

平成16年度大学入試センター試験前期日程における 志願状況と科目別得点の特徴について

大津 起夫*

橋本 貴充*

荘島宏二郎*

石塚 智一*

要 約

平成16年度大学入試センター試験受験者の科目選択および科目得点と、前期日程における志願状況および合否状況との関係を調べた。受験者数の多い「国語Ⅰ・Ⅱ」,「数学Ⅰ・A」,「数学Ⅱ・B」,および「英語」について、募集単位別に入学者の得点平均と不合格者の得点平均とを図示した。科目により入学者と不合格者の得点平均の差が大きい部分の異なることが確認された。さらに前期日程における募集単位を、上記4科目の入学者の科目別得点平均および不合格者の科目別得点平均によって特徴付けし、10個のクラスターに分類した。これらのクラスターごとに、前期日程における合否状況と科目選択状況を調べた。入学率はクラスターによらず多くの場合類似した値であるが、志願者の性別比率は違いが大きい。また、理科の科目選択において性別による顕著な違いが見られた。さらに、複数の科目得点の関係について同時分布を表示するとともに、クラスター別の平均得点と非線形因子分析による推定結果について検討を加えた。

キーワード：科目選択, 可換マトリックスグラフ, 非線形因子分析

1 はじめに

現在の日本の国公立大学への入学は、大学入試センター試験（以下ではセンター試験と表記する）と個別試験の成績の両者を用いて選抜が行われている。センター試験は、高校での学力の達成度を測る目的の下に作題が行われており、多くの国公立大学における入学試験において、選抜のための

主要な資料として利用されている。また、現在では私立大学における相当数の入学定員の選抜にも用いられている。

ここでは、センター試験の科目別得点と国公立大学における合否判定との関連の全体像を把握することを目的とする。

センター試験を選抜のための資料としている大学（より正確には学部、履修コースなどの募集単位）には様々の特徴を持つものがあり、合否判定

* 大学入試センター 研究開発部 試験作成支援研究部門

2005年11月15日 受理

† 本研究は、平成17年度大学入試センター共同研究Ⅰ「試験問題統計情報の整備と利用に関する研究」（研究代表者大津起夫）、平成16年度大学入試センター所長裁量経費「試験作成支援研究のための情報共有基盤サーバの構築」（研究代表者大津起夫）、および平成16-17年度科学研究費補助金（基盤研究（C））「大規模テストのための統計情報管理システムの開発」（課題番号16530441、研究代表者大津起夫）による支援を受けた。

がなされる基準も多様である。以下では、受験者が出願している募集単位と、受験科目および科目別の得点の関係に注目する。これらは出題科目の構成およびそれぞれの科目の試験の難易度の設定などを検討するために重要であり、センター試験の役割を検討する上で必須の情報であるが、最近ではこれらを概観しうる分析が見当たらない。このために、平成16年度の前期日程における受験者の志願と合否判定の状況を、研究開発部が保持する受験者の個人情報秘匿したデータに基づき検討する。分析対象は1年度の前期日程に限ったものであるが、ここで有効な分析の枠組みを明らかにすることができれば、他の年度や後期日程などについても、同様の分析を行うことにより有益な情報を得ることが期待できる。

センター入試の前身である共通第1次試験については、本研究と同様の関心に基づく分析がかなり詳細に行われている。研究の多くは共通1次試験の発足間もないころに行われており、大学入試センターの研究紀要およびその非公開版に出版されている。荒井(1982)は、共通第1次学力試験の総得点を説明変数とし、募集単位における合否を被説明変数としたロジスティック回帰を行い、そこで得られる回帰係数を「相関度」と命名し、それらの大きさを学部種別に分析している。また、清水(1985)は科目別の得点を説明変数とし、募集単位別に合否判定を予測する線形判別分析を行い、そこで得られた科目別得点の係数の特徴を検討している。

これらはいずれも、共通第1次試験の得点が大学の合否にどの程度影響しているかを検討するものである。この分析の視点は本研究にも共通しているが、ここではより単純なデータ記述を用いることにより状況を概観する。以降に示す得点と合否の関係の分析では比較的単純な分析手法しか用いていないが、分析の視点と結果の図的な表示方法を工夫することにより、直観的に明快な結果を得ることは可能である。

共通第1次試験およびセンター試験の実施において重要なもうひとつの視点は、選択科目間の難易度の比較である。いくつかの教科においては受験者による受験科目の選択が行われる。しかも少なからぬ募集単位において特定の教科科目の中からいくつかを背反的に選択して受験することが要

求されてきた。これらの募集単位の合否判定においては、互いに異なる受験科目を選択した受験者の得点が同一の基準において比較されることになるため、難易度の比較は公平性の確保からも重要な点である。現在ではこの問題に対応するため得点調整のルールが定められているが、これについては共通1次試験の発足当初より注意が払われてきた。

清水(1981)は不完備配置の線形モデルを用いて、昭和54および55年の共通第1次学力試験における社会と理科の選択科目間の難易度の比較を行っている。ただしここでの推定は通常の最小2乗法ではなく、残差絶対値の最小化に基づいている。また、清水(1983)では不完備配置の線形モデルと最小2乗法の枠組み(分散分析)を用い、科目の難易度および受験者群の能力推定を昭和54年から昭和58年までの5年間について行っている。この論文では科目間の難易度の差が年とともに次第に減少していることが報告されている。また、科目間の難易度比較ではないが、最近では吉村・荘島(2004)が傾向スコア(propensity score)を用いた難易度比較の方法を法科大学院適性試験のモニター調査を用いて示している。

1980年代前半においては線形モデルの推定を受験者全体について行うこと自体、かなり労力を要するものであったと推測されるが、現在ではより複雑な計算を低コストで実現できる。本論文では、この問題を大津(2004)によって開発された欠測値に対応する非線形因子分析を用いて取り扱う。この方法は計算コストはかかるものの、科目の難易度の比較に洗練された統一的分析のための枠組みを与える。

以下では、まず募集単位を志願者の得点における特徴によってタイプ(クラスター)別に分類し、次いでこれらの各クラスターにおいて、入学者および不合格となった志願者の得点における特徴を比較することにより、センター試験が大学入試選抜において果たしている役割を明らかにする。また、募集単位のクラスターにおける科目選択率、および志願者全体における科目間の難易度の比較についても併せて検討する。

ここで重視するのは、センター試験が大学入学の合否にどのように関係しているかの全体像を把握することである。以降で明らかになるように、個

別大学における成績と合否の関係においては、志願者がかなり均質な学力を持つ集団に制約されているため、センター試験の全体像を的確に把握することが難しい。ここでは、受験者全体を分析することにより、個別大学における受験者の分析では得難い知見を示す。

2 平成 16 年度センター試験の概要

平成 16 年度大学入試センター試験（平成 16 年 1 月実施）の志願者総数は 587,350 人であり、このうち 540,446 人（志願者のうち 92.0%）が少なくとも 1 科目以上を本試験あるいは追試験、再試験において受験した。これらのいずれか 1 科目以上を受験した志願者を「受験者」と定義する（試験に出席したものの科目選択の誤りにより有効な得点を得ていない者も含む）。志願者のうちいずれの科目も受験しなかった者は残りの 46,904 名（志願者のうち 8.0%）である。また、志願者のうち前期日程に出願したものは 274,171 名であり、313,179 名は前期日程での出願を行っていない。前期日程に出願を行っている志願者のほとんどは、当然ながらセンター試験のいずれかの科目を受験しているが、いずれの科目も受験せずに出願している者が 9 名存在する。これ以外の前期日程への出願者が、個別大学の出願要件を満たしているかどうかは必ずしも明らかではないが、受験者のうち前期日程に出願したものは 274,162 名であり、受験者における出願率は 50.73%となる。

試験は、国語、地理歴史、公民、数学①、数学②、理科①、理科②、理科③、外国語の 9 つの試験時間（実施順ではない）にわたって行われた。このうち、最大の受験者数があったものは外国語であり、537,057 人（受験者のうち 99.4%）が受験した。さらにこのうち、536,213 人（受験者のうち 99.2%）が英語を受験している。

3 前期日程における募集単位の特徴

まず、平成 16 年度の国公立大学の前期日程試験における志願状況をもとに、募集単位の特徴づけを試みる。前期日程で募集を行った募集単位は 536 あるが、これらを志願者のセンター試験における 4 つの科目「国語 I・II」、「数学 I・A」、「数

学 II・B」、および「英語」における得点によって特徴づける。ここでは各募集単位がこれらの科目を採点の対象としているか否かに係わらずに志願者の得点のみにもとづいて特徴づけを行う。これら 4 つの科目を用いるのは、536 の募集単位のいずれにおいても、合格し入学手続きを行ったものおよび不合格となったもののうち、これらの科目を受験しているものが少なくとも 1 名存在するためである。

以下ではセンター試験の受験者を前期日程において志願した募集単位別に類別し、さらに上記の各 4 科目について、その科目の本試験を受験して合格し入学手続きを行ったものと、不合格（2 段階選抜における不合格を除く）になったものを抽出する。また、以下では入学手続きを行った者を「入学者」と呼ぶ。厳密には手続きをしても入学するとは限らないが、簡明さを保つためこのように略称する。上記の科目について、各募集単位別に入学者の平均点と不合格者の本試験の平均点を求める。各募集単位について上記 4 つの科目の各々について入学者の平均点と不合格者の平均点が得られる。これら 8 個の変数が募集単位の特徴を表すものとみなして以下では分析を行う。

前述のように平成 16 年度のセンター試験において志願しながら、前期日程においていずれの国公立大学の募集単位にも出願を行っていないものが 313,179 名存在するが、これらの志願者の特徴については別途検討することにする。

まず、上記 4 科目の本試験における入学者および不合格者の募集単位別平均点を図 1 から図 7 に示す。当然ながら、いずれの場合も少数の例外を除いて入学者の科目別平均点が不合格者の科目別平均点より大きい。また、合格者数が 15 人に満たない募集単位は十字の印で識別した。各募集単位における志願者数の分布を検討すると、最小が 26、第 1 四分位が 260.75、中位数が 419、第 3 四分位が 653.5、最大が 2718 である。

図 1 は、「国語 I・II」の募集単位別の得点平均を示している。横軸は入学者の平均点、縦軸は不合格者の平均点である。入学者と不合格者の差は 15 点あたりを中心として広がっている。これらの得点差の分布を検討すると、半数以上が 10 点から 20 点の間にあることがわかる。成績が下位の部分で、入学者と不合格者の差が小さくなる傾向が

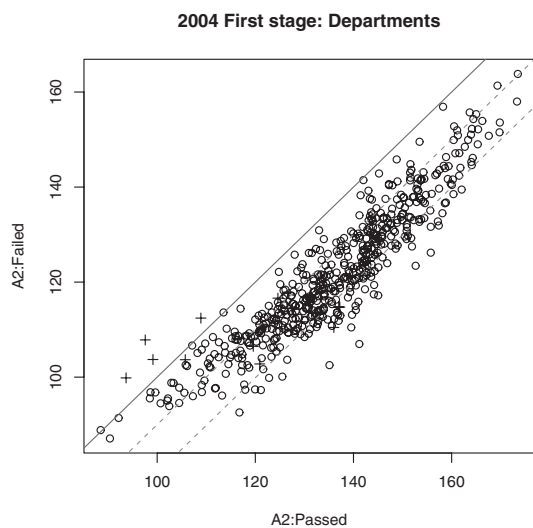


図 1 募集単位別の国語 I・II の平均点 (横軸：入学者, 縦軸：不合格者)

