

統計解析ソフトを使ったアンケートの因子分析

牧野 浩二

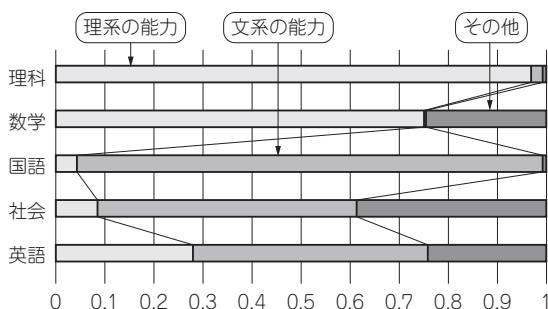


図1 因子分析は観測可能なデータから観測不可能なものの傾向をあぶり出す

本格的な応用ができるように、「R」という統計解析ソフトウェアを使った因子分析の実例を示します(図1)。

統計解析環境 R の準備

● データを準備する

本章でも英語、国語、社会などの成績表データを使います。

因子分析を始める前に、study_esk.csv をドキュメント・フォルダに移動します。なお、他のフォルダに置いたデータを使いたいときには「ファイルメニューからディレクトリの変更」を選択することで行ってください。

リスト1 まずは統計解析ソフトウェアRでread.csvコマンド&表示コマンドを実行してみる

	英語	社会	国語
A	9	7	10
B	4	5	8
C	8	5	5
D	4	2	3
E	9	10	9
F	8	5	7
G	6	4	4
H	7	7	9
I	9	4	9
J	7	6	5

● Rのコマンドに触れてみる…データの読み出し / 表示と共分散を求める

Rでstudy_esk.csvのデータを読み出すにはread.csvコマンドを用います。このコマンドは、dataだけ入力することで読み込まれたデータを表示させることができます。コマンドを次に示します。

```
>data = read.csv("study_esk.csv",
header=T)
>data
```

実行結果をリスト1に示します。

次に、共分散を次のcorコマンドで求めてみます。

```
>cor(data)
```

実行結果をリスト2に示します。この結果を見ると、英語と社会の共分散は0.5749711、英語と国語は0.5670651、社会と国語は0.6521236となっていることが分かります。これはExcelで求めた値と一致します。

まずは成績表サンプルで動かしてみる

● 共通因子が1つのとき

それでは、Rを用いて因子分析を行います。Rでの因子分析は、factanal関数を用います。そして、factors=1とすることで共通因子の数を1に設定しています。次のコマンドを実行します。

```
>f <- factanal(x=data,factors=1)
>print(f,cutOFF=0)
```

結果をリスト3(a)に示します。

リスト3(a)のUniquenessesが独自因子の分散、Loadingsが因子負荷量を示しています。これも先ほど求めた値と一致しています。

国語、数学、体育の3教科のデータはstudy_

リスト2 共分散を求めるコマンドを実行してみる

英語、社会、国語の成績に対する共分散(対角成分は1)

	英語	社会	国語
英語	1.0000000	0.5749711	0.5670651
社会	0.5749711	1.0000000	0.6521236
国語	0.5670651	0.6521236	1.0000000