 であるから各評点について $Z/\sqrt{\lambda}$ を求め Z の標準化を行なうと、各年度毎の評点を同一の尺度で表わすことになり相対的の大きさに変えることが出来る。これを Z^* とおき各年度毎に地域別平均評点の Z^* 、入学者の平均評点の Z^* を求め年度による変化を調べてみる。この際、2で概観したことより、入学者の割合が13年間平均して10%以上であるA, B, E, G, Hの5地域に注目してみた。地域別 Z^* の年度による平均評点は表7に示した。これによると昭和52年までの13年間の平均でE地域の受験者、入学者がともに成績が一番良いようである。入学者の地域別では、次いでH地域がよく、A, B, G, の地域はほぼ等しい。更にこれらについて今一步詳しく見たのが図8, 図9で年度内の比較は縦に、年度別変化は横にみればよい。総じてE地域が受験者、入学者とも上位の変化をしているのがみられる。

次に入学者全体について年度別平均評点の Z^* を計算したものが表8である。各年度の (Z_1^*, Z_2^*) を図10に示す。 Z_1 軸の正方向での値が大きければ、その年度の受験者中での入学者の平均値の位置が高いことを示している。

○印を附した年度は Z_2 での意味が逆転している。 Z_1^* 軸に関して対称に移動してみれば、他の年

図-11 入学者平均評点の Z_1^*

度での Z_2^* の比較ができる。40、44、45 年度は特に理数系科目の良いものが多かったと読みとれる。

図 11 は、入学者の平均評点の Z_1^* に注目して、年度による推移をみたものである。下降減少を辿り 48 年度を境に再び上昇の傾向にある。この数値の大小は、各年度の全受験者における入学者平均点の相対的大きさを示すから、各集団内のどの部分が入学しているかを測る数値である。大きい程、その年度の全体の中で試験成績の良いものが多いとも考えられる。

これにより各年度の成績そのものを比較することは出来ないが、各年度とも等質で辞退者数も一定と仮定すれば、下降は即成績低下を意味する。年度毎の受験者数や辞退者数の大小が主にこの数値に影響を与えると思われる。51 年度から 52 年度にかけての上昇傾向には、更に判定規準の変更による影響も考えられるが、来年度以降の変化に注目する必要がある。この Z_1^* の値は必ずしも成績の良さを示さないことも先に述べたが、強いて言えば成績の意味での歩止りの程度を示していると考えてよさそうである。その意味で入学者の質的低下の問題についてのある一面については語るができそうであるが、決定的な材料は提供出来ない。もっと別な面での調査、検討が必要である。ここにこの種の問題に関連する一つの資料を図 12 で示そう。

これは本校において比較的一定の質を維持しているとみられるある学科の入学者について、初年度からの中学校の内申成績を国語、社会、数学、理科、英語の 5 科目について、最高評定を持つものの割合である。入学者についての中学校評定には学校格差がそれ程大きくないものとして取扱った。これによると 48 年度を境に明らかな変化がみられ、最近の実感を裏付ける資料の一つと思われる。

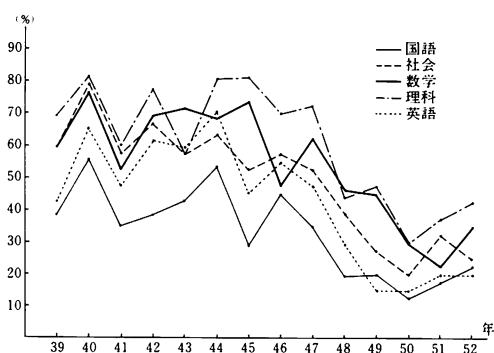


図-12 中学校内申成績の変化例

5 おわりに

以上、13 年間分の入試データに主成分分析法を適用し、分析結果を受験者と入学者について地域別、年度別の観点から述べ、入学者の最近における質的变化の一面についても触れて来た。得られた数値のもつ意味を考えることに非常に難しさを感じるが、目立った点に注目して一応次のようにまとめられると考える。

- (1) 入学者が 10% 以上の 5 地域の中で、平均的にみて E 地域が受験者、入学者共に成績が良い。
- (2) 同一年度の入試成績における入学者平均成績の相対的变化は 48 年を境に下降から緩やかな上昇の傾向を示す。
- (3) (2) における変化は入学者の質的变化を示すとは限らないが、中学校の内申成績でみると 48 年以降かなり低下しているのがみられる。

本稿で示した分析のプロセスや結果に現われた数値は、入試データのもつ性質の一面や日常の体験から語られている学生の質的な一面をその変化を以って語っているものと理解したい。

その意味で受験者や入学者に関する傾向把握のための一資料を提供できたものと思う。

最後に、本分析を着手するに当り、この種の調査の示唆と励ましを与えられた本校前教務主事の 小野寺隆教授、資料の取扱いに当って常に慎重な配慮と助力を示された本校教務係の菅原修三事務官、遠藤豊事務官、膨大な量のパンチ作業に協力された電子計算機室の山内幸子さんに対して厚く感謝の意を表するものである。

参 考 文 献

- 〔1〕金田 峯：「入試成績への主成分分析法の適用について」苫小牧高専紀要 13 号, 1978
- 〔2〕文部省大学局技術教育課：「高等専門学校入学者等状況調査」1977
- 〔3〕芳谷大和：「入学者選抜試験における調査書の評価について」福井高専紀要, 自然科学 2 号, 1969
- 〔4〕山本 孟・井関正嗣：「電気工学科を中心とした各種成績相互の関連について」舞鶴高専紀要 8 号, 1973
- 〔5〕山本 孟・井関正嗣：「電気工学科を中心とした各種成績相互間の関連について」(第 2 報) 舞鶴高専紀要 9 号, 1974
- 〔6〕河口至商：「多変量解析入門」森北出版 (1973)
- 〔7〕奥野忠一他：「多変量解析法」日科技連 (1971)
- (昭和 52 年 12 月 3 日受理)