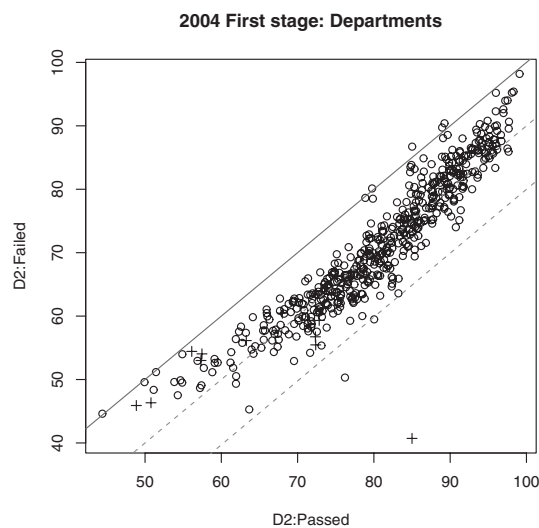


図 3 募集単位別の数学 I・A の平均点 (横軸：入学者, 縦軸：不合格者)

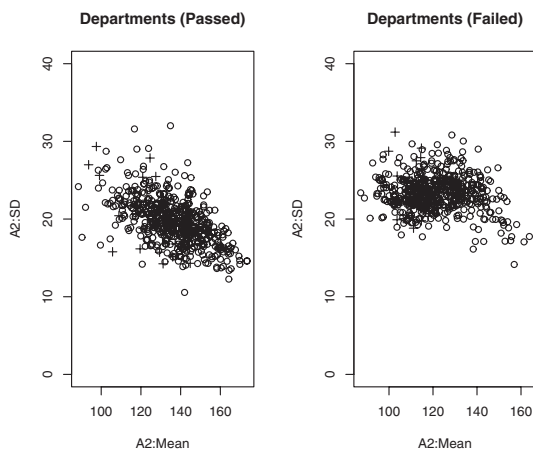


図 2 募集単位別合格否別の国語 I・II の平均と標準偏差 (左：入学者, 右：不合格者)

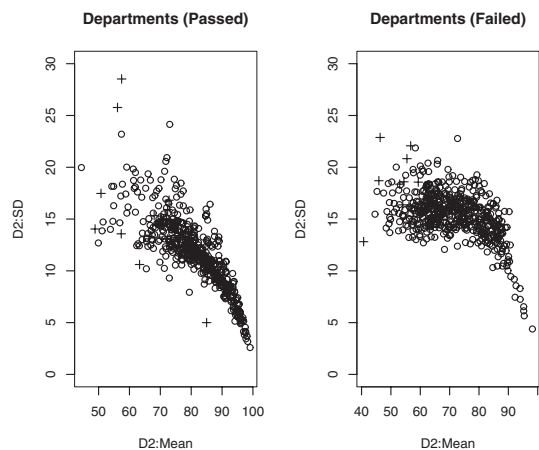


図 4 募集単位別合格否別の数学 I・A の平均と標準偏差 (左：入学者, 右：不合格者)

多少あるが、全体としてはほぼ一定の得点差がある。これに続く図 2 には、それぞれの募集単位における入学者と不合格者の別々に、縦軸に標準偏差 (SD)、横軸に平均 (Mean) を示した。合格者のグループは、それぞれの募集単位において得点の範囲が制限される。このため特に高得点の募集単位において入学者の標準偏差が小さくなる傾向が見られる。

図 3 は、「数学 I・A」についての入学者と不合格者の平均得点の図である。こちらでは、低得点の募集単位と高得点の募集単位の両端で、入学者の得点平均と不合格者の得点平均とが近づく傾向が見られる。入学者と不合格者の平均得点差は 10 点前後が多く、半数以上の募集単位が 5 点から 15

点の間に収まる。「数学 I・A」の場合には得点平均が満点に近い募集単位が少なくないため、これらにおいては標準偏差は極めて小さくなり、志願者の得点差が生じにくいことがわかる。

図 5 には「数学 II・B」の得点平均を示した。「数学 I・A」とは対照的に、高得点の募集単位で入学者と不合格者の平均得点差が開いている。また、図 6 では平均点 60 点から 80 点あたりで、募集単位別入学者の標準偏差が大きなもの全体から外れて存在する。入学者の平均得点が 60 点以上で、標準偏差が 15 より大きい条件には 44 の募集単位が該当するが、このうち 38 が医学部であり、2 つが歯学部である。少なからぬ医学・歯学系の学部において、入学者の「数学 II・B」の成績

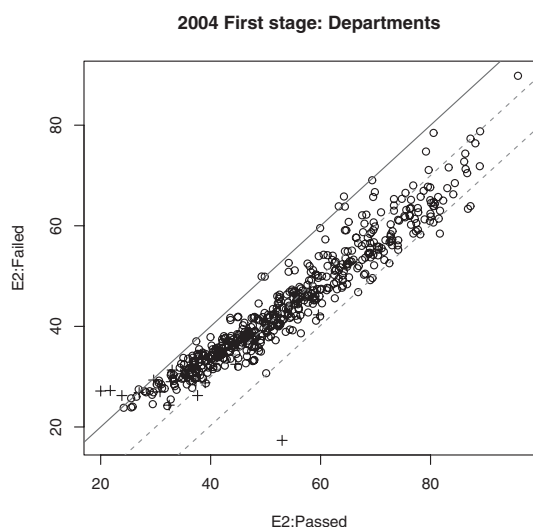


図 5 募集単位別の数学Ⅱ・Bの平均点（横軸：入学者、縦軸：不合格者）

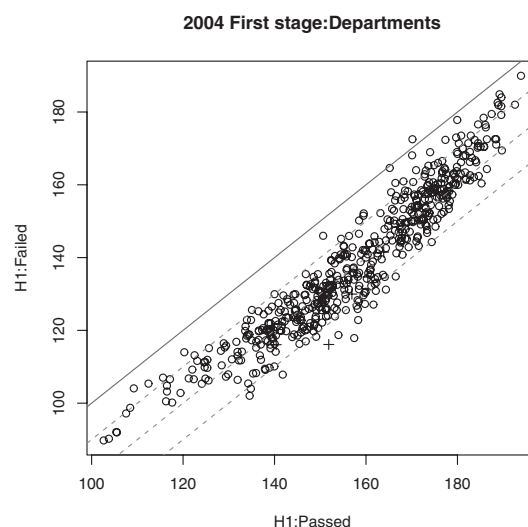


図 7 募集単位別の英語の平均点（横軸：入学者、縦軸：不合格者）

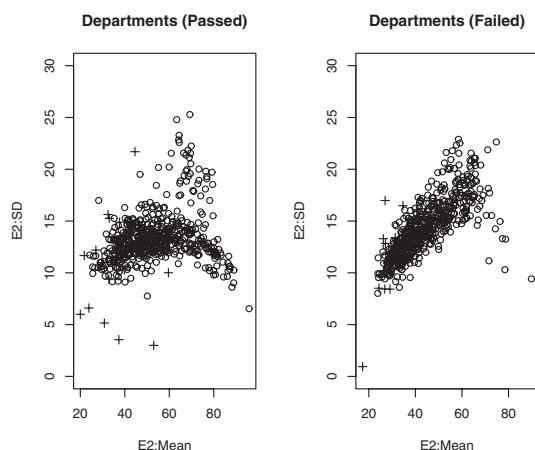


図 6 募集単位別合否別の数学Ⅱ・Bの平均と標準偏差（左：入学者、右：不合格者）

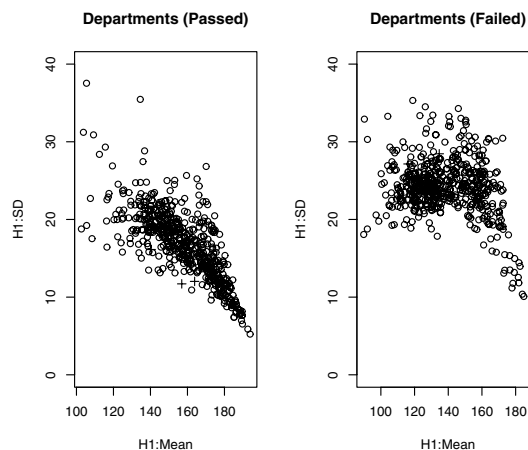


図 8 募集単位別合否別の英語の平均と標準偏差（左：入学者、右：不合格者）

には大きなばらつきのあることがわかる。入学者と不合格者の得点平均の違いはここでも 10 点前後が多く、募集単位の半数以上が 5 点から 15 点の間の得点差となっている。

「数学Ⅰ・A」と「数学Ⅱ・B」の図の違いは、前者が比較的易しく満点に近い受験者が多い一方、後者の試験問題がより難しく、すべての志願者が満点に近い得点をとるような募集単位が極めて少ないため、募集単位における「数学Ⅱ・B」の得点の分散がいくつかの募集単位では大きくなる場合があるためと思われる。

図 7 は「英語」の得点平均である。「数学Ⅱ・B」とは異なり、高得点の募集単位で得点差が小さくなる傾向が見られる。また合格者の標準偏差が特

異的に大きくなる募集単位も存在しない。合格者と不合格者の得点平均差は「国語Ⅰ・Ⅱ」よりも幾分大きく、20 点前後に分布しており、半数以上の募集単位が 15 点から 25 点の間にある。

4 募集単位のクラスタリング

ここでは、前節で示した 4 科目の 8 つの値を用いて募集単位のクラスタリングを行う。国語Ⅰ・Ⅱ、英語の試験は 200 点満点であり、一方、「数学Ⅰ・A」および「数学Ⅱ・B」の試験は各々 100 点満点であり尺度が異なるが、これらの標準化は行わず、4 科目の素点平均をそのまま分析に用いた。

センター試験の利用においては、素点に加重し

表 1 募集単位の分類

合格者と不合格者平均 によるクラスター	合格者平均によるクラスター										合計
	A'	B'	C'	D'	E'	F'	G'	H'	I'	J'	
A	30	0	9	0	0	0	0	0	0	0	39
B	0	36	26	7	0	0	0	0	0	0	69
C	0	0	34	0	0	14	0	0	0	0	48
D	0	0	2	72	2	3	3	0	0	0	82
E	0	12	0	9	21	0	0	0	0	0	42
F	0	0	0	0	0	40	17	5	0	0	62
G	0	0	0	0	21	0	18	0	9	0	48
H	0	0	0	0	0	0	7	66	0	5	78
I	0	0	0	0	0	0	8	4	20	3	35
J	0	0	0	0	0	0	0	2	0	31	33
合計	30	48	71	88	44	57	53	77	29	39	536

た得点が合否判定に利用されることが多い。ここで、数学と国語および英語を同じ尺度とすると数学の合計得点が 400 点となり、多くの利用大学における加重と異なり、記述の方法として妥当ではないと考えられる。加重しない得点によるクラスター分類が最適であるとの保証はないが、4 つの科目を同一の尺度に変換するよりは、素点のまま分析を行うほうがより望ましいと考えた。

