社会システム分析のための統合化プログラム１４

ー コレスポンデンス分析・数量化Ⅲ類 ー

福井正康

福山平成大学経営学部経営学科

概要

我々は教育分野での利用を目的に社会システム分析に用いられる様々な手法を統合化したプログラム College Analysis を作成してきた。今回は、質的なデータに関する分類手法の１つであるコレスポンデンス分析と数量化Ⅲ類についてプログラムを作成した。数量化Ⅲ類についてはすでに College Analysis に含まれているが、今回コレスポンデンス分析と対応させるために、定義を少し変更して作り直した。

キーワード

College Analysis，社会システム分析，統計，OR，意思決定，コレスポンデンス分析，数量化Ⅲ類

URL:　http://www.heisei-u.ac.jp/ba/fukui/

１．はじめに

これまで社会システム分析ソフトウェア College Analysis に多くのプログラムを組み込んできたが、今回は質的データの分類手法である、コレスポンデンス分析[1]と数量化Ⅲ類についてプログラムを作成した。数量化Ⅲ類についてはすでに組み込んでいるが[2]、コレスポンデンス分析との対応関係を明確化するために定義を少し変えて再度作り直した。

コレスポンデンス分析は、２次元分割表で表される２つの変数のカテゴリの中で似たものを探し、分類する手法である。２つの変数に含まれるカテゴリは同列に扱われ、散布図上の表示ではすべて同一平面上に表される。それに対して数量化Ⅲ類は、分割表でなく個々のデータを用いて、カテゴリ同士、個体同士で似たものを探す手法で、散布図上の表示ではカテゴリと個体それぞれ別の図として表される。理論とプログラムの動作についてはそれぞれ２章と３章で詳述する。

最後に４章で重回帰分析に新しく追加した変数自動選択の機能や、因子分析の新しい実行画面について説明する。

２．コレスポンデンス分析

2.1 コレスポンデンス分析の理論

今 2 つの質的な変数、変数 1 と変数 2 があるとする。変数 1 のカテゴリ数を p、変数 2 のカテゴリ数を q（一般性を失わず$p \leq q$）とする。この 2 つの変数に対して p 行 q 列の 2 次元分割表を考え、変数 1 のカテゴリ i、変数 2 のカテゴリ j に属するデータ数をn_{ij}とする。またデータ数の合計を以下のように定義する。

$$n_{i\cdot} \equiv \sum_{j=1}^{q} n_{ij}, \quad n_{\cdot j} \equiv \sum_{i=1}^{p} n_{ij}, \quad n \equiv \sum_{i=1}^{p} \sum_{j=1}^{q} n_{ij}$$

次に変数 1 のカテゴリ i のデータに点数u_i、変数 2 のカテゴリ j のデータに点数v_jを与え、これらの点数の値によって各カテゴリ間の特徴的な関係を考えることとする。但し、これらの関係は変数 1 の点数と変数 2 の点数との相関係数を最大にするものとして与える。

これらの点数に対して、２つの変数の相関係数ρは以下のように与えられる。

$$\rho = \frac{S_{uv}}{S_u S_v} \ ,$$

$$S_{uv} = \frac{1}{n} \sum_{i=1}^{p} \sum_{j=1}^{q} n_{ij} u_i v_j, \quad S_u^2 = \frac{1}{n} \sum_{i=1}^{p} n_{i\cdot} u_i^2, \quad S_v^2 = \frac{1}{n} \sum_{j=1}^{q} n_{\cdot j} v_j^2$$

ここに、S_{uv}は共分散、S_u^2とS_v^2は分散であり、2 つの変数の点数について平均は 0 としている。

$$\bar{u} = \frac{1}{n} \sum_{i=1}^{p} n_{i\cdot} u_i = 0, \quad \bar{v} = \frac{1}{n} \sum_{j=1}^{q} n_{\cdot j} v_j = 0$$

この相関係数ρについて、点数の分散を 1 とする制約条件を付けて最大値を求めるために Lagrange の未定乗数法を用いる。

$$L = S_{uv} - \lambda(S_u^2 - 1) - \mu(S_v^2 - 1)$$

ここにλとμは未定乗数である。これをu_iとv_jで微分して、以下の方程式を得る。

$$\sum_{k=1}^{q} n_{ik} v_k - 2\lambda n_{i.} u_i = 0, \qquad \sum_{k=1}^{p} n_{kj} u_k - 2\mu n_{.j} v_j = 0$$

これらの式を行列で表示すると上式は以下のようになる。

$$\boldsymbol{N}\boldsymbol{v} = 2\lambda \boldsymbol{D}_r \boldsymbol{u}, \quad \boldsymbol{N}'\boldsymbol{u} = 2\mu \boldsymbol{D}_c \boldsymbol{v}$$

