⑦ 数量化の一問題

青山　博次郎

§1. 緒　　言

数量化の問題としてよく取扱う問題は一つの item と他の幾つかの item との関係を，何れか一方より，或は同時に数量化して予測に役立てようとする問題である。

これらは統一的には重相関係数を最大とする方法によつて論ずることが出来る。

それらの方法は，林知己夫，石田正次両氏によつて取扱われている。[*1]

こゝでは直接的方法を以て凡ての場合について考察してみることにする。

§2. 幾つかの item より一つの item を数量化する場合

今 m 組の item $x^{(i)}$ $(i=1,2,\cdots\cdots,m)$ があつて，これは既に数量で与えられているものとする。

更に一つの item y は定性的なものであつて，これを $x^{(i)}$ の値から，最も妥当な方法で数量化しようという問題である。

例えば各学校の算数の評価（これは5段階法で表わされているものとする）と，ある共通のテストを幾種類か行つた結果より，

逆にこのテストの点数を用いて算数の評価を各校共通の立場から数量化するという如き場合がこれに当る。*2

さて y, $x^{(i)}$ は凡て k_i 個の category に分れていて，次の表の如くなつているものとする。

$x_\rho^{(i)}$ （ $i = 1, 2, \cdots\cdots, m$; $\rho = 1, 2, \cdots\cdots, k_i$ ）を與えられたとき

$$y = w_1 x^{(1)} + w_2 x^{(2)} + \cdots\cdots + w_m x^{(m)} \qquad (1)$$

によって推定し，このときの重相関係数 $r^2_{0\cdot 12\cdots\cdots m} = 1 - \frac{R}{R_{00}}$ を最大になるように y をえらべばよい。

$y \backslash x^{(i)}$	$x_1^{(i)}$	$x_2^{(i)}$	---	$x_\rho^{(i)}$	---	$x_{k_i}^{(i)}$	計
y_1	$f_{11}^{(i)}$	$f_{21}^{(i)}$	---	$f_{\rho 1}^{(i)}$	---	$f_{k_i 1}^{(i)}$	f_1
⋮			----				⋮
y_s	$f_{1s}^{(i)}$	$f_{2s}^{(2)}$	----	$f_{\rho s}^{(i)}$	----	$f_{k_i s}^{(i)}$	f_s
⋮			------				⋮
y_t	$f_{1t}^{(i)}$	$f_{2t}^{(i)}$	----	$f_{\rho t}^{(i)}$	----	$f_{k_i t}^{(i)}$	f_t
計	$f_1^{(i)}$	$f_2^{(i)}$	----	$f_\rho^{(i)}$	---	$f_{k_i}^{(i)}$	n

このとき

$$w_i = - \frac{\sigma_0 R_{0i}}{\sigma_i R_{00}} \qquad (2)$$

$$R = \begin{vmatrix} 1 & r_{01} & r_{02} & \cdots & r_{0m} \\ r_{10} & 1 & r_{12} & \cdots & r_{1m} \\ \vdots & & \ddots & & \vdots \\ r_{m0} & r_{m1} & \cdots & & 1 \end{vmatrix} \tag{3}$$

r_{ij} は $x^{(i)}$ と $x^{(j)}$ の相関係数，r_{0j} は y と $x^{(j)}$ の相関係数，R_{0i} は行列式 R における r_{0i} の余因数．σ_0^2 は y の分散，σ_i^2 は $x^{(i)}$ の分散を示す。

$r^2_{0.12\cdots m}$ の中で R_{00} は y_s を含んでいないので重相関係数を最大にするためには，R を最小にすればよいことが分る。

故に

$$\frac{\partial R}{\partial y_s} = \sum_{i=1}^{m} \frac{\partial R}{\partial r_{0i}} \frac{\partial r_{0i}}{\partial y_s} = 0 \tag{4}$$

より y_s を解けばよい。

ここで $\bar{y} = 0$ と仮定しても一般性を失わないから

$$\frac{\partial R}{\partial r_{0i}} = 2R_{0i} \tag{5}$$

$$r_{0i} = \frac{\frac{1}{n} \sum_s \sum_p f_{ps}^{(i)} x_p^{(i)} y_s}{\sigma_0 \sigma_i} \tag{6}$$

$$\sigma_0^2 = \frac{1}{n} \sum_s f_s y_s^2 \tag{7}$$

$$\sigma_i^2 = \frac{1}{n} \sum_p \sum_s f_{ps}^{(i)} x_p^{(i)2} - \bar{x}^{(i)2} \tag{8}$$