クラスタリングの方法は k -means 法と呼ばれる手法を用い、R-2.1 システム (Ihaka & Gentleman, 1996; R Development Core Team, 2005) に実装されている関数を用いた。 k -means 法のアルゴリズムには複数のものが存在するが、ここでは Hartigan & Wong (1979) によるものを用いた。

この手法はあらかじめ指定した個数 K のクラスターに標本を分割する手法のひとつである。サンプルを表す添え字を $i = 1, \dots, n$ とし、第 i 番のサンプルの多変量データを $(x_{ij} | j = 1, \dots, p)$ とする。サンプルのクラスターへの所属を表す関数を $m(i)$ とする。関数 $m(i)$ はサンプル i が第 1 番目のクラスターに所属するなら 1、第 2 番目のクラスターに所属するなら 2、さらに一般的に k 番目のクラスターに所属するなら k となる関数である。

この $m(i)$ によって導出される第 k クラスターの平均を μ_k とする。ここで、次の値をクラスタリングの「良さ」を表す指標とする。

$$f = \sum_{i=1}^n \|x_i - \mu_{m(i)}\|^2 \quad (1)$$

この指標は、簡単な計算によって最適化すること

は困難であるため、経験的な反復計算を用いて、最適（と思われる）クラスタ分割法を求める。

対象を分類するために k -means 法を用いると、指標となる値 (1) の最小化をヒューリスティックに行うために、初期値の設定などによって結果が異なる場合がある。ここでは、求めるべきクラスター数を 10 個に固定し、各クラスターの中心の初期値をデータからランダムに選んだ。上記のアルゴリズムによる計算を 20 回繰り返し、指標が最適となるものを結果として採用した。

募集単位のクラスター数を 10 と設定したのは、合否別あるいは性別の分析をおこなうと、区分数は非志願群 Z を加え 22 となり、直観的な把握を行うためにはこの程度が限界であろうという判断に基づく。また、受験者の得点の構造は、これまでの分析を考慮すると、緩やかに関連する「国語、英語、社会科」および「数学、理科」の 2 つの次元で説明されることが期待されたため、それぞれを上位、中位、下位 3 群に分割することを想定すると 9 個の群に分類されることになる。分析のためのクラスター数は少なくともこれを上回るのが望ましい。

このような指標を用いたのは入学者の得点だけではなく、不合格になった志願者集団の特徴をも考慮する必要があると判断したためである。クラスターの名称は「英語」の科目得点の平均の順にアルファベットを与えた。不合格者の平均を用いず合格者の平均のみを用いて同様の計算を行うと多少の結果の違いが生じる。クラスターの安定度を検討するための資料として、表 1 にこれら 2 種

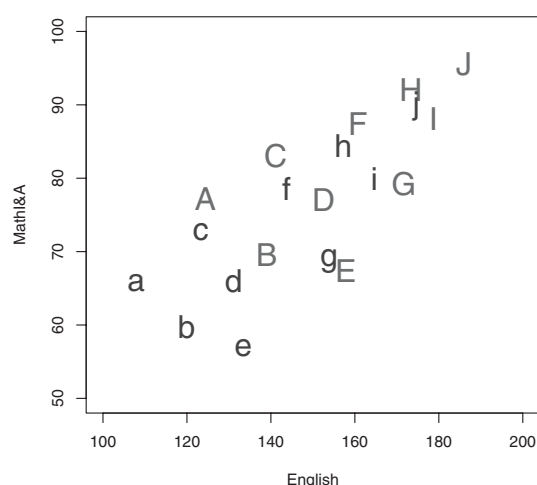


図 9 募集単位クラスターの入学者と不合格者の平均（横軸：英語，縦軸：数学 I・A，大文字は入学者平均，小文字は不合格者平均を示す。）

類のクラスタリングの結果を示す。クラスターの構成は少なからず変化するが，所属の変化の多くはクラスターの周辺部分で生じている。以降では最初の方法によって得られたクラスター（表 1 の行にあたる）に基づいて分析を行う。

図 9 には，各々のクラスターに含まれる募集単位の入学者の平均得点と不合格者の平均得点とを「英語」と「数学 I・A」について示す。クラスターを表す文字の位置が平均得点を表す。大文字は入学者の平均であり，小文字は不合格者の平均である。当然ながらいずれのクラスターにおいても，不合格者の平均は「英語」，「数学 I・A」とともに低く，各クラスターの入学者の得点平均の配置は，不合格者の得点平均の配置を右上方向にほぼ平行移動したものになっている。ここに示したのは 2 変数のみだが，各クラスターの「国語 I・II」の平均得点は「英語」とほぼ線形の関係にあり，また「数学 II・B」は「数学 I・A」と類似の関係にある（図 20，図 22，および表 5 を参照）。

また表 4 には，それぞれのクラスターに多い学部の種別を示した。

表 2 には，クラスター別の志願者数および入学者数等を示す。表側はクラスターを表し，表頭が可否の種別を表す。ここで「辞退」と表示したのは，合格したが入学手続きを行わなかった者の数である。表のうち Z で表された行は，前期日程においていずれの国公立大学への志願も行っていない志願者である。志願者数は Z を除いてクラスター

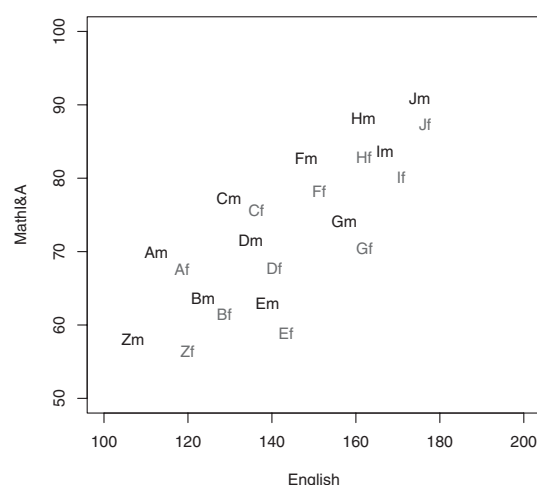


図 10 クラスター別・性別の志願者得点平均（横軸：英語，縦軸：数学 I・A）（m は男子，f は女子）

H が最も大きく，次いで D が多い。最も少ないのは，クラスター I であり，次に A である。各行の下段の括弧内には，クラスターの志願者のうちの各可否種別に該当する人数をパーセントで示した。

志願者数に対する入学者の比率は，C が最も大きく 37.82 パーセントである。一方，最もこの比率が小さいのは B であり 25.53 パーセントである。また他のクラスターはいずれも 30 パーセントを下回っている。2 段階不合格となったのは志願者全体のうち 3,561 名である。このうち 2,258 名（63.41 パーセント）がクラスター J の志願者である。また，志願者に対する辞退者の率は，クラスター E で最も大きく 6.01 パーセント，次いで B が 5.04 パーセントとなっている。

志願者における女子の比率は，クラスターによってかなり異なる（表 3）。最も女子志願者の比率が多いのは，E であり次いで D と G である。入学者における女子の比率も，志願者における比率とほぼ同様であり，これら 3 つのクラスターで半数以上の入学者が女子となっている。一方，女子の志願者が少ないのは C であり，次いで A である。表 3 の下段には，入学者における女子比率から志願者における女子比率を引いたものを示した（表中では「入学女子率 - 志願女子率」と表記した。文系の学部が多いクラスター G でこの値が最も大きく，逆にクラスター H で小さい。この理由は明らかではないが，図 9 の左上方向のクラスターでこの値が小さく，逆方向で大きい傾向が見取れる。図 10 には，「英語」と「数学 I・A」の科目得

表 2 クラスター別合否別受験者数 (括弧内は%)

クラスター	受験者数					計	2 段階不合格 実施募集単位数
	入学者	辞退	追加合格	2 段階不合格	不合格		
A	4973 (28.46)	746 (4.27)	21 (0.12)	22 (0.13)	11714 (67.03)	17476 (100)	1
B	7661 (25.53)	1512 (5.04)	34 (0.11)	0 (0)	20806 (69.32)	30013 (100)	0
C	9533 (37.82)	944 (3.75)	29 (0.12)	0 (0)	14698 (58.32)	25204 (100)	0
D	11292 (28.86)	1068 (2.73)	62 (0.16)	105 (0.27)	26601 (67.98)	39128 (100)	1
E	4643 (26.75)	1043 (6.01)	9 (0.05)	0 (0)	11663 (67.19)	17358 (100)	0
F	9411 (29.90)	694 (2.20)	52 (0.17)	98 (0.31)	21224 (67.42)	31479 (100)	4
G	8568 (28.21)	1343 (4.42)	14 (0.05)	26 (0.09)	20424 (67.24)	30375 (100)	1
H	12077 (27.85)	429 (0.99)	58 (0.13)	540 (1.25)	30259 (69.78)	43363 (100)	6
I	4789 (29.81)	267 (1.66)	11 (0.07)	512 (3.19)	10487 (65.27)	16066 (100)	4
J	6798 (28.67)	55 (0.23)	15 (0.06)	2258 (9.52)	14583 (61.51)	23709 (100)	19
合計	79745 (29.09)	8101 (2.95)	305 (0.11)	3561 (1.30)	182459 (66.55)	274171 (100)	36
Z	—	—	—	—	—	313179	
総計						587350	

「辞退」とは、合格したが入学手続きを行わなかった受験者を意味する。クラスター Z は前期日程の出願を行っていない志願者からなる。

表 3 クラスター別志願者・入学者の女子比率 (%)

	A	B	C	D	E	F	G	H	I	J	Z
志願者の女子比率	16.6	46.4	15.5	53.7	62.1	30.3	53.5	30.5	39.1	26.4	42.4
入学者の女子比率	15.3	47.1	14.4	55.3	64.8	28.6	57.8	28.0	42.0	25.5	—
入学女子率-志願女子率	-1.3	0.7	-1.1	1.6	2.7	-1.7	4.3	-2.5	2.9	-0.9	—

表 4 クラスターの特徴

クラスター	募集単位数	多い学部
A	39	工学部 (夜間) 10, 工学部 9, 経済学部 (夜間) 3
B	69	看護学部 9, 経済学部 8, 教育学部 5
C	48	工学部 23, 理学部 10, 看護学部 2, 情報学部 2
D	82	教育学部 28, 看護学部 8, 経済学部 7
E	42	人文学部 8, 文学部 5, 法学部 3
F	62	理学部 11, 工学部 11, 農学部 9
G	48	文学部 10, 法学部 8, 教育学部 4, 外国語学部 4
H	78	医学部 28, 歯学部 10, 薬学部 8
I	35	文学部 5, 法学部 5, 経済学部 5
J	33	医学部 12, 薬学部 4, 法学部 3

