

成績追跡システム

— 高校評定/入試成績/大学成績の統計解析 —

石岡 恒憲*, 野島 正之**, 石川 毅***

* リコー システムソリューション事業本部

** ヒワサキコンピュータシステム

*** 会津大学企画運営室

データ解析とグラフィックスのためのプログラミング環境である S-Plus を用いて、高校時代の成績、および科目履修の有無が、入試成績や大学での成績にどのように関与するかを統計的に解析・表示するシステムを開発した。ペアワイズの散布図や数種類の箱髭図の他、主成分分析やクラスター分析を行いその結果を図示する。

Grading Pursuit System

— Statistical analysis among students' evaluation in high-school,
entrance examination and university class —

Tsunenori Ishioka*, Masayuki Nojima**, Tsuyoshi Ishikawa***

* System Solution Business Group, Ricoh Co., Ltd.

** Hiwasaki Computer System Co., Ltd.

*** Office for Planning and Management, The University of Aizu

We developed a statistical analysis system which analyzes how students' evaluation in high-school relates to that in entrance examination and/or in university class. The system is implemented on the S-Plus, which is a programming environment for data analysis and graphics. It displays pairwise scatter plots, box-whiskers plots, hierarchical clustering plot trees, and principle components analysis charts.

1 はじめに

文部省は毎年、年度始めに「大学入学者選抜方法実施要項について」という通知を各大学に出している。これは、そこに列記されている事項に留意して、各大学が入試改善に努力して欲しいという要望書である。

93年度の通知では、13項目が挙げられているが、その最初の項目において、推薦入学や小論文、面接や職業高校の卒業生に対する選抜方法の工夫など、「評価尺度の多次元化・複数化」が求められ、「大学教育を受けるにふさわしい能力・適性等を多面的に判

定し、公正かつ妥当な方法で実施」することが要求されている。

当社（株式会社リコー）はこのような状況を踏まえ、成績追跡システム パッケージを開発・販売した。このシステムは、以下の機能を有する。

- 高校の成績および入試の成績と、入学後の成績とを比較することにより、情報科学に素質のある要因を統計的に抽出する。
- 高校時代の物理・数学などの科目履修の有無による違いを一对比較する。
- 一般入試合格者と推薦入試合格者との大学での成績の差異を一对比較する。現役/浪人による差異も比較する。
- 科目間の相関分析をおこなう。

これらにより推薦入学の決定の際の指針を与えるとともに、大学における科目間の関係を明らかにする。

2 システム概要

2.1 特長

- 統計処理に世界的に定評のある S-Plus[1]を用いており、信用のおける解析結果を、みやすい X-Window 上に表示する。
- ユーザは S-Plus を全く意識する必要がない。操作は数字あるいは 1 文字入力とマウス・クリックだけの簡単操作で済む。
- プラットホームを選ばない。X-Window システムが搭載されている UNIX 上で動作する。

2.2 ハードウェア構成

ハードウェアプラットフォームに UNIX ワークステーションを、業務端末に X 端末を利用する。

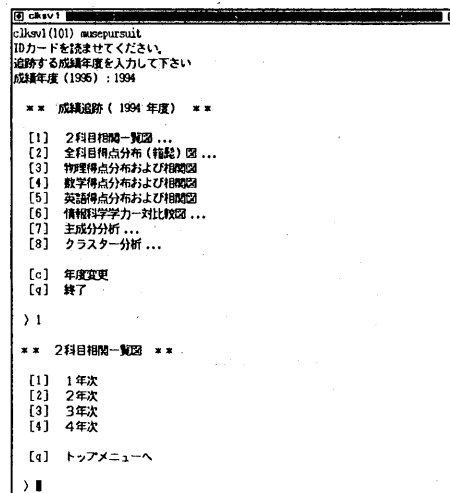
2.3 ソフトウェア構成

データベース管理システム(DBMS)にはリコー製拡張リレーショナル型データベース管理システム G-BASE を使用する。高校評定のデータ入力のための操作画面は tig(Terminal Interface for G-BASE)を使う。アプリケーション・プログラムは、S 言語[2]、および C 言語と、G-BASE が提供するプログラミングインターフェースやコマンドにて開発した。