ここに

$$\boldsymbol{N} = \begin{pmatrix} n_{11} & \cdots & n_{1q} \\ \vdots & \ddots & \vdots \\ n_{p1} & \cdots & n_{pq} \end{pmatrix}, \quad \boldsymbol{D}_r = \begin{pmatrix} n_{1.} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & n_{p.} \end{pmatrix}, \quad \boldsymbol{D}_c = \begin{pmatrix} n_{.1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & n_{.q} \end{pmatrix},$$

$$\boldsymbol{u}' = (u_1 \quad \cdots \quad u_p), \quad \boldsymbol{v}' = (v_1 \quad \cdots \quad v_q)$$

上の方程式で、左式に左から$\boldsymbol{u}'$を掛けると$\rho = 2\lambda$、同様に右式に左から$\boldsymbol{v}'$を掛けると$\rho = 2\mu$を得る。右式を$\boldsymbol{v}$について解いて左式に代入すると以下となる。

$$\boldsymbol{D}_r^{-1}\boldsymbol{N}\boldsymbol{D}_c^{-1}\boldsymbol{N}'\boldsymbol{u} = \rho^2\boldsymbol{u}, \quad \text{また、} \boldsymbol{v} = \rho^{-1}\boldsymbol{D}_c^{-1}\boldsymbol{N}'\boldsymbol{u} \qquad (2.1)$$

また$\boldsymbol{v}$についても同様の関係が示されるが、ここでは省略する。

ここで$S_u^2 = 1$としたことから、$\boldsymbol{u}$の規格化条件を$\frac{1}{n}\boldsymbol{u}'\boldsymbol{D}_r\boldsymbol{u} = 1$として、新たに以下のベクトル$\boldsymbol{z}$を考える。

$$\boldsymbol{z} \equiv \frac{1}{\sqrt{n}}\boldsymbol{D}_r^{1/2}\boldsymbol{u}, \quad \text{ここに } \boldsymbol{D}_r^{1/2} = \begin{pmatrix} \sqrt{n_{1.}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sqrt{n_{p.}} \end{pmatrix}$$

これを用いて(2.1)式は最終的に以下となる。

$$\boldsymbol{A}\boldsymbol{z} = \rho^2\boldsymbol{z}, \quad \boldsymbol{z}'\boldsymbol{z} = 1, \ \boldsymbol{A} \equiv \boldsymbol{D}_r^{-1/2}\boldsymbol{N}\boldsymbol{D}_c^{-1}\boldsymbol{N}'\boldsymbol{D}_r^{-1/2} \qquad (2.2)$$

異なる固有値ρ_α^2（$\alpha = 1, \cdots, p$）に対する固有ベクトルを$\boldsymbol{z}^\alpha$とすると、各点数は以下のように表される。

$$\boldsymbol{u}^\alpha = \sqrt{n}\boldsymbol{D}_r^{-1/2}\boldsymbol{z}^\alpha, \quad \boldsymbol{v}^\alpha = \rho_\alpha^{-1}\sqrt{n}\boldsymbol{D}_c^{-1}\boldsymbol{N}'\boldsymbol{D}_r^{-1/2}\boldsymbol{z}^\alpha$$

ところで、(2.1) 式には $\rho^2 = 1, \ \boldsymbol{u} = \boldsymbol{1}$ の自明な解が存在し、それに基づく固有値と固有ベクトルが得られるが、この解は除外される。

その他、点数 $\boldsymbol{u}, \boldsymbol{v}$ の与え方には、以下のように相関係数を掛ける方法もある。

$$\tilde{\boldsymbol{u}}^\alpha = \rho_\alpha \boldsymbol{u}^{\boldsymbol{\alpha}}, \ \tilde{\boldsymbol{v}}^\alpha = \rho_\alpha \boldsymbol{v}^{\boldsymbol{\alpha}}$$

各成分の重要性を表すために、自明な解に対する固有値をρ_p^2として、以下で与えられる寄与率λ_αを考える場合もある。

$$\lambda_\alpha = \rho_\alpha^2 / \sum_{\beta=1}^{p-1} \rho_\beta^2 \qquad (\alpha \neq p)$$

2.2 プログラムの動作

メニュー「分析−多変量解析−コレスポンデンス分析」を選択すると図 2.1 に示される分析メニューが表示される。分析は通常の質的データと図 2.2 のような分割表の２通りから選択できる。