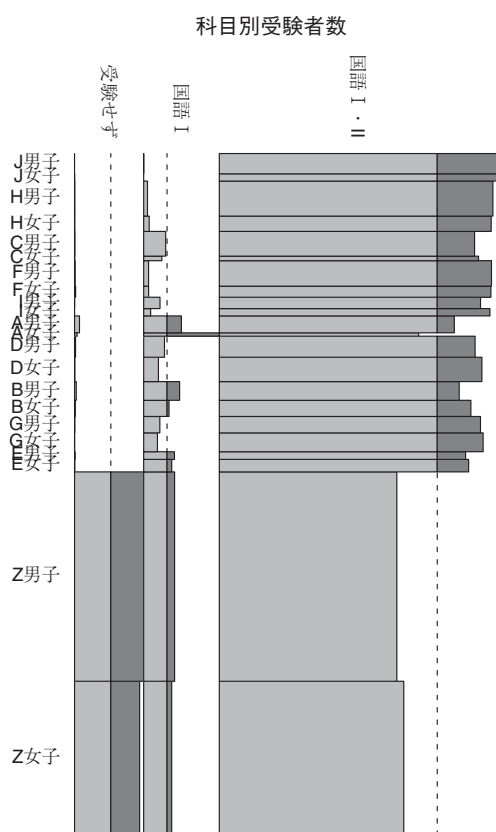


図 11 クラスター別国語科目選択者数

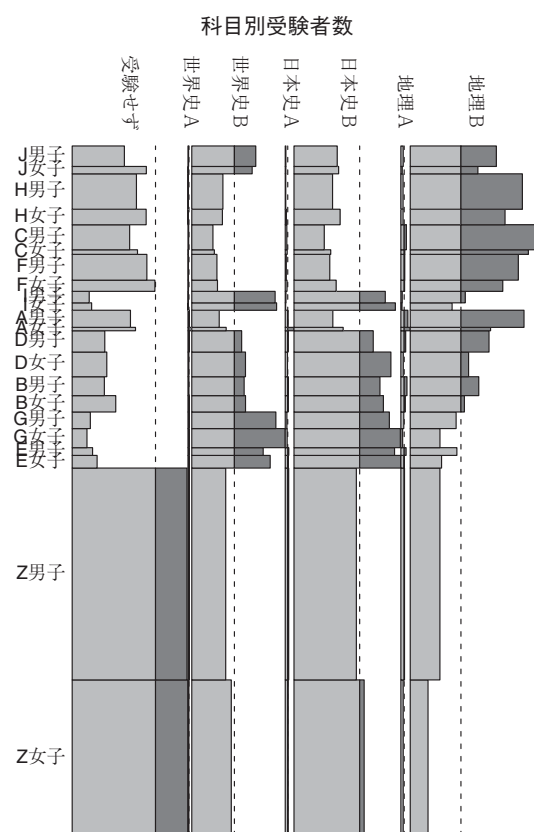


図 12 クラスター別地理歴史科目選択者数

点の平均をクラスター別、性別ごとに示した。表示してある文字はクラスターと性別（m は男子，f は女子）を示す。

5 クラスター別科目選択の状況

次に、各クラスターにおける科目選択状況を検討する。平成 16 年センター試験の実施は 9 つの時間単位からなる。各実施時間は、「国語」、「地理歴史」、「公民」、「数学①」、「数学②」、「理科①」、「理科②」、「理科③」、および「外国語」である。それぞれの実施時間単位において受験者は 1 つの科目のみを受験することができる。異なる実施時間の間の受験制限はないため、受験者は任意の科目を受験できる。ただし、センター試験の出願時に 2 科目以内かあるいは 3 科目以上の受験を行うかの選択がある。

これら 9 つの実施時間の内、ほとんどの志願者が「英語」を選択している「外国語」を除く 8 つについて、ジャック・ベルタンによる重み付き可換マトリックス表示 (Bertin, 1981) を用いて科目

選択の状況を示す。煩雑になるため詳細な数値についてはここでは省略する。

図 11 から 18 まではクラスター別、男女別の科目選択者数を示している。これらの図は科目選択率を示す横向きの棒グラフを並べたものである。各行は各クラスターにおける男女別の志願者の集合を表しており、行方向の幅は該当する志願者の数に比例するよう設定してある。このため各セルの部分に表示された矩形の面積は、該当する部分の志願者数に比例している。また、志願者全体における平均科目選択率を破線で示し、これを超える部分の表示を強調した。

最初の図 11 は「国語」の科目別受験者数を、次の図 12「地理・歴史」のものを示す。クラスターの表示順は、「数学 II・B」の選択率の高いものから低いものへと配置してある。「地理 B」の選択率に一貫した傾向が見て取れる。また、「世界史 B」、「日本史 B」の選択傾向が「地理 B」と相補的であることもわかる。

次の図 13 は「公民」の科目別受験者数を示している。いずれのクラスターにおいても「現代社

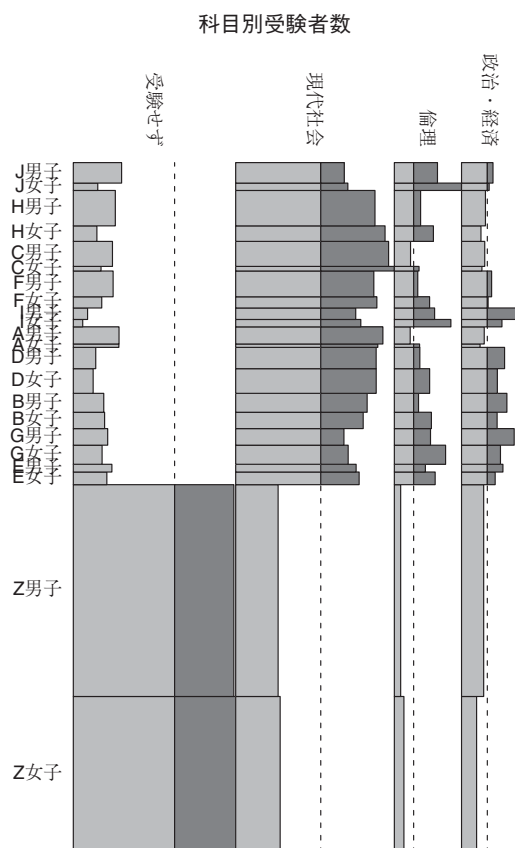


図 13 クラスター別公民科目選択者数

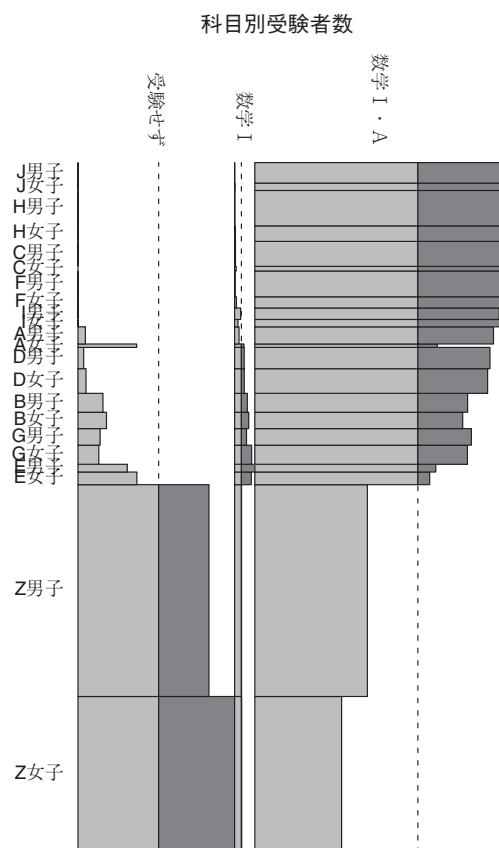


図 14 クラスター別数学①科目選択者数

会」の選択率が高いが、特に一貫した傾向はみられない。ただし、いずれのクラスターにおいても「倫理」の選択率は男子よりも女子が大きく、また「政治・経済」の選択率は男子が大きい。

図 14 は「数学①」における科目別受験者数である。上から次第に「数学 I・A」の選択率が小さくなる傾向が見られる。ただしクラスター A の女子は、男子に比べて顕著にこの選択率が低く（男子 95%，女子 73%），クラスター E と同程度であることがわかる。クラスター E の「数学 I・A」の選択率は男子が 72%，女子が 70% である。

図 15 は「数学②」の科目別受験者数を示している。クラスターの提示順は、この枠の「数学 II・B」の選択率に基づいて定めた。「数学①」と類似の傾向がここでも見られるが、「数学 II・B」の選択率はより急に減少している。クラスター E では、「数学 II・B」の選択率は男子で 57%，女子で 53% である。

理科は 3 つの時間区分にわたって行われた。「理科①」においては、「総合理科」、「物理 IA」、「物理

IB」が行われた。図 16 はこれらの科目の受験者数を示している。これらの科目の中では、「物理 IB」の選択率が最も高く、「数学 II・B」の選択率の低いクラスターでは「物理 IB」の選択率が低く、一方「総合理科」の選択率は相対的に高い。数学における科目選択と異なるのは、クラスター I で例外的に「物理 IB」の選択率が低いこと（男子 17%，女子 8%），および性別により「物理 IB」の選択率が著しく異なることである。いずれのクラスターにおいても、男子の選択率が大きい。クラスター J では、男子の「物理 IB」選択率は 58%，女子は 30% である。また、クラスター H では男子は 77%，女子は 39% であり、いずれも 2 倍近い違いがある。この傾向は、前期日程において国公立大学へ志願していないクラスター Z においても同様に認められる。

図 17 は、「理科②」の科目選択者数を示している。「化学 IB」の選択者が全体で最も多いが、数学の選択率の低いクラスターでは「化学 IB」の選択率も低い。「物理 IB」と同様にクラスター I の

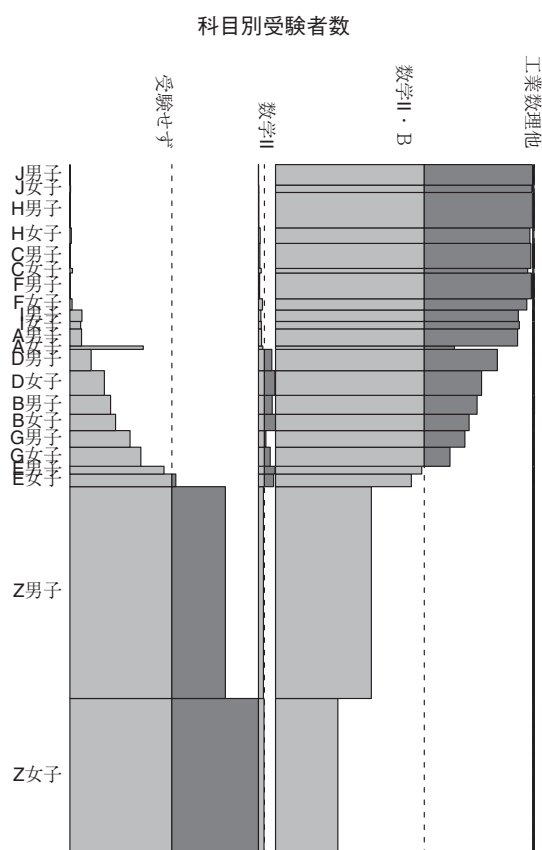


図 15 クラスター別数学②科目選択者数

選択率が低く（男子 29%，女子 27%），ここではクラスター B と近い値になっている．ただし，「物理 IB」において見られたような性別による選択率の違いは認められない．また，「地学 IB」の選択率は，クラスター H，C，F，A で低く，他のクラスターで相対的に高い．

