Algorithm と呼ばれていて，1960年頃に得られているものである．

PAVA で知られていることを拡張すると次の結果となる．「初速度$(v_1, \ldots, v_N)$がランダムで球面的(spherical)な確率密度関数をもち質量が置換不変(exchangeable)ならば，任意の位置にたいして，衝突が有限時間で終わり，j個の質点から成るクラスタがc_j個$(c_j \geqq 0,\ \sum_{j=1}^{N} jc_j=N)$となる確率は$1\Big/\prod_{j=1}^{N} j^{c_j}c_j!$となる．極限$N\to\infty$で，$(c_j)_{j=1}^{\infty}$は独立な平均$1/j$のポアソン分布となる．」

この結果は質点を区別していない．第n粒子がj個の質点から成るクラスタは，速度の分布に依存して，計算は難しい．数値実験と$N=3, 4$の場合の計算によると，中央部分ほど大きなクラスタが生じやすい．

参 考 文 献

Sibuya, M., Kawai, T. and Shida, K. (1990). Equipartition of particles forming clusters by inelastic collisions, *Physica A*, **167**, 676-689.

統計データ解析センター

定量的・定性的データの成分分析による医学的分類応用

駒 澤 勉

今回はカテゴリカルデータについて成分分析による判別・予測分析の効用を心機能評価データで示した．

まずデータの総合化と分析法について，

〔1〕アイテム・カテゴリー・データの総合化

$$\langle\text{全体での総合量}\rangle = Y = X_1+X_2+\cdots+X_m=\sum_{j=1}^{m} X_j$$

$$\langle\text{グループ } G_t \text{ での総合量}\rangle = Y^{(t)}=X_1^{(t)}+X_2^{(t)}+\cdots+X_m^{(t)}=\sum_{j=1}^{m} X_j^{(t)} \qquad (t=1, 2, \ldots, T)$$

$$X_j=\sum_{k=1}^{l_j} \delta_{(jk)}\, x_{(jk)}$$

$$X_j^{(t)}=\sum_{k=1}^{l_j} \delta_{(jk)}^{(t)}\, x_{(jk)}$$

数量$x_{(jk)}$がアイテム・カテゴリー$C_{(jk)}$に与える数量である．

ここで，

$$\delta_{(jk)}=\begin{cases}1\,;\ C_{(jk)}\text{ に該当}\\ 0\,;\ C_{(jk)}\text{ に非該当}\end{cases}$$

$$\delta_{(jk)}^{(t)}=\begin{cases}1\,;\ \text{グループ } G_t,\ \text{かつ } C_{(jk)}\text{ に該当}\\ 0\,;\ \text{グループ } G_t,\ \text{かつ } C_{(jk)}\text{ に非該当.}\end{cases}$$

〔2〕判別・予測のための$C_{(jk)}$への数量化方法

(i) 数量化II類 … 相関比η^2の最大化による数量化$HX=\eta^2FX$（医学データのように説明アイテム間(X_j, X_k)に多くの従属関係があるときは利用できない.）

(ii) 数量化III類 … 相関係数r_{XY}の最大化による数量化$FX=r_{XY}^2GX$

(iii) 主成分数量化 … 分散σ_Y^2の最大化による数量化$FX=\sigma_Y^2CX$(説明アイテム間(X_j, X_k)に従属

関係があるときの対策法として，主成分判別分析的な考え方で判別・予測を行う.）

ここで，

$$
\begin{aligned}
&F=\{f_{(jk)(uv)}\}\,;\ f_{(jk)(uv)}=n_{(jk)(uv)}-\frac{n_{(jk)}n_{(uv)}}{n}\\
&H=\{h_{(jk)(uv)}\}\,;\ h_{(jk)(uv)}=\sum_{t=1}^{T}\frac{n_{(jk)}^{(t)}n_{(uv)}^{(t)}}{n_t}-\frac{n_{(jk)}n_{(uv)}}{n}\\
&G=\{g_{(jk)(uv)}\}\,;\ g_{(jk)(uv)}=\begin{cases} m\cdot n_{(jk)}\,; & j=u,\ k=v\\ 0 & ;\ \text{others}\end{cases}\\
&C=\{c_{(jk)(uv)}\}\,;\ c_{(jk)(uv)}=\begin{cases} n\,;\ j=u,\ k=v\\ 0\,;\ \text{others}\end{cases}\\
&X=\{x_{(jk)}\} \qquad\qquad \begin{pmatrix} j,\ u=1,2,\ldots,m\\ k=1,2,\ldots,l_j\\ v=1,2,\ldots,l_u\end{pmatrix}.
\end{aligned}
$$

なお，

$$
n_{(jk)}=\sum_{i=1}^{n}\delta_{i(jk)},\quad n_{(jk)}^{(t)}=\sum_{i=1}^{n}\delta_{i(jk)}^{(t)},\quad n_{(jk)(uv)}=\sum_{i=1}^{n}\delta_{i(jk)}\delta_{i(uv)}.
$$

外的基準のある方法 (i) と外的基準のない方法 (ii)，(iii) を同じ説明アイテムの総合化モデル式で解析したとき，判別・予測の評価尺度の判別成功率より，アイテム・カテゴリーに得られた数量 $x_{(jk)}$ で分析した方が有用性・妥当性が得られる．このことを説明アイテム間に従属関係が多くある医学データで示した．

男女の生まれかわり～日本人の国民性調査から

中 村 隆

「日本人の国民性調査」は第I次調査が1953年（昭和28年）に実施されて以来，5年ごとに継続調査され，1988年には第VIII次調査が行なわれた．現在，第VIII次調査までの結果をまとめた報告書を関係委員とともに執筆中である．ここでは，#6.2‘男・女の生まれかわり’という質問項目を分析した結果を，継続的調査データから年齢・時代・コウホートの3効果を分離するコウホート分析を適用した結果を含めて報告する．

質問文は次のようなものである．

#6.2‘男・女の生まれかわり’

もういちど生まれかわるとしたら，あなたは男と女の，どちらに，生まれてきたいと思いますか？
1 男に　2 女に　3 その他［記入］　4 D.K.

各カテゴリの回答比率を調査時点・男女別に示すと図1のようになる．これをみると，男と女の回答比率の時代的推移は好対照であることがわかる．男の‘男に生まれてきたい’という比率が1958年以来ほぼ90%で変わらないのに対して，女の‘男に生まれてきたい’という比率は，1958年に64%だったものが68年まで5年ごとにほぼ10ポイントずつ減少し，その後70年代には微減傾向だったが，80年代に入って再び減少傾向が加速して，1988年には34%となっているのである．男性が“変わろうとしない”のに対して，女性は確実に変わってきている．

このような女全体についての意見の変化は，すべての年齢層あるいは世代の意見が同じ方向に変わったことによってもたらされたのであろうか，それとも違う意見をもった若い世代が入ってくることに