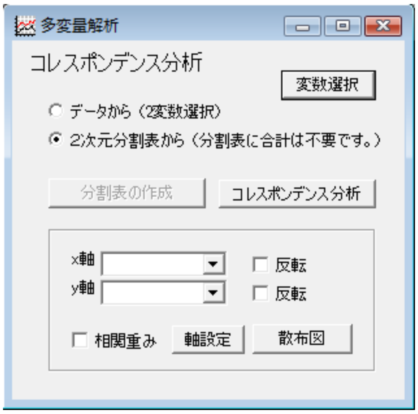
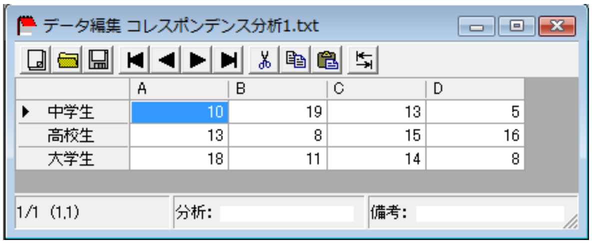


図 2.1 分析メニュー



	A	B	C	D
中学生	10	19	13	5
高校生	13	8	15	16
大学生	18	11	14	8

図 2.2 分割表データ

変数を選択して、「コレスポンデンス分析」ボタンをクリックすると図 2.3 のような分析結果が表示される。

	群	第1成分	第2成分	重み1成分	重み2成分
固有値		0.0763	0.0183		
相関係数		0.2762	0.1352		
中学生	1	-1.3287	-0.6528	-0.3670	-0.0883
高校生	1	1.1333	-0.7748	0.3130	-0.1048
大学生	1	0.0690	1.3916	0.0190	0.1882
A	2	0.2373	1.5238	0.0655	0.2060
B	2	-1.4691	-0.6411	-0.4058	-0.0867
C	2	0.0596	-0.1102	0.0165	-0.0149
D	2	1.5032	-1.1547	0.4152	-0.1561

図 2.3 コレスポンデンス分析実行結果

出力される成分数は 2 つの変数のカテゴリ数の小さい方から自明な固有値の数の 1 を引いた数であり、この例の場合 2 である。重み成分はそれぞれの成分に相関係数をかけたものである。

この結果を図の上で表示するには、まず「軸設定」ボタンをクリックし、図 2.4 のように x 軸と y 軸に表示される成分の中で適切なものを選択する。通常は x 軸に第 1 成分、y 軸に第 2 成分を表示する。「散布図」ボタンをクリックすると図 2.5 のような結果が表示される。

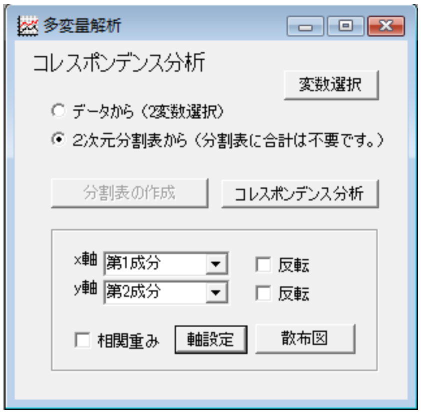


図 2.4 軸設定された分析メニュー

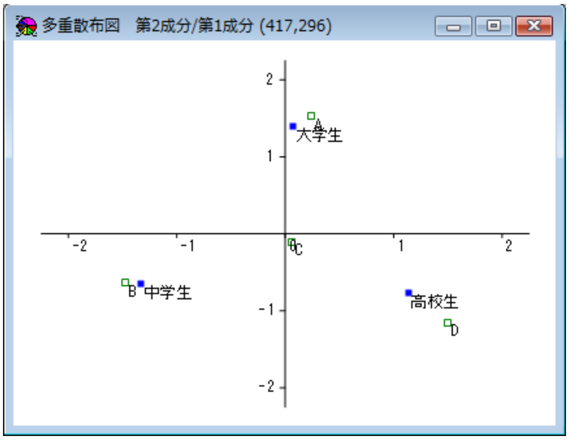


図 2.5 散布図画面

相関係数の重みを付ける場合は、「相関重み」チェックボックスにチェックを入れ、軸を反転させて表示したい場合は、それぞれの軸の「反転」チェックボックスにチェックを入れて散布図を表示する。