図 18 は，「理科③」の科目選択者数を示す．「生物 IA」と「生物 IB」の試験がこの時間に実施されている．志願者全体の 31% が「生物 IB」を受験している．「物理 IB」とは逆の性別についての傾向が，ここでも一貫して見られる．全てのクラスターにおいて，女子の「生物 IB」の選択率は男子よりもかなり大きい．特に，クラスター H（男子 24%，女子 61%），C（男子 15%，女子 48%），F（男子 27%，女子 65%）で差が大きい．

6 クラスター別の科目得点

上述の募集クラスター単位における科目別得点の平均値を，入学者と不合格者の群について検討

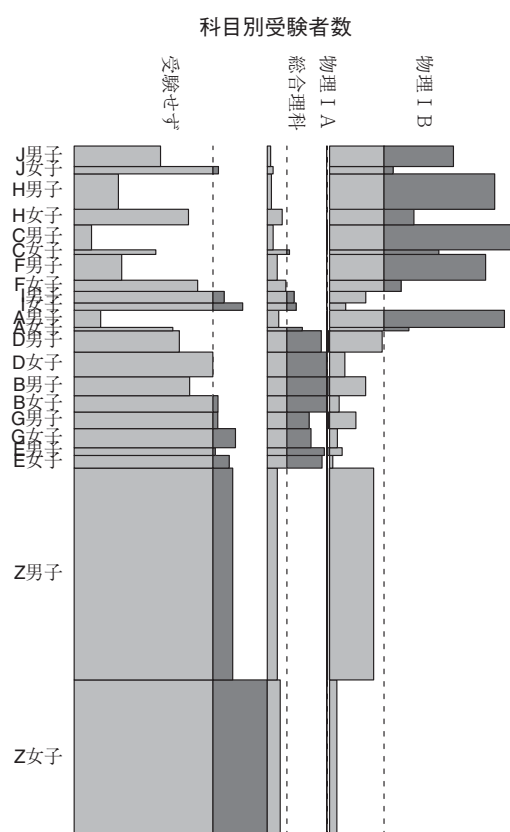


図 16 クラスター別理科①科目選択者数

する．

ここでは科目を限定し，受験者数の多い「国語 I・A」，「現代社会」，「数学 I・A」，「数学 II・B」，「化学 IB」，「生物 IB」，「英語」について検討を加える．ただし，2 段階選抜での不合格者および 2 次募集による合格者は含んでいない．センター試験の得点が合否とどの程度強く関係しているかを見るため，次の指標を各科目得点についてクラスターごとに求めた．この指標 d を「分離度」と呼ぶことにする．ここで $\bar{x}_1$ を入学者の得点平均， $\bar{x}_2$ を不合格者の得点平均とし， s_1 と s_2 をそれぞれの群の得点の標準偏差とする．また n_1 と n_2 をそれぞれの該当する受験者数とする．

$$d = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(n_1 s_1^2 + n_2 s_2^2) / (n_1 + n_2)}}$$

各クラスターにおける合否別受験者数，科目得点の平均と標準偏差は表 5 に示した．

上記の値 d は，得点のばらつきを考慮した入学者と不合格者のそれぞれの平均得点の分離の程度を表しており，芝・南風原 (1990) では効果量と呼

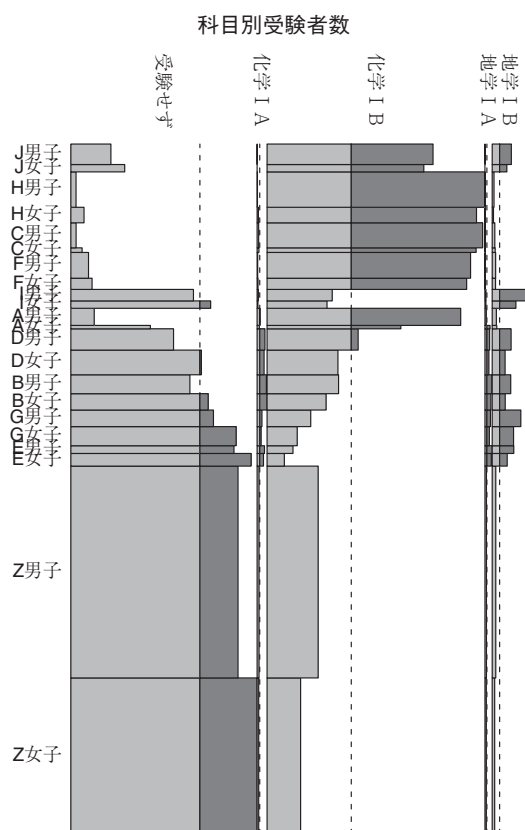


図 17 クラスター別理科②科目選択者数

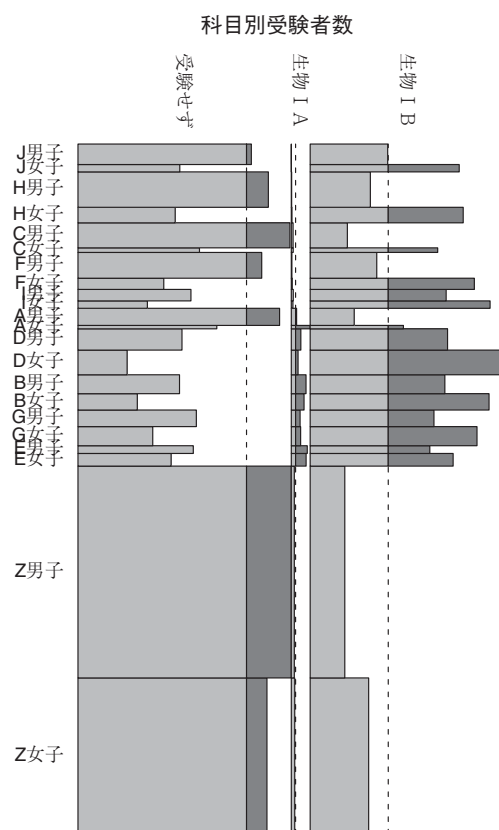


図 18 クラスター別理科③科目選択者数

ばれている。この値を上記 7 科目について求めた結果を表 6 に示す。ただし、クラスター J においては、多数の 2 段階選抜での不合格者がおり、これをここでは考慮に入れていないため、分離度を過大評価している可能性がある。また、図 19 には、表の値を多重の棒グラフとして表示した。各棒グラフの高さは分離度の大きさを表す。下端は 0.2 であり、上端は 1.0 となるよう行間を設定してある。横方向の破線は 0.7 の水準を表し、これを超える部分については表示色を変更して強調した。

それぞれの科目で分離度の傾向が異なることがわかる。ここで 0.7 より大きな値を示している箇所に注目すると (70 個の数値のうち、30 個が 0.7 を超えている)、「国語 I・II」ではクラスター D, E, G が該当する。また「現代社会」ではクラスター E のみで分離度が 0.7 を超えている。「数学 I・A」はクラスター A, C において高い分離度を示しており、「数学 II・B」では B, E, G を除くクラスターで分離度がよく、D では 0.63 とやや低い。数学の 2 科目は、相補的な性質を持っている。

また、「化学 IB」は B と E を除くクラスターで分離度が 0.7 を超えており、特に C と G で 0.9 に近い。一方、「生物 IB」では D, E および G において分離度が高く、「国語 I・II」と類似したパターンを示している。「英語」は A と H, J を除くクラスターで 0.7 を超える分離度を示している。ただし、A と H では 0.68 であり全般的に高い値を示している。

「化学 IB」がここに示した科目の中では高い分離度を示している。また「英語」も比較的高い分離度を示し、合否との関係が強いことがわかる。また、「化学 IB」と「生物 IB」とは同じ「理科」の科目ではあるが、それらの性質はかなり異なり、分離度の傾向は対照的である。

科目間の得点の関係を検討するために、「英語」と他の科目について得点の同時分布と、各クラスターの入学者および不合格者の得点平均を図 20 から図 24 に示す。

図 20 は、横軸に「英語」の得点、縦軸に「国語 I・II」の得点を示している。図中の小円は、それ

表 5 クラスター別合否別科目得点

群	合否	国語 I・II			現代社会			数学 I・A			数学 II・B		
		N	平均	S.D.	N	平均	S.D.	N	平均	S.D.	N	平均	S.D.
A	入学	4094	108.51	23.49	3115	52.11	11.16	4698	77.27	13.27	4653	47.82	13.22
	不合格	9609	100.62	23.61	7410	49.10	11.10	10544	65.75	15.81	10314	38.18	12.88
B	入学	6643	124.52	22.27	4253	60.28	11.95	6736	69.53	15.49	6282	39.98	13.47
	不合格	18484	109.99	23.31	11693	54.46	11.60	17257	59.76	15.88	15651	33.09	11.90
C	入学	8777	119.65	22.30	6683	56.24	11.30	9508	83.10	11.32	9466	55.22	13.13
	不合格	13527	109.20	24.02	9379	52.34	11.36	14618	72.67	15.44	14506	44.49	14.17
D	入学	10471	132.61	20.91	7041	63.01	11.94	10743	77.02	13.92	9888	46.55	14.45
	不合格	24914	115.42	23.79	15984	56.17	11.88	24638	66.07	16.65	21497	37.81	13.65
E	入学	4185	140.18	19.33	2471	64.64	11.84	3679	67.41	16.64	3048	35.43	12.86
	不合格	10310	120.66	24.00	6180	55.96	12.28	7863	56.97	17.08	5774	29.96	11.77
F	入学	9246	132.33	21.76	6108	61.54	12.04	9386	87.44	10.71	9335	62.57	15.00
	不合格	20770	120.72	24.59	12322	58.14	12.32	21140	78.62	15.83	20945	51.08	16.96
G	入学	8004	148.73	19.68	4306	71.62	11.39	7851	79.31	13.75	6844	47.22	14.90
	不合格	19391	132.23	24.10	9621	63.55	12.69	17091	68.74	17.63	13673	39.35	15.12
H	入学	11881	144.00	20.87	7838	66.00	12.20	12068	92.12	9.02	12033	72.90	15.91
	不合格	29761	130.94	24.99	18197	61.87	12.97	30198	84.45	14.81	30078	59.89	18.54
I	入学	4574	154.88	18.57	2481	76.03	11.64	4726	88.10	10.27	4608	61.70	14.97
	不合格	10099	141.74	23.01	5602	69.93	12.71	10309	80.01	14.99	9946	50.56	16.10
J	入学	6798	160.36	19.11	3264	74.15	12.07	6792	95.59	6.67	6783	81.17	14.40
	不合格	14579	147.50	22.70	6961	69.53	12.65	14569	89.83	12.21	14560	67.88	18.15
		化学 IB			生物 IB			英語					
		N	平均	S.D.	N	平均	S.D.	N	平均	S.D.			
A	入学	4315	50.70	13.20	814	57.55	19.10	4950	124.42	22.75			
	不合格	9266	40.64	13.10	2626	52.64	17.93	11652	107.94	25.03			
B	入学	2427	49.55	14.22	4796	65.78	13.20	7649	138.97	21.85			
	不合格	5877	40.28	13.89	12904	56.89	13.94	20774	119.90	24.93			
C	入学	9233	59.50	12.33	1719	68.73	17.82	9529	141.16	20.79			
	不合格	13833	47.29	14.34	3200	59.21	18.24	14687	123.29	25.67			
D	入学	4254	57.13	15.15	7530	70.53	12.33	11285	152.53	19.35			
	不合格	9197	45.32	15.35	17921	60.56	14.20	26571	131.14	25.40			
E	入学	414	46.43	15.54	2750	65.43	13.59	4641	157.82	18.45			
	不合格	1074	37.40	14.46	5976	55.44	14.43	11649	133.40	26.98			
F	入学	8682	68.69	12.42	3335	77.53	13.07	9400	160.69	18.15			
	不合格	18838	56.57	16.19	8355	69.18	15.68	21205	143.74	26.65			
G	入学	1429	63.54	14.33	5600	74.61	10.95	8545	171.63	15.92			
	不合格	3294	49.79	16.51	11527	64.97	14.33	20369	153.82	25.51			
H	入学	11563	77.57	12.03	3449	81.69	12.01	12066	173.40	15.97			
	不合格	28930	65.51	16.20	11333	73.12	15.87	30220	157.27	26.20			
I	入学	1332	71.61	12.69	2909	80.79	9.70	4781	178.76	12.79			
	不合格	3003	60.92	15.97	6506	72.95	12.96	10480	164.67	21.84			
J	入学	5031	84.32	10.40	2167	86.64	9.31	6779	186.03	11.02			
	不合格	10574	73.42	14.86	5756	79.39	13.26	14551	174.64	19.89			