3 機能概要と画面例

UNIX のプロンプトから、コマンドを入力し、システムを起動する (図 1)。このときに、磁気カードにより本システムを起動する権限の認証を行う。

ユーザは 1 から 8 の数字を入力し、解析・表示する図を選択する。項目 (例えば[1] 2 科目相関一覧図...) の行末にある「...」はサブメニューがあることを意味し、サブメニューでは学年を指定する (図 1)。



```
G-BASE 1
cjksw1(101) @asepursuit
IDカードを読ませてください。
追跡する成績年度を入力して下さい
成績年度 (1996) : 1994

** 成績追跡 ( 1994 年度 ) **

[1] 2科目相関一覧図 ...
[2] 全科目得点分布(相関)図 ...
[3] 物理得点分布および相関図
[4] 数学得点分布および相関図
[5] 英語得点分布および相関図 ...
[6] 情報科学学力-対比較図 ...
[7] 主成分分析 ...
[8] クラスター分析 ...

[6] 年度変更
[9] 終了

> 1

** 2科目相関一覧図 **

[1] 1年次
[2] 2年次
[3] 3年次
[4] 4年次

[9] トップメニューへ

> |
```

図 1: メインメニュー

以下は、各項目で表示される図とその出力例である。

2 科目相関一覧図

任意の2科目間の散布図を1枚にまとめる(図2)。

大学における科目間だけの相関だけでなく、入試の成績と大学における全科目の相関が見てとれる。散布図は、データ構造を見る上で最も基本的なグラフである。

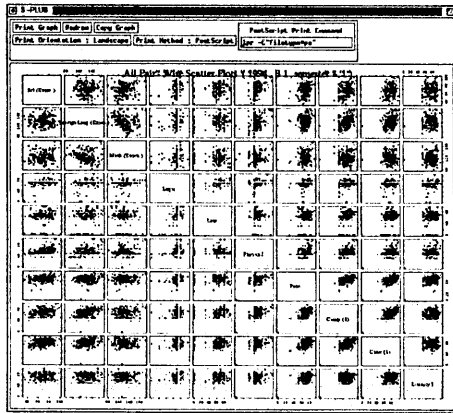


図 2: 2科目相関一覧図

全科目得点分布(箱髭)図

大学での全科目の得点分布を箱髭図で示す(図3)。これにより、中央値の位置だけでなく、得点分布の歪み具合や、裾の引き方、異状値の様子が見て取れる。

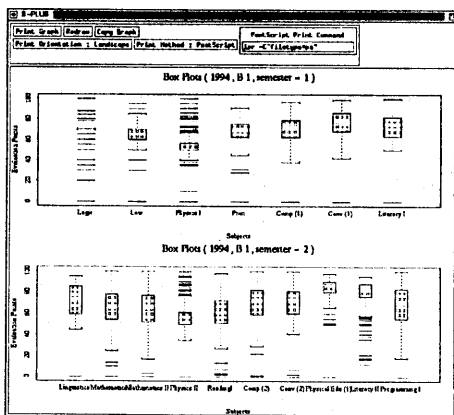


図 3: 全科目得点分布(箱髭)図

物理得点分布および相関図

高校時代に物理を履修したか否かで、大学での物理の成績にどのような違いがあるかをみるために、物理得点分布を2つの箱髭図で示す。

また、高校時代の理科の評定と、物理得点との間にどのような相関があるかを見るために、散布図を示す(図4)。

近年、大学入試科目が削減される方向にあったため、高校時代に物理を履修しないものが多く、あるいは履修しても受験科目としない場合もあり、大学での授業に支障をきたしていると言われている。したがって、本当に高校時代における物理履修の有無で、大学時代の物理成績が違ってくるのかをこれにより観察する。もし、高校での物理履修が大学時代の物理成績に影響があり、大学時代の物理がカリキュラム上、重要であると考えられるならば、高校時代の物理の成績がどの程度であるのかは、推薦入学の際の重要な指標となると考えられる。