に注意して，(4) は

$$\frac{f_s y_s}{\sum_i \sum_\rho \frac{R_{0i}}{\sigma_i} f_{\rho s}^{(i)} x_\rho^{(i)}} = \frac{\sigma_0^2}{\sum_i^m \frac{R_{0i}}{\sigma_i} \mathrm{cov}(x^{(i)}, y)} = \lambda \tag{9}$$

ここで λ は s には無関係な量で容易に分る如く

$$\lambda = \frac{\sigma_0}{R - R_{00}} \tag{10}$$

となる。

(9) 式に於て第一辺の R_{0i} を展開すれば，R_{00} の r_{ij} の余因数を Δ_{ij} とおくとき

$$f_s y_s = -\lambda \sum_{j=1}^m r_{j0} \sum_i \sum_\rho \frac{\Delta_{ij}}{\sigma_i} f_{\rho s}^{(i)} x_\rho^{(i)}$$

$$= -\frac{\lambda}{\sigma_0} \sum_j^m \sum_i^m \sum_\rho^{k_i} \sum_{\rho'}^{k_j} \sum_{s'}^{t} \frac{\Delta_{ij}}{\sigma_i \sigma_j n} f_{\rho s}^{(i)} x_\rho^{(i)} f_{\rho' s'}^{(j)} x_{\rho'}^{(j)} y_{s'} \tag{11}$$

ここで

$$\frac{1}{n} \sum_i \sum_j \sum_\rho \sum_{\rho'} \frac{\Delta_{ij}}{\sigma_i \sigma_j} f_{\rho s}^{(i)} x_\rho^{(i)} f_{\rho' s'}^{(j)} x_{\rho'}^{(j)} = H_{ss'} \tag{12}$$

$$\frac{1}{\mu} = -\frac{\lambda}{\sigma_0} \tag{13}$$

とおくと，(11) は

$$\sum_{s'=1}^t H_{ss'} y_{s'} = \mu f_s y_s \tag{14}$$

となる． この式で $H_{ss'} = H_{s's}$ なることはすぐ分る。

(14) を解いて y_s を求めればよい。

このとき (10) より

$$r_{0.12\cdots\cdots m} = \sqrt{\frac{\mu}{R_{00}}} \tag{15}$$

なることが分る。

例

	$y_1=0$	$y_2=1$	$y_3=2$	計
$x_1=0$	1	2		3
$x_2=1$		4		4
$x_3=2$		2	1	3
計	1	8	1	10

	x_1	x_2	x_3	計
Z_1	2	2		4
Z_2	1	2	1	4
Z_3			2	2
計	3	4	3	10

	y_1	y_2	y_3	計
Z_1	1	3		4
Z_2		3	1	4
Z_3		2		2
計	1	8	1	10

上の如き分布表が與えられたとき，(14) 式より Z_i を求めよう。

$$\sigma_1^2 = 0.6,\quad \sigma_2^2 = 0.2,\quad r_{12} = \frac{\sqrt{3}}{3},\quad R_{00} = \frac{2}{3},\quad \Delta_{12} = -\frac{\sqrt{3}}{3},$$

$$\Delta_{11} = \Delta_{22} = 1$$

従って

$$\begin{cases} 4Z_1 = \dfrac{1}{30\mu}(95Z_1 + 155Z_2 + 50Z_3) \\[2ex] 4Z_2 = \dfrac{1}{30\mu}(155Z_1 + 255Z_2 + 90Z_3) \\[2ex] 2Z_2 = \dfrac{1}{30\mu}(50Z_1 + 90Z_2 + 60Z_3) \end{cases}$$

$120\mu = \xi$ とおくとき，上式より係数のつくる行列式を 0 とおいて

$$\xi^3 - 470\xi^2 + 21000\xi = 0$$

$$\therefore\quad \xi = 0,\quad 50,\quad 420$$

この根の中で，$\bar{Z} = 0$ となるものは $\xi = 50$ のときであって

$$Z_1 = -k,\qquad Z_2 = -k,\qquad Z_3 = 4k$$ (kは比例定数)