N は当該部分の志願者数を表す。S.D. は標本標準偏差を示す。

ぞれ得点の範囲を 11 区分し、該当する箇所の受験者の数を面積で表している。区分は 0 以上 10 未満, 10 以上 30 未満, 以後 190 点まで 20 点刻みの区間とし、最後は 190 点以上 200 点までとしている。中央の曲線は、複数の科目に共通な能力の存

在を仮定し、その能力が一定であるときの条件付の 2 つの科目のそれぞれ平均値を非線形因子分析を用いて推定し、それらの点を結んだものである。この曲線はある意味において、難易度の同等性の推定値とも解釈しうるものであるが、計算手続き

表 6 入学者と不合格者の各クラスターにおける分離度

クラス ター	国語 I・II	現代 社会	数学 I・A	数学 II・B	化学 IB	生物 IB	英語
A	0.33	0.27	0.76	0.74	0.77	0.27	0.68
B	0.63	0.50	0.62	0.56	0.66	0.65	0.79
C	0.45	0.34	0.75	0.78	0.90	0.53	0.75
D	0.75	0.58	0.69	0.63	0.77	0.73	0.90
E	0.86	0.71	0.62	0.45	0.61	0.70	0.98
F	0.49	0.28	0.61	0.70	0.80	0.56	0.70
G	0.72	0.66	0.64	0.52	0.87	0.72	0.77
H	0.55	0.32	0.57	0.73	0.80	0.57	0.68
I	0.60	0.49	0.59	0.71	0.71	0.65	0.72
J	0.59	0.37	0.53	0.78	0.80	0.59	0.65

入学者と不合格者の分離度

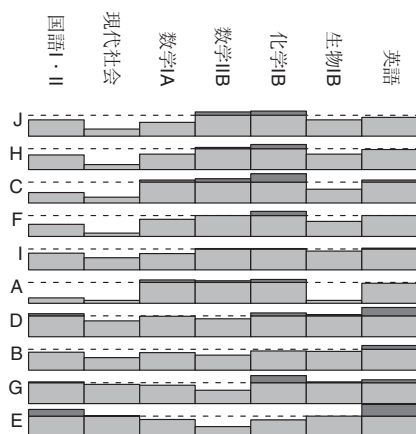


図 19 クラスター別合格者不合格者分離度

各セルの上端は1, 下端は0.2, 破線は0.7を示す。志願者中の数学 II・B の選択率順でクラスターを表示した。

については次節で説明する。

文字はクラスター別の平均点を表すが、集計の対象としているのは「国語 I」または「国語 I・II」を本試験で受験し、「英語」も本試験で受験している志願者である。「地理歴史」および「公民」を再試験または追試験で受験した志願者も対象から除いている。小円で表される同時分布の表示の対象となるのは、「国語 I・II」と「英語」の両者を本試験で受験したもののみであるが、「英語」のクラスター平均では、「国語 I」の受験者も集計対象となっている。

二つの科目得点の間には両者を受験した者について約 0.60 の相関があるが、募集単位のクラスター別得点平均には、ほぼ線形の明白な関係が認められる。これらのクラスターは「国語 I・II」「数

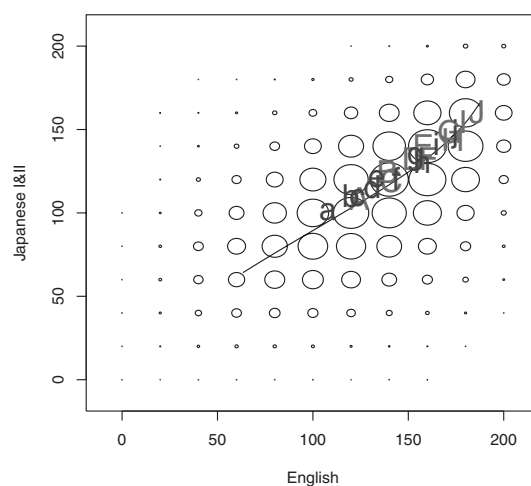


図 20 「英語」と「国語 I・II」の同時得点分布とクラスター別平均

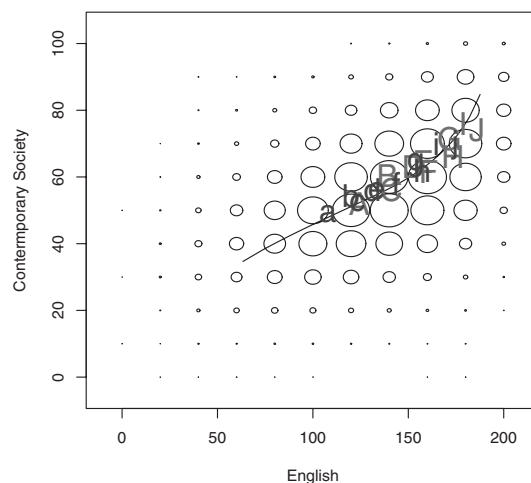


図 21 「英語」と「現代社会」の同時得点分布とクラスター別平均

学 I・A」「数学 II・B」および「英語」の得点に基づいて構成したものである。前掲の図 9 からは、数学と「英語」の間に 2 次元的な広がりがあることがわかるが、ここではそのような構造は示されていない。

図 21 は、同様に「英語」と「現代社会」の同時分布とクラスター別の得点平均を示したものである。百点満点の場合には、小円の表示のための区分は 0 から 5 点未満の区分、次いで 5 点以上 15 点未満の区分、以降 10 点刻みで 85 点以上 95 点未満の区分まで続き、最後は 95 点以上としている。図を見ると英語が 170 点より低い部分では、合否別クラスターの得点平均は、ほぼ直線状に配置されているが、より高得点の部分では傾きが変化し

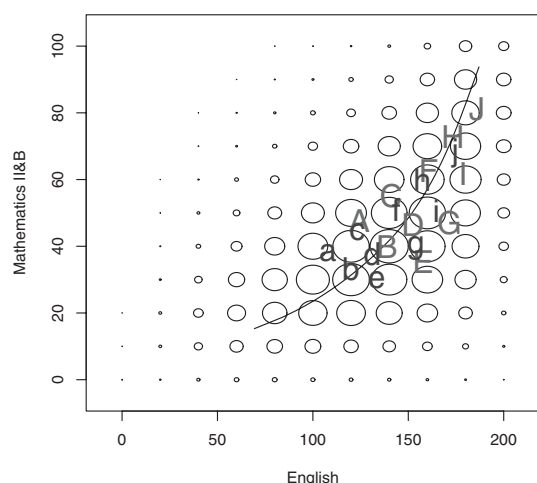


図 22 「英語」と「数学Ⅱ・B」の同時得点分布とクラスター別平均

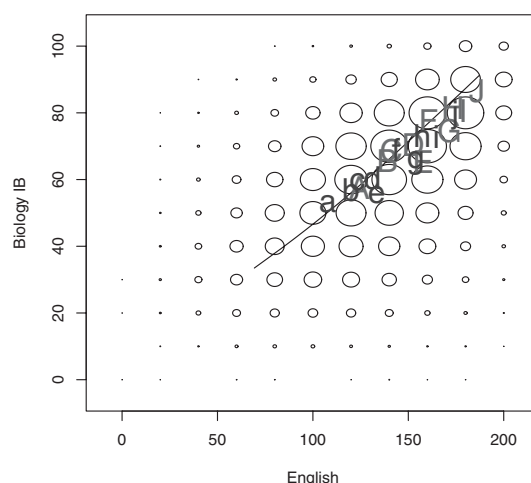


図 24 「英語」と「生物 IB」の同時得点分布とクラスター別平均

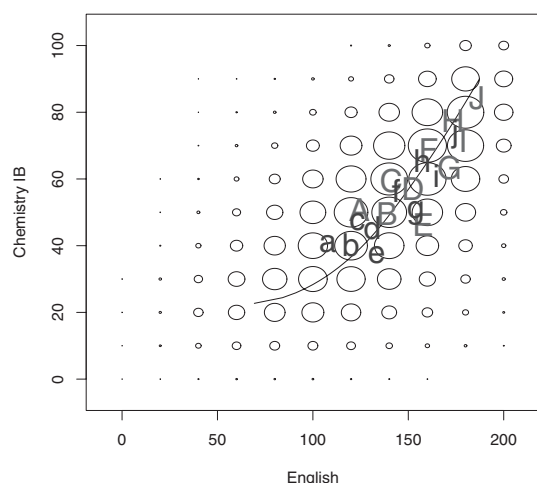


図 23 「英語」と「化学 IB」の同時得点分布とクラスター別平均

ているのが認められる。各クラスターの「現代社会」と「英語」の得点平均には、必ずしも線形ではない単調関係が見られる。

図 22 は、「英語」と「数学Ⅱ・B」の同時分布とクラスター別の科目得点の平均を示している。ただし、これ以降の図においては、クラスター平均の計算の対象となっているのは、「数学①」において「数学Ⅰ・A」を本試験で受験し、「外国語」では「英語」を本試験で受験したものに限る。また、「理科①」「理科②」「理科③」のいずれかを追試験で受験したものは、分析対象から除いてある。ここでは、図 9 と同様に三角状のクラスター平均の広がりが見られる。

次の図 23 は、「英語」と「化学 IB」についての

ものである。この図から「化学 IB」についてのクラスターの得点平均の構造は、「英語」と「数学」との中間的なものになっていることがわかる。

図 24 は、「英語」と「生物 IB」についてである。同じく理科の科目ではあるが、「生物 IB」のクラスター別の得点平均は「英語」と単調に近い関係にあることがわかる。少なくともクラスター単位の平均で見ると、限りは、「英語」や「国語」に近い得点パターンを持っていることが示されている。