３． 数量化Ⅲ類

3.1 数量化Ⅲ類の理論

数量化Ⅲ類はカテゴリと個体にそれぞれ数値を与えて、特徴的な量を作りだし、データの持つ構造を解明しようとするものである。個々のデータはカテゴリに反応した場合 1、反応しない場合は 0 で与えられる。

$$x_{i\lambda} \in \{0, 1\}$$

ここに、i はカテゴリ、λ は個体を表わす。また、カテゴリ数をp、データ数をn（$p \le n$）とする。

カテゴリと個体に対してカテゴリウェイトと個体ウェイトと呼ばれる特徴的な点数u_iとv_λを得るために、まず以下のような点数u_iとv_λの分散と共分散を考える。

$$S_u^2 = \frac{1}{T}\sum_{i=1}^{p} c_i u_i^2,\ S_v^2 = \frac{1}{T}\sum_{\lambda=1}^{n} d_\lambda v_\lambda^2,$$

$$S_{uv} = \frac{1}{T}\sum_{i=1}^{p}\sum_{\lambda=1}^{n} x_{i\lambda} u_i v_\lambda,$$

ここに、

$$c_i = \sum_{\lambda=1}^{n} x_{i\lambda},\ d_\lambda = \sum_{i=1}^{p} x_{i\lambda},\ T = \sum_{i=1}^{p}\sum_{\lambda=1}^{n} x_{i\lambda},$$

であり、2 つの点数についての平均は 0 としている。

$$\bar{u} = \frac{1}{T}\sum_{i=1}^{p} c_i u_i = 0,\ \bar{v} = \frac{1}{T}\sum_{\lambda=1}^{n} d_\lambda v_\lambda = 0$$

これからカテゴリと個体の相関係数を$\rho = S_{uv}/S_u S_v$と表わし、この相関係数ρについて、点数の分散を 1 とする制約条件を付けて最大値を求めるために Lagrange の未定乗数法を用いる。

$$L = S_{uv} - \eta(S_u^2 - 1) - \mu(S_v^2 - 1)$$

ここにηとμは未定乗数である。これをu_iとv_λで微分して、以下の方程式を得る。

$\sum_{\lambda=1}^{n} x_{i\lambda} v_\lambda - 2\eta c_i u_i = 0,\ \sum_{i=1}^{p} x_{i\lambda} u_i - 2\mu d_\lambda v_\lambda = 0$

これらの式を行列で表示すると以下のようになる。

$$\boldsymbol{Xv} = 2\eta \boldsymbol{Cu},\quad \boldsymbol{X'u} = 2\mu \boldsymbol{Dv}$$

ここに

$$\boldsymbol{X} = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{p1} & \cdots & x_{pn} \end{pmatrix},\quad \boldsymbol{C} = \begin{pmatrix} c_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & c_p \end{pmatrix},\quad \boldsymbol{D} = \begin{pmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_n \end{pmatrix},$$

$$\boldsymbol{u}' = (u_1 \quad \cdots \quad u_p),\quad \boldsymbol{v}' = (v_1 \quad \cdots \quad v_n)$$

上の方程式で、左式に$\boldsymbol{u}'$を掛けると$\rho = 2\eta$、同様に右式に$\boldsymbol{v}'$を掛けると$\rho = 2\mu$を得る。右式を$\boldsymbol{v}$について解いて左式に代入すると以下となる。

$$\boldsymbol{C}^{-1}\boldsymbol{X}\boldsymbol{D}^{-1}\boldsymbol{X}'\boldsymbol{u} = \rho^2 \boldsymbol{u},\quad \text{また、}\boldsymbol{v} = \rho^{-1}\boldsymbol{D}^{-1}\boldsymbol{X}'\boldsymbol{u} \qquad (3.1)$$

また$\boldsymbol{v}$についても同様の関係が示されるが、ここでは省略する。

ここで$S_u^2 = 1$としたことから、$\boldsymbol{u}$の規格化条件を$\frac{1}{T}\boldsymbol{u}'\boldsymbol{C}\boldsymbol{u} = 1$として、新たに以下のベクトル$\boldsymbol{z}$を考える。

$$\boldsymbol{z} \equiv \frac{1}{\sqrt{T}}\boldsymbol{C}^{1/2}\boldsymbol{u},\quad \text{ここに}\ \boldsymbol{C}^{1/2} = \begin{pmatrix} \sqrt{c_{1\cdot}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sqrt{c_{p\cdot}} \end{pmatrix}$$

これを用いて最終的に以下となる。

$$\boldsymbol{Az} = \rho^2 \boldsymbol{z},\quad \boldsymbol{z}'\boldsymbol{z} = 1,\ \boldsymbol{A} \equiv \boldsymbol{C}^{-1/2}\boldsymbol{X}\boldsymbol{D}^{-1}\boldsymbol{X}'\boldsymbol{C}^{-1/2} \qquad (3.2)$$