高校時代における物理を履修している者に対しては、高校時代の物理の評定平均値と大学時代の全物理科目との散布図を示す。

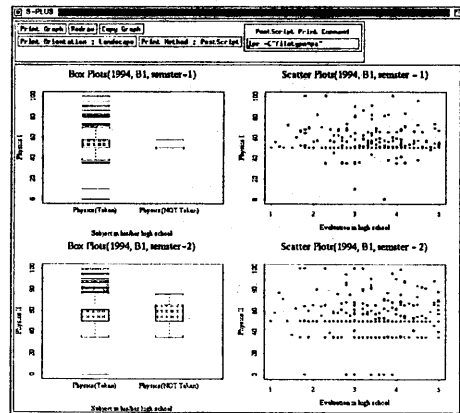


図 4: 物理得点分布および相関図

数学得点分布および相関図

高校時代に数学の上級コース（確率・統計など）を履修したか否かで、大学での数学の成績にどのような違いがあるかをみるために、数学得点分布を2つの箱髷図で示す。

また、高校時代の数学の評定と、数学得点との間にどのような相関があるかを見るために、散布図を示す（図5）。

現在の高校（普通科）の数学科目には、「数学Ⅰ」、「数学Ⅱ」、「代数・幾何」、「基礎解析」、「微分・積分」、「確率・統計」があるが、「微分・積分」や「確率・統計」の履修率が少なく、大学での授業に支障をきたしていると言われている。したがって、本当にそのような科目履修の有無で、大学時代の数学成績が違ってくるのか否かを観察する。もし本当に高校時代での科目履修が大学数学での成績に影響があるのであれば、高校時代にどの科目を履修し、その成績がどの程度であるのかは、推薦入学の際の重要な指標となると考えられる。

システムとしては高校時代における数学の上級コースを特に規定せずに、ユーザが着目したい科目を最大3つまで選択・指定できるように設計してある。これは、現高校2年生からカリキュラムが変わること、および職業科では数学科目が異なることによる。

高校時代においてその数学の上級コースを履修している者に対しては、高校時代のその科目の評定平均値と大学時代の全数学科目との散布図を示す。

英語得点分布および相関図

高校時代の英語の評定と、英語得点との間にどのような相関があるかを見るために、散布図を示す（図6）。

英語の場合は、ほとんど全ての学生が高校時代に履修していると考えられる。したがって、高校時代に英語を履修したか、いないかの別による大学時代の

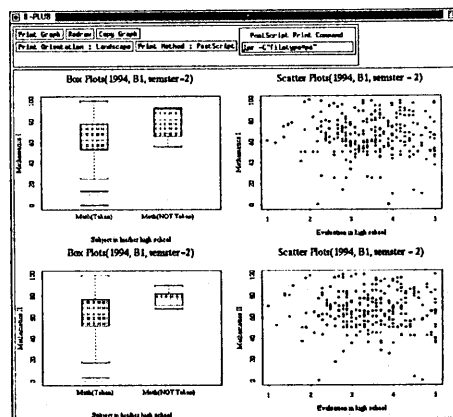


図5: 数学得点分布および相関図

成績分布の層別をしない。高校時代の英語の評定平均値と大学時代の全英語科目との散布図のみを示す。

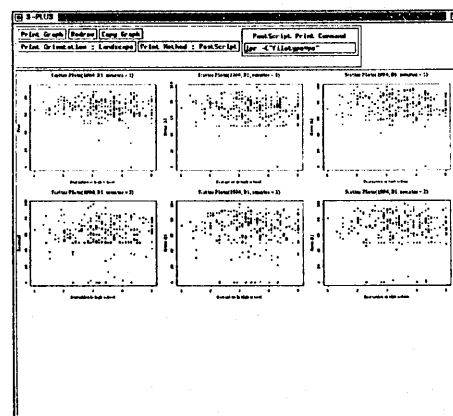


図6: 英語得点分布および相関図

情報科学学力一対比較

一般入試合格者と推薦入試合格者との間で、あるいは現役/一浪/二浪以上で、あるいは男女の別で、大学での成績がどのように違うのか、あるいは違わないのかを比較する（図7）。