7 複数科目の得点比較

センター試験得点の構造を分析する際の困難のひとつは、受験者が多様な科目選択パターンをとっているために、統一的な視点によってデータをまとめることが困難であることである。複数の科目の困難度を平均点によって単純に比較しようとする場合にも、「英語」を除いてはほとんど全ての受験者が受験している同一科目は存在しないために、結論はあいまいなものになる。ここではこの問題に対応するひとつの解決法として、欠測値に対応した非線形因子分析（大津，2004）を用いて得点分布についての推論を行う。

因子分析は複数の変数間の関係を、直接には観測されない潜在変数を仮定することにより簡潔に表現する統計的分析方法であり、心理テストの分析に伝統的に用いられてきた。ただし、通常用いられている手法は変数間の線形の関係のみを分析の対象としている。ここでは変数間の非線形の関

係を分析するために、スプライン関数によって変換された潜在変数（因子）によって観測変数の期待値が表現されるモデルを用いて分析を行う。またここで用いている推定方法では、潜在変数を離散近似することにより数値的な取り扱いが容易になり、MAR と呼ばれる仮定のもとでの欠測値への対応が容易となる。

非線形因子分析は観測変数と潜在変数の間に非線形の関係性を想定する統計モデルである。このような方法はかなり早くから提案されている (McDonald, 1962, 1967)。ここで提案する方法により密接な関係を持つ研究は項目反応曲線の推定にスプライン変換を用いた Ramsay & Winsberg (1991) である。McDonald の 1960 年代の研究は、多項式変換を用いた共分散構造の当てはめに基づいている。一方、Ramsay & Winsberg の研究は、2 値反応を前提としているものの潜在変数の非線形変換にスプライン曲線を用い、最尤法を用いた推定を行う点で、ここに示す方法と共通点が多い。また、センター試験におけるような大問構造を持つテストデータの分析法であるテストレット分析 (石塚・中畝・内田・前川, 2001; 石塚・前川・菊地・中畝・内田, 2003) ととも分析意図において共通する部分が多い。

以下で非線形因子分析モデルの説明を行う。

ここで Y_j は j 番目の観測変数であり、 $\mathbf{x}$ は正規分布に従う潜在変数を表す。

$$Y_j | \mathbf{x} \sim N(\eta_j(\mathbf{x}), \psi_j). \quad (2)$$

また η_j は潜在変数を固定した場合の Y_j の条件付平均を表し、また ψ_j は変数に固有の分散（独自分散）を示す。さらに、線形の因子分析と同様に、 $\mathbf{x}$ を固定したもとでの各 Y_j は統計的に独立であると仮定する。

非線形へのモデルの拡張を行うことにより、パラメータの同定が難しくなる可能性が大きい。線形の因子分析の推定においては、しばしば独自因子の分散の推定値がゼロとなる場合があり、非線形への拡張はこのようなケースをより頻繁にもたらす。また、周辺分布が複雑になり、簡明な式によっては尤度を表すことができない。これらの問題を回避するため、独自因子の分散に緩やかな事前分布を仮定する。また、潜在変数を離散近似することによって計算を容易にし、さらに非線形変

換 η にスプライン関数を用いることにより、多様な非線形関係を柔軟に表現できるようにする。

独自因子の分散 ψ_j の値を制約するために逆カイ 2 乗分布を事前分布として仮定することにより、パラメータ同定が困難になる問題を回避する。逆カイ 2 乗分布は、ベイズ推定において分散の事前分布としてしばしば用いられている。

確率密度関数を付録 A の (4) 式に示す。

以下の推定ではこの事前分布の自由度を $\nu = 2$ とし、尺度パラメータ s は観測変数の標準偏差に等しくなるように設定した。また潜在変数は 1 次元の正規分布を仮定し、50 個の等分位区間の平均値によって離散近似を行った。分布の両端（最小値および最大値の離散近似点）と中央に節点を設け、2 次の区分多項式をスプライン関数として用いた。

以下の分析において、 Y_j の条件付分散 ψ_j は潜在変数 $\mathbf{x}$ の値によらず一定であると仮定している。しかし、試験得点の場合には満点に近い場合には明らかに分散が小さくなると考えられるため、次の式によって科目得点を変換した上で分析を行った。

$$Y'_j = f_j(Y_j) = \arcsin \sqrt{\frac{Y_j}{\text{科目 } j \text{ の満点}}} \quad (3)$$

欠測を伴うデータの推定においては、欠測構造についてのなんらかの仮定の下に推定を行う必要がある。一般的な欠測メカニズムに対応した推定は容易ではなく、データ以外の情報を用いて欠測メカニズムをモデル化する必要がある。しかし、欠測メカニズムが MAR (Missing At Random) と呼ばれる性質を満たす場合には、データに基づく推定方法が簡明になる (Rubin, 1976)。

確率変数 Y_{ij} の欠測を示す指標を I_{ij} とする。ここで $I_{ij} = 1$ ならば Y_{ij} が観測されており、 $I_{ij} = 0$ ならば欠測を表す。この指標が確率変数であるとすると、そのモデルには様々なものが考えうる。一番単純なものは、 $\{Y_{ij}\}$ のいずれの値とも無関係な確率変数として I_{ij} が定義される場合である。これは MCAR (Missing Completely At Random) と呼ばれ、欠測の発生が無作為抽出によって生じていることを意味する。MCAR である場合には、欠測の発生したデータを除いて推定を行うと、無作為に標本を減少させて推定を行うのと同様であり、

推定効率とは下がるが、特別な偏りが生じることはない。

この条件を緩め、より適用範囲を広げたものが MAR である。確率変数 $\{Y_j\}$ のうち、観測されている値の集合を $\mathbf{y}_o$ とし、欠測している値の集合を $\mathbf{y}_m$ で表す。ここで、観測変数と同様に欠測の指標 $\{I_j\}$ を確率変数と考え、これらの全体、すなわち $(I_1, \dots, I_p)$ を I で表すものとし、その確率分布を $p(I)$ とする。一般的に I は $\mathbf{y}_o$ と $\mathbf{y}_m$ の両者に依存する可能性があるが、ここで I が $\mathbf{y}_m$ には依存せず、 $\mathbf{y}_o$ のみによって分布が定まるものとする。つまり、

$$p(I|\mathbf{y}_o, \mathbf{y}_m, \phi) = p(I|\mathbf{y}_o, \phi)$$

が成立しているとする。ここで ϕ は欠測発生分布にかかわるパラメータである。各 Y の分布を定めるパラメータを ζ とする。利用可能なのは $\mathbf{y}_o$ と I の値であり、これらの確率分布は

$$\begin{aligned} p(\mathbf{y}_o, I|\zeta, \phi) &= \int p(\mathbf{y}_o, \mathbf{y}_m, I|\zeta, \phi) d\mathbf{y}_m \\ &= \int p(\mathbf{y}_o, \mathbf{y}_m|\zeta) p(I|\mathbf{y}_o, \mathbf{y}_m, \phi) d\mathbf{y}_m \end{aligned}$$

で定義される。ここで MAR の仮定が成立しているならば、

$$p(I|\mathbf{y}_o, \mathbf{y}_m, \phi) = p(I|\mathbf{y}_o, \phi)$$

であり、上の式は

$$p(\mathbf{y}_o, I|\zeta, \phi) = \int p(\mathbf{y}_o, \mathbf{y}_m|\zeta) d\mathbf{y}_m \times p(I|\mathbf{y}_o, \phi)$$

となる。ここで求めたいのは ζ の値であるが、 ζ についての尤度は

$$\int p(\mathbf{y}_o, \mathbf{y}_m|\zeta) d\mathbf{y}_m$$

によって定まり、欠測メカニズム $p(I|\mathbf{y}_o, \mathbf{y}_m, \phi)$ を考慮する必要がない。

潜在変数を離散近似した非線形因子分析モデルの尤度は MAR の仮定のもとで簡明に表すことができる。これを付録 A の (5) に示す。

MAR の仮定は、欠測の発生確率が観測されている変数の値によって予測可能であり、欠測している値自身には依存しないことを主張するものである。ここで、国語、数学、英語の 3 科目からなる仮想的な試験の例を考えよう。一部の受験者は数

学を受験していないが、国語と英語とは全員が受験しているものとする。また、数学を受験しないことを表す欠測変数を I とする。MAR の仮定は、 I の分布が国語と英語の得点にのみ依存し、数学を受験していない学生の（受験したら得られるであろう）数学の得点には依存しないことを意味する。つまり、国語と英語の得点と同じならば、数学を受験したなら得られるであろう得点の値にはよらず、数学の受験率は一定であると仮定するものである。

入学試験の場合、このような仮定が厳密に成り立っていると想定することは明らかに困難であり、受験したなら得られるであろう得点が低い受験者において欠測が多く発生していると仮定するほうが自然であろう。しかしながら、この条件を表現するためには、得点と欠測発生との関係を明示的にモデル化する必要があるため、科目成績と受験率との関係についての情報を与える他のデータを利用しなければ推定が難しい。また、多くの科目についての得点が利用できる場合については、厳密に MAR が成立しておらずともある程度妥当な推定が可能であると考えられる。

以下では、「国語」、「地理歴史」、「公民」、および「外国語」のうち「英語」についての分析と、「数学」および「理科」の主要科目についての分析例について結果を簡単に示す。

最初の分析では、「国語 I」、「国語 I・II」、「世界史 A」、「世界史 B」、「日本史 A」、「日本史 B」、「地理 A」、「地理 B」、「現代社会」、「倫理」、「政治・経済」、「英語」の 12 科目の得点を対象とする。また、受験者のうち、「国語」のいずれかの科目を本試験で受験し、「外国語」の試験において「英語」を本試験で受験した者に分析対象を限定する。また、その他の科目においても追試験または再試験を受験したものは分析対象から除くことにする。この結果分析対象となった受験者は 507,163 名となり、これは英語の全受験者の 94.6 パーセントである。

ここで潜在変数は 50 個の離散値 $x_k, k = 1, \dots, 50 (= K)$ で近似しているが、変換後のそれぞれの科目得点の分布は $\eta_j(x_k)$ で近似される。潜在変数についての条件付平均を逆変換した値は

$$g_j(\eta_j(x_k)), (k = 1, \dots, K)$$

で与えられる。

図 20 と 21 における曲線は、上の分析における $(g_j(\eta_j(x_k)), g_{j'}(\eta_{j'}(x_k)))$ を繋いだものである。10 個の募集単位のクラスターの入学者平均および不合格者平均ともに、この曲線に近い箇所に位置している。同一クラスターに属する募集単位への志願者は、互いに類似した学力層に属すると考えられるため、これらのクラスター平均は複数科目の難度比較のひとつの基準を与えるものと考えられる。異なる推論によって得られた結論が、ほぼ同一の結果を与えることは、それぞれの推論の有効性を強める。

もうひとつの分析は、「英語」および数学と理科の 7 科目（「数学 I・A」, 「数学 II・B」, 「総合理科」, 「物理 IB」, 「化学 IB」, 「生物 IB」, 「地学 IB」）を対象とするものである。こちらは、「数学①」の本試験において「数学 I・A」を受験し、かつ「外国語」の本試験において「英語」を受験した者を分析対象とした。また最初の分析と同様に、「数学②」, 「理科①」, 「理科②」, 「理科③」において追試験または再試験を受験したものは除いてある。分析対象となった受験者は 380,463 名であり、先の分析の場合と比べ分析対象者がかなり限定されている。また、「数学②」, および「理科①」, 「理科②」, 「理科③」における受験者数の少ない他の科目は分析が困難と思われるためここでは検討から除いた。