異なる固有値ρ_α^2（$\alpha = 1, \cdots, p$）に対する固有ベクトルを$\boldsymbol{z}^\alpha$とすると、各点数は以下のように表される。

$$\boldsymbol{u}^\alpha = \sqrt{T}\boldsymbol{C}^{-1/2}\boldsymbol{z}^\alpha,\quad \boldsymbol{v}^\alpha = \rho_\alpha^2\sqrt{T}\boldsymbol{D}^{-1}\boldsymbol{X}'\boldsymbol{C}^{-1/2}\boldsymbol{z}^\alpha$$

ところで、(3.1) 式には $\rho^2 = 1,\ \boldsymbol{u} = \boldsymbol{1}$ の自明な解が存在し、それに基づく固有値と固有ベクトルが得られるが、この解は除外される。点数 $\boldsymbol{u},\ \boldsymbol{v}$ の与え方には、以下のように相関係数を掛ける方法もある。

$$\tilde{\boldsymbol{u}}^\alpha = \rho_\alpha \boldsymbol{u}^{\boldsymbol{\alpha}},\ \tilde{\boldsymbol{v}}^\alpha = \rho_\alpha \boldsymbol{v}^{\boldsymbol{\alpha}}$$

ここで$p \le n$を仮定してきたが、$p > n$の場合、先に$\boldsymbol{v}$について求め、後で$\boldsymbol{u}$について求めるが、方法は同様であるので省略する。

このカテゴリウェイト$\boldsymbol{u}^\alpha$と個体ウェイト$\boldsymbol{v}^\alpha$を用いてカテゴリ得点$\boldsymbol{y}^\alpha$と個体得点$\boldsymbol{w}^\alpha$をそれぞれ以下のように定義する場合もあるが、ここでは省略する。

$$\boldsymbol{y}^\alpha = \boldsymbol{X}\boldsymbol{v}^\alpha,\ \boldsymbol{w}^\alpha = \boldsymbol{X}'\boldsymbol{u}^\alpha$$

各成分の重要性を表すために、自明な解に対する固有値をρ_p^2として、以下で与えられる寄与率λ_αを考える場合もある。

$$\lambda_\alpha = \rho_\alpha^2 / \sum_{\beta=1}^{p-1} \rho_\beta^2 \qquad (\alpha \neq p)$$

3.2 プログラムの動作

メインメニューの中の「分析-多変量解析-数量化Ⅲ類」メニューを選択すると図 3.1 に示される分析メニューが表示される。分析は図 3.2 のような {0, 1} の値を持つデータから実行される。

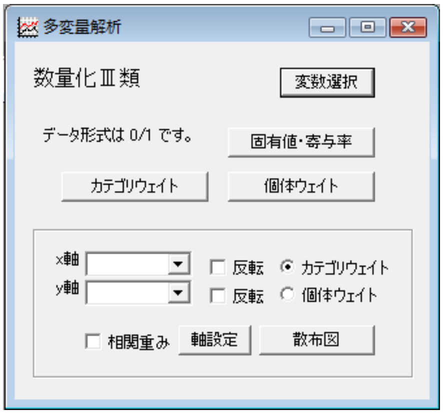


図 3.1 分析メニュー

データ編集 数量化Ⅲ類1.txt

	ご飯	パン	うどん	そば	ラーメン	スパゲッティ
1	1	0	1	1	1	0
2	1	0	1	0	0	0
3	0	1	0	0	1	1
4	1	1	1	1	0	1
5	0	1	0	1	1	1
6	1	0	1	1	1	0
7	1	0	0	0	0	0
8	1	1	1	1	0	1
9	0	1	0	1	1	1
10	1	0	1	0	1	0
11	1	1	1	0	1	1
12	1	0	1	0	1	1
13	1	1	0	0	1	1
14	0	1	0	1	0	0
15	0	1	0	1	1	0

1/1 (1,1) 分析: 数量化Ⅲ類 備考:

図 3.2 分割表データ

変数を選択して、「固有値・寄与率」ボタンをクリックすると図 3.3 のような結果が表示される。

固有値・寄与率

	第1次元	第2次元	第3次元	第4次元	第5次元
固有値	0.3505	0.1556	0.0948	0.0605	0.0252
相関係数	0.5921	0.3945	0.3079	0.2459	0.1589
寄与率	0.5105	0.2267	0.1381	0.0880	0.0368
累積寄与率	0.5105	0.7371	0.8752	0.9632	1.0000