平均的な差異はほとんどなくとも、分布の裾、たとえば成績優秀者や落ちこぼれの部分で、違いがある場合が少なくないようである。

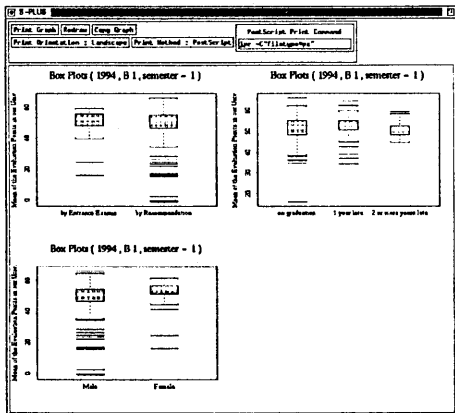


図 7: 情報科学学力一対比較

主成分分析

多変量データの背後にある現象の構造を明らかにする「記述モデル」の一つで、データを最もよく説明する成分(主成分と呼ぶ)を抽出し、その成分でどの程度、説明されているか(寄与率と呼ぶ)を定量的に示す(図8)。

第1主成分が総合的な学力を計る指標であるのに対し、第2主成分は一般に理科系・文科系を計る指標になるようである。第1主成分は、各科目の偏差値の合計で与えられることが知られている[3]。

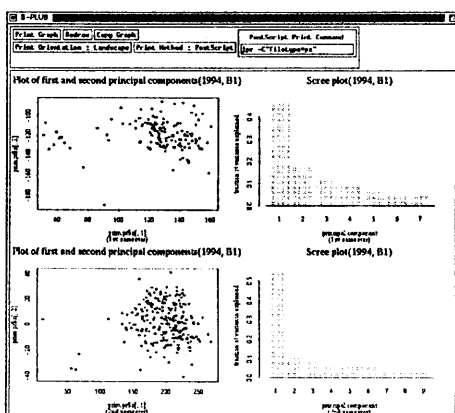


図 8: 主成分分析

クラスター分析

多次元空間上の空間的距離の近い科目をグルーピング化してツリーにして表示する(図9)。

グラフの下の部分でグルーピング化されているものほど、空間的距離が近いことを示している。これにより、入試科目に近い大学の科目は何か、すなわち入試ではどのような科目に近い学力を計っているのかがわかるようになる。若干の試行によると、入試の数学と関係の深いのが大学の数学ではなく、大学の物理であったりして、興味深い結果を与えることがあるようである。入試問題策定の指針を与えるものと考え

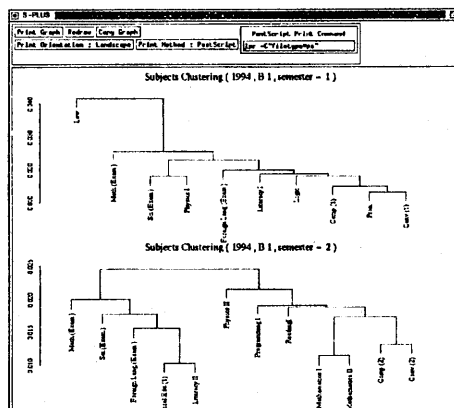


図 9: クラスター分析

これらの図は全て印刷時に EPS (Encapsulated PostScript)形式でカレントディレクトリにファイルとして保存される。したがって、 \LaTeX で書かれる報告書などの文書中に容易に取り込むことができる。

4 おわりに

本システムは、株式会社リコーシステムソリューション事業本部 ソフトウェア事業部が販売する大学管理事務システムのパッケージとして、既に会津大学様、および神

戸商科大学様に納入されることが決定している。今後も、より良いシステムに改良してゆく所存ですので、賢兄諸氏の忌憚ないご意見が戴ければ有難く存じます。お問合せは下記まで。

株式会社リコー システムソリューション事業本部 ソフトウェア事業部
〒112 東京都文京区小石川 1-1-17
とみん日生春日町ビル
Phone:(03)3815-7261
FAX:(03)3818-0348

謝辞

仕様策定に際し、鎌倉稔成教授（中央大学理工学部管理工学科）より有益なご助言を賜りました。ここに記して厚くお礼申し上げます。また、執筆の機会を与えてくれた國井秀子副事業本部長（株式会社リコー）、ならびに金崎克己課長代理研究員（株式会社リコー）に感謝申し上げます。

参考文献

- [1] S-Plus Reference Manual Ver.3.0, Statistical Sciences, Inc., 1991.
- [2] 「S 言語」(I, II), R.A.ベッカー, J.M.チェンバース, A.R.ウィルクス (渋谷政昭, 柴田里程, 訳), 共立出版, 1991.
- [3] “順序がかわる 2 つの偏差値,” 小川潤次郎, 数学セミナー, 31 巻, pp72-74, 1992.