図 22 から図 24 における中央の曲線は、これら「英語」を含む 7 科目の分析によって得られた潜在変数の条件付平均を前例と同様に表示したものである。これらの場合においても、「数学 II・B」を除き、クラスター毎の科目得点の平均がほぼ単調な関係にある場合には、それらの位置と潜在変数の条件付平均とはよく一致している。ただし、「数学 II・B」については、図 9 に見られるのと類似の構造が得られているのがわかる。興味深いことに、「化学 IB」と「英語」の関係を示す図 23 は、「数学 II・B」（図 22）と「生物 IB」（図 24）との中間的な特質を示していることがわかる。

8 ま と め

学部等の募集単位別にセンター試験の科目別得点を求めることにより、どのような得点層において、センター試験の得点が合否と強い関係をもっ

ているかが示された。ここでは、募集要項に示された条件は考慮せず、センター試験と合否の見かけ上の関係に基づいて分析を行ったことによるあいまいさは残るが、センター試験の科目別得点が、どのような得点水準において、合否と良く関連しているかの状況を示すことができた。

また、センター試験の科目別得点の平均値によって募集単位をクラスターに分割することにより、それぞれのクラスターにおける科目選択状況と科目別得点の特徴が明らかになった。ここでのクラスターは、「国語 I・II」「数学 I・A」「数学 II・B」「英語」の得点にのみ基づいて構成したものであり、他の科目得点を考慮していない。しかしながら、他の科目のクラスター別の平均得点を検討すると、クラスターを構成した 4 科目の得点と強い関連があり、他の科目得点の構造もこれらのクラスターによって表現されることがわかった。

クラスターは「国語」および「英語」の得点を表す空間軸、およびその軸と緩やかに関係している「数学」の得点を表す空間軸とによって表現される。ほぼ全ての科目において最も成績のよいクラスターである J は、文系、理工学系、医学系などが混在しており、ここでの分析では幾分特徴があいまいになっている。クラスター J に含まれる募集単位は、科目別得点の良さという点で一致しているのであって、より詳細な特徴では異質なものが混在していると思われる。これらの特徴は、センター試験の成績によっては判別しがたい。

さらに、クラスターによって志願者の男女比、科目別の選択率、辞退率などに大きな違いのあることがわかった。特に、理科の科目選択においては、同一のクラスターにおいても性別により、「物理 IB」と「生物 IB」の選択率に大きな違いのあることが示された。

試験内容の改善をはかるためには、それぞれの試験がどのような受験者によって利用されており、それが合否判定にどのように関わっているかを把握する必要がある。ここで示した分析は限定的なものではあるが、このような試みを実現し有益な知見を得たと判断される。

また、異なる科目の得点の比較においては、クラスターの得点平均と MAR の仮定に基づく非線形因子分析の結果が多くの場合一致することから、なお少人数科目での推論に問題はあるものの、こ

れら 2 つの分析方法が現実的に有効であることが強く示唆された。

これまで行われている科目得点の比較に関わる多くの研究は、暗黙のうちに推定の前提として MAR を仮定しているとみなせる。ここでの分析は、推定の前提を明示することにより、適用の限界をより明確に示すことができる。これは実用上重要な点であり、優れた理論研究がデータ分析の実際に与える影響の広範さを示すものでもある。

欠測に対応した非線形因子分析の枠組みは、広いパターンのテストデータの記述を的確に行える。一般的に利用されることの多いテスト得点の等化法に比べ、理論的な枠組みの柔軟さに特徴があるが、潜在変数による表現を用いるため（線形の因子分析においても共通する性質であるが）個人の得点の事後推定値は、平均に引き寄せられたものになる。純粋に統計的な推定の視点からみればこれは適切ではあるが、全てのテスト得点を平均に引き寄せた推定が行われるため、実用的な得点変換においては当事者の同意は得られにくいと思われる。一般的に利用されている等分位点法や Tucker 法は、2 つのテスト得点が同等となる変換を求めようとするものであるが、ここでの分析はむしろ必ずしも同等とはいえない複数の変数の関係に内在する構造を記述しようとするものである。

また、非線形の推定が適切に行われるためには、相当数のデータが必要であり、データ件数が少ない場合には、非線形性によってもたらされる推定の不安定さに注意を払う必要がある。また、ここでの推定は独自因子の分散 ψ_j が潜在変数 x の大きさによらず一定であることを仮定しているが、非線形モデルはこの仮定にかなり敏感であり、この前提が崩れるとしばしば不可解な推定結果をもたらすことに留意する必要がある。

ここでの分析は、平成 16 年度のセンター試験に限定したものであり、この年度に特有の特徴と複数の年度にわたって一般的に成立している特徴とを区別することができない。科目得点についての他の分析結果などを参照すると、国語、数学、英語の科目得点の相互関係は比較的安定したものであり、複数の年度にわたって成立していると思われる。ただし、科目別の難易度は、年度によってかなり変化することに留意する必要がある。

ここで試みた分析の枠組みはセンター試験デー

タの分析にかなり有効であることが示されたので、さらに複数年度にわたる分析を行うことにより、より有益な知見を得ることも可能と思われる。

謝辞：査読者の先生方の丁寧な指摘により、論文の内容と表記が大きく改善されました。感謝いたします。

参考文献

- 荒井克弘 (1982). 国公立大学における合否判定と共通第一次学力試験との相関, 大学入試センター研究紀要, **A4**, 1-34.
- Bertin, J. (1981). *Graphics and graphic information-processing*. translated by Berg, J.W. and Scott, P. Berlin: de Gruyter.
- Bickel, P.J., Hammel, E.A., & O'Connell, J.W. (1975). Sex bias in graduate admissions: Data from Berkely, *Science*, **187**, No.4175, 398-404.
- Ihaka, R., & Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, **5**, 299-314.
- 石塚智一・中畝菜穂子・内田照久・前川眞一 (2001). テストレットモデルによる英語試験問題の分析, 大学入試センター紀要, **No.30**, 21-38.
- 石塚智一・前川眞一・菊地賢一・中畝菜穂子・内田照久 (2003). テストレットモデルに基づく適応型テストシステム, 大学入試センター紀要, **No.32**, 1-12.
- Harman, H.H. (1976). *Modern factor analysis*. Chicago: The University of Chicago Press.
- Hartigan, J.A., & Wong, M.A. (1979). A *K*-means clustering algorithm. *Applied Statistics*, **28**, 100-108.
- McDonald, R.P. (1962). A general approach to nonlinear factor analysis. *Psychometrika*, **27**, 397-415.
- McDonald, R.P. (1967). *Nonlinear factor analysis*. Psychometric monograph **15**. Psychometric corporation.
- 大津起夫 (2004). 潜在変数の区分多項式変換を用いた非線形因子分析. 行動計量学, **31**, 1-15.
- R Development Core Team (2005). *R: A language and environment for statistical computing*, Vienna Austria: R Foundation for Statistical Computing. <http://www.R-project.org>.
- Ramsay, J., & Winsberg, S. (1991). Maximum marginal likelihood estimation for semipara-

metric item analysis, *Psychometrika*, **56**, 365–379.

Rubin, D.B. (1976). Inference and missing data. *Biometrika*, **63**, 681–592.

芝祐順・南風原朝和 (1990). 行動科学における統計解析法. 東京：東京大学出版会.

清水留三郎 (1981). 共通第 1 次学力試験の社会と理科における選択科目で分類した受験者群平均学力と問題の科目間難易度の推定, 大学入試センター研究紀要, **A1**, 1–14.

清水留三郎 (1983). 共通第 1 次学力試験の「社会」と「理科」の選択科目間における差異の統計解析, 大学入試センター研究紀要, **No.6**, 1–31.

清水留三郎 (1985). 共通第 1 次学力試験の成績から見た国公立大学・学部による入学者選抜の特性, 大学入試センター研究紀要, **A6**, 1–54.

柳井晴夫・繁榊算男・前川眞一・市川雅教 (1990). 因子分析 その理論と方法. 東京：朝倉書店.

吉村宰・莊島宏二郎 (2004). 本追モニター調査の結果を利用した傾向スコア加重法による本追試験間の難易度比較, 大学入試センター紀要, **No.33**, 19–28.

A 事前分布とモデルの尤度

独自因子の事前分布を表す逆カイ 2 乗分布の確率密度関数は次の式で定義される.

$$\text{Inv}\chi^2_\nu(\theta) = \frac{(\nu/2)^{\nu/2}}{\Gamma(\nu/2)} s^\nu \theta^{-\nu/2-1} \exp\left(\frac{-\nu s^2}{2\theta}\right) \quad (4)$$

ここで s は尺度パラメータである. この θ がゼロに近づくとも関数の値もゼロに近づく. この関数を独自因子の事前分布と仮定することにより独自因子の分散の推定値がゼロに近づくことを避ける.

離散近似されたモデルの尤度は次の式で表される. ここで, 添字集合 o_i は標本 i において観測された変数の集合を表す. 添え字 k は潜在変数を離散近似する点を表し K はそれらの個数を表す. また各 k に対応して設定された $w(\mathbf{x}_k)$ は, 離散化された $\mathbf{x}_k$ における重みを表す.

$$L(\boldsymbol{\theta}, \boldsymbol{\psi}|\mathbf{y}) = \prod_j \text{Inv}\chi^2(\psi_j|\nu, s_j^2) \times \prod_{i=1}^n \left\{ \sum_{k=1}^K \prod_{j \in o_i} p(y_{ij}|\eta_j(\mathbf{x}_k|\theta_j), \psi_j) w(\mathbf{x}_k) \right\} \quad (5)$$

On the Analysis of Relationship between Applicant's Test Scores and their Application Behavior in the First Period of The National Center Test 2004

OTSU Tatsuo*
HASHIMOTO Takamitsu*
SHOJIMA Kojiro*
ISHIZUKA Tomoichi*

Abstract

We analyzed the relationship between an applicant's scores on the National Center Test and the subsequent success or failure of his or her university application for the first period in 2004. Scatter plots that describe the mean scores of applicants who were admitted to a national or a public university are shown as well as those of applicants who failed to admit. The figures describe applicants' scores in four subjects: Japanese I and II, Mathematics I and A, Mathematics II and B, and English. The characteristics that related to an applicant's test result are different for subjects.

For detailed consideration, we classified university departments into ten clusters. The classification was based on the mean scores of successful and unsuccessful applicant's in the four subjects. The pass-fail status in the clusters is shown in a table. Although female applicants' ratio varied between departments, there was no significant difference in the pass ratio between male and female applicants in the clusters. We found a significant gender difference in the subject choice ratio for Physics IB and Biology IB. A simultaneous distribution of applicants' scores in several subjects is shown in scatter plots. The mean scores of applicants in the clusters are shown in the figures. Estimations on the relationship between subject scores by nonlinear factor analysis are shown.

Key words: subject selection, reorderable matrix graph, nonlinear factor analysis.

* Department of Applied Statistics and Measurement, Research Division, The National Center for University Entrance Examinations