図 3.3 固有値・寄与率画面

ここで表示される固有値は、(3.2) 式のρ^2、相関係数は同じくρである。

図 3.1 の分析メニューで「カテゴリウェイト」ボタンをクリックすると図 3.4 のような結果が表示される。

カテゴリウェイト

	第1次元	第2次元	第3次元	第4次元	第5次元
ご飯	-1.3676	-0.0171	0.7543	1.3384	0.2630
パン	1.2003	-0.0615	0.8530	0.3157	-1.6177
うどん	-1.2744	0.4328	-0.2079	-1.7697	-0.7989
そば	0.8258	1.9285	-0.0892	-0.0449	1.1017
ラーメン	0.2012	-0.6217	-1.8909	0.5360	-0.1014
スパゲッティ	0.5563	-1.4935	0.7582	-0.8836	1.3151

図 3.4 カテゴリウェイト画面

ここでは自明な解に対応する結果は表示されていない。

分析メニューの「個体ウェイト」ボタンをクリックすると、図 3.5 の個体ウェイト画面が表示される。

個体ウェイト

	第1次元	第2次元	第3次元	第4次元	第5次元
1	-0.6819	1.0915	-1.1640	0.0608	0.7308
2	-2.2312	0.5268	0.8872	-0.8772	-1.6862
3	1.1022	-1.8392	-0.3028	-0.0432	-0.8476
4	-0.0201	0.4001	1.3434	-0.8493	0.3312
5	1.1754	-0.1573	-0.2995	-0.0780	1.0977
6	-0.6819	1.0915	-1.1640	0.0608	0.7308
7	-2.3099	-0.0435	2.4495	5.4433	1.6554
8	-0.0201	0.4001	1.3434	-0.8493	0.3312
9	1.1754	-0.1573	-0.2995	-0.0780	1.0977
10	-1.3742	-0.1741	-1.4554	0.1419	-1.3368
11	-0.2311	-0.8928	0.1732	-0.3768	-1.1830
12	-0.7957	-1.0770	-0.4760	-0.7920	1.0664
13	0.2492	-1.3903	0.3853	1.3285	-0.2218
14	1.7111	2.3661	1.2403	0.5508	-1.6237
15	1.2540	1.0521	-1.2200	1.0939	-1.2951

図 3.5 個体ウェイト画面

カテゴリウェイトや個体ウェイトを図で表示するには、まずどちらを表示するかをラジオボタンで選択し、「軸設定」ボタンをクリックして x 軸と y 軸の成分を選択する。その後、「散布図」ボタンをクリックすると図 3.6 や図 3.7 のような散布図が表示される。

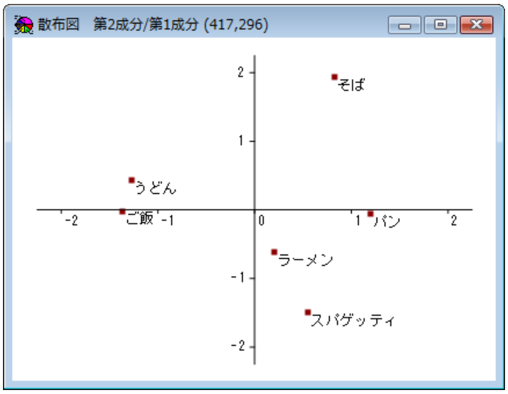


図 3.6 カテゴリウェイトの散布図

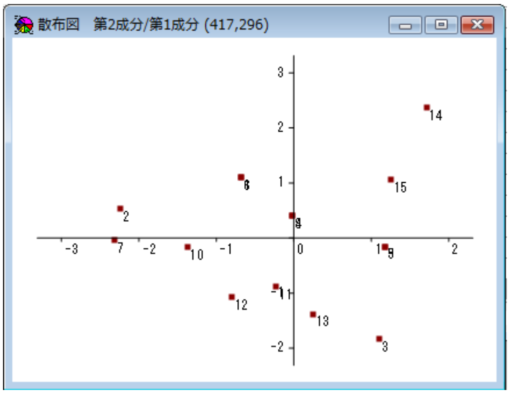


図 3.7 個体ウェイトの散布図

散布図の各成分には相関係数をかけて表示する場合があるが、その時には図 3.1 の「相関重み」チェックボックスにチェックを入れて散布図を表示する。また、成分を反転させて表示する場合は、反転チェックボックスにチェックを入れる。

4. 多変量解析に関する変更点

この章ではこれまで作成した多変量解析のプログラムで[2)]、その後機能追加を行った分析について、

その概略を説明する。

4.1 重回帰分析

重回帰分析については、新しく変数の自動選択機能を加えた。変数の追加と削除の基準は、追加と削除の変数の係数についての検定確率またはF検定値のどちらかで与えられる。「Pin」左側のラジオボックスをチェックすると検定確率で指定し「Fin」左側のラジオボックスをチェックするとF検定値で指定することになる。デフォルトは検定確率になっている。

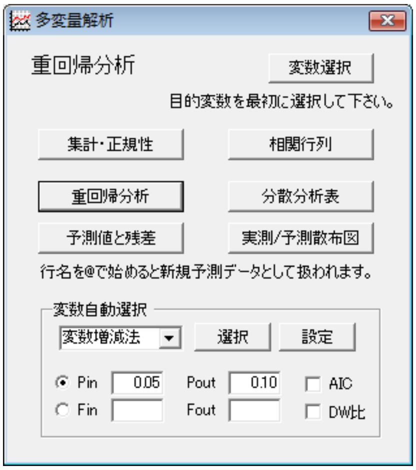
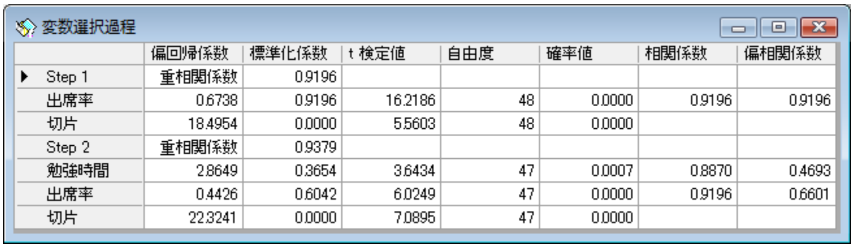


図 4.1 重回帰分析メニュー画面

変数の選択法として、変数増加法、変数減少法、変数増減法のどれかを選び、「選択」ボタンをクリックすると図 4.2 のように選択過程での種々の統計量が表示される。


変数選択過程

	偏回帰係数	標準化係数	t 検定値	自由度	確率値	相関係数	偏相関係数
Step 1	重相関係数	0.9196					
出席率	0.6738	0.9196	16.2186	48	0.0000	0.9196	0.9196
切片	18.4954	0.0000	5.5603	48	0.0000		
Step 2	重相関係数	0.9379					
勉強時間	2.8649	0.3654	3.6434	47	0.0007	0.8870	0.4693
出席率	0.4426	0.6042	6.0249	47	0.0000	0.9196	0.6601
切片	22.3241	0.0000	7.0895	47	0.0000		

図 4.2 変数選択過程表示画面

この場合は、２段階で変数が２つ選択されている。図 4.1 で「AIC」チェックボックスや「DW 比」チェックボックスにチェックを入れると、各過程での AIC の値やダービン・ワトソン比が図 4.2 の画面上に追加して表示される。

4.2 因子分析

因子分析では、因子負荷量推定法に主成分分析を加えた。これによって因子数を変数の数まで任意に選ぶことができるようになり、主成分分析の同じ主成分数の場合と累積寄与率が等しくなる。また、他の推定法に比べても累積寄与率の値は向上する。その他に、出力変数の並びをこれまでの変数選択順の他に、因子負荷量の大きさで２通りに並べ替える方法を加えた。これによって因子ごとに因子負荷量の大きい変数同士を並べて表示できるようになり、因子の解釈がより容易になる。また、因子ごとに主な変数をまとめた Cronbach の α 係数も表示するようにした。但し α 係数は、変数によって因子への寄与に正負の違いがある場合があるので、因子負荷量の符号により、変数の符号を変え、寄与を統一させて計算している。以上の機能を加えた分析メニューを図 4.3 に示す。

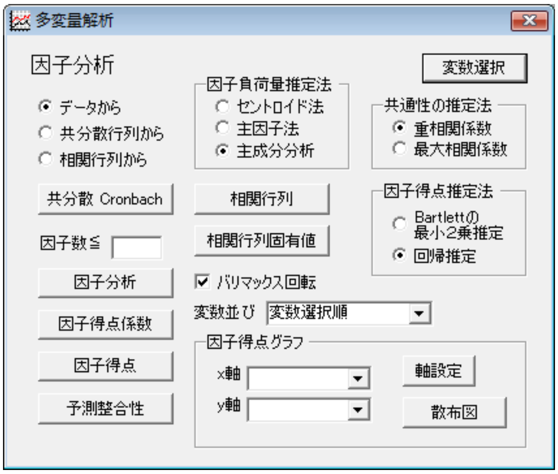


図 4.3 因子分析メニュー画面

4.3 クラスター分析

クラスター分析では、デンドログラムで表示するだけでなく、分類を表形式に表示する機能を加えた。

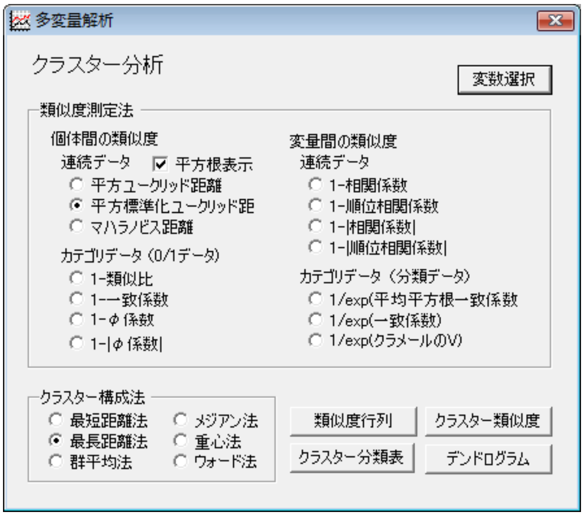


図 4.4 クラスター分析メニュー画面

図 4.4 のクラスター分析メニューで、「クラスター分類表」ボタンをクリックすると、図 4.5 のデンドログラムを表形式で表した図 4.6 のクラスター分類表が表示される。これはクラスター構成の各段階での分類を表示している。これによって例えば全体を 2 分割するときに各個体がどちらのクラスターに属するか簡単に知ることができる。また、これを利用して 2 つのクラスター間での有意差検定などを行いたい場合、この表の列をコピーして元データに加え、簡単に群分けすることができるようになる。

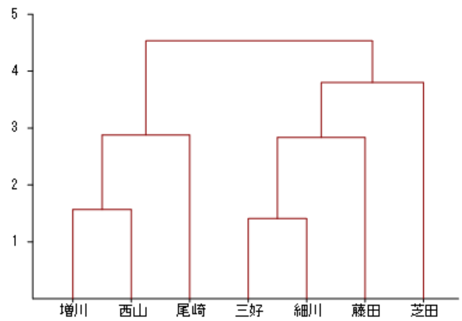


図 4.5 デンドログラム


クラスター分類表

	並び	7	6	5	4	3	2	1
増川	1	1	1	1	1	1	1	1
西山	2	2	2	1	1	1	1	1
三好	4	3	3	3	3	3	3	1
芝田	7	4	4	4	4	4	3	1
尾崎	3	5	5	5	5	1	1	1
藤田	6	6	6	6	3	3	3	1
細川	5	7	3	3	3	3	3	1

図 4.6 クラスター分類表

他の分析でも同様であるが、これまで予測値は欠損値データを除いて表示していたが、新しいデータを作成することを考えると欠損値を加えたままで表示し、元のデータに簡単に追加できるようにする方が賢明である。例えばこのクラスター分類表で、芝田のデータに欠損がある場合、図 4.7 の形式で表示すべきである。

クラスター分類表

	並び	6	5	4	3	2	1
増川	1	1	1	1	1	1	1
西山	2	2	2	1	1	1	1
三好	3	3	3	3	3	3	1
芝田							
尾崎	6	4	4	4	4	3	1
藤田	5	5	5	5	3	3	1
細川	4	6	3	3	3	3	1

図 4.7 欠損値のある場合の分類表の表示

この考えをすべての多変量解析に適用し、予測値には欠損値も加えて表示するように変更した。特に予測値の並びが変わった分析は、判別分析と数量化Ⅱ類である。これらは今まで群ごとに予測値を表示していたが、新たにデータ並びの順に表示するように作り変えた。

謝辞

この論文で作成したコレスポンデンス分析は、利用者からの要望で作成した。また、重回帰分析での変数選択についても、当初開発する予定はなかったが、要望があって追加した。また、因子分析での因子負荷量の変数並びもクラスター分析の分類表も同様である。これらの要望に答えることで、最近よく利用されている分析手法や、多変量解析と有意差検定などの連携を考えることができるようになった。心よりお礼を申し上げたい。

参考文献

1）Excel で学ぶコレスポンデンス分析, 高橋信, オーム社, 2005.

2）社会システム分析のための統合化プログラム７－多変量解析－, 福井正康, 細川光浩, 福山平成大学経営情報研究, 7 号, (2002) 85-106.

Multi-purpose Program for Social System Analysis 14

- Correspondence Analysis, Quantification Method type III -

Masayasu FUKUI

Department of Business Administration, Faculty of Business Administration,
Fukuyama Heisei University

Abstract

We have been constructing a unified program on the social system analysis for the purpose of education. This time we created new programs of correspondence analysis and quantification method, type III. These are classification methods of qualitative data. The latter program was already produced, but we remade it to fit the definition of former analysis.

Keywords

College Analysis, social system analysis, OR, statistics, correspondence analysis, quantification method type III

URL: http://www.heisei-u.ac.jp/ba/fukui/