となる。

またこのときの重相関係数 ρ は

$$\rho = \sqrt{\frac{\mu}{R_{00}}} = \sqrt{\frac{5}{8}} = 0.7906$$

となる。

§ 3. 幾つかの item を一つの item (outer criterion) に対して数量化する場合

この場合が林知己夫氏（数値計算には至っていないが，後に，石田氏と同一結果が得られている。），石田正次氏（最小二乗法を利用）によって論じられている場合である。

我々の直接法を用いると次のようになる。

一般性を失うことなく $\bar{x}^{(i)} = 0$ と仮定して差支えない。

このとき

$$r_{0i} = \frac{\frac{1}{n}\sum_{s}\sum_{p} f_{ps}^{(i)} x_p^{(i)} y_s}{\sigma_0 \sigma_i} \tag{16}$$

$$r_{ij} = \frac{\frac{1}{n}\sum_{p}\sum_{u} f_{pu}^{(ij)} x_p^{(i)} x_u^{(j)}}{\sigma_i \sigma_j} \tag{17}$$

である。

重相関係数を最大にするには

$$\frac{\partial R}{\partial x_p^{(i)}} = \frac{R}{R_{00}} \frac{\partial R_{00}}{\partial x_p^{(i)}} \tag{18}$$

となる如く $x_p^{(i)}$ を定めればよい。

然るに

$$\frac{\partial r_{oi}}{\partial x_p^{(i)}} = \frac{1}{\sigma_o \sigma_i^3}\left\{\frac{\sigma_i^2}{n}\sum_s f_{ps}^{(i)} y_s - \frac{f_p^{(i)} x_p^{(i)}}{n} cov(x^{(i)}, y)\right\} \tag{19}$$

$$\frac{\partial r_{ij}}{\partial x_p^{(i)}} = \frac{1}{\sigma_j \sigma_i^3}\left\{\frac{\sigma_i^2}{n}\sum_u f_{pu}^{(ij)} x_u^{(j)} - \frac{f_p^{(i)} x_p^{(i)}}{n} cov(x^{(i)}, x^{(j)})\right\} \tag{20}$$

$$\therefore \frac{\partial R}{\partial x_p^{(i)}} = \frac{\partial R}{\partial r_{oi}} \frac{\partial r_{oi}}{\partial x_p^{(i)}} + \sum_{j \neq i} \frac{\partial R}{\partial r_{ij}} \frac{\partial r_{ij}}{\partial x_p^{(i)}}.$$

$$\frac{\partial R_{oo}}{\partial x_p^{(i)}} = \sum_j \frac{\partial R_{oo}}{\partial r_{ij}} \frac{\partial r_{ij}}{\partial x_p^{(i)}}$$

に代入すると（18）より重相関係数を ρ と略記して

$$\frac{R_{oi}}{\sigma_o}\sum_s f_{ps}^{(i)} y_s + \sum_{j \neq i} \frac{R_{ij}}{\sigma_j}\sum_u f_{pu}^{(ij)} x_u^{(j)} - (1-\rho^2)\sum_{j \neq i}\frac{\Delta_{ij}}{\sigma_j}\sum_u f_{pu}^{(ij)} x_u^{(j)}$$

$$= \frac{f_p^{(i)} x_p^{(i)}}{\sigma_i R_{oo}} (R\Delta_{ii} - R_{ii} R_{oo}) \tag{21}$$

我々の $x_p^{(i)}$ は定数をかけておいても ρ には影響がないから

$$x_u^{(j)} = -\frac{\sigma_j R_{oo}}{\sigma_o R_{oj}} Z_u^{(j)} \tag{22}$$

とおくと，（21）式により

$$\frac{R_{ii} R_{oo} - R\Delta_{ii}}{R_{oi}^2} f_p^{(i)} Z_p^{(i)} = \sum_{j \neq i}^{m} \sum_{n=1}^{k_j} \frac{R\Delta_{ij} - R_{ij} R_{oo}}{R_{oi} R_{oj}} f_{pu}^{(ij)} Z_u^{(j)} + \sum_s f_{ps}^{(i)} y_s \tag{23}$$

然るに R は対称行列式だから

$$\left.\begin{aligned} R_{ii} R_{oo} - R\Delta_{ii} &= R_{oi}^2 \\ R_{ij} R_{oo} - R\Delta_{ij} &= R_{oi} R_{oj} \end{aligned}\right\} \tag{24}$$

なる関係があり，結局 $Z_p^{(i)}$ を求める式は

$$f_p^{(i)} Z_p^{(i)} + \sum_{j \neq i} \sum_u f_{pu}^{(ij)} Z_u^{(j)} = \sum_s f_{ps}^{(i)} y_s \qquad (25)$$

となる。

このとき ρ の値は $\sqrt{1-\frac{R}{R_{00}}}$ によつて計算すればよい。（多重分類がしてあつたり，item や Category の数の少いときは他の便法もある。*1 ）

例．　§2の例で $Z_1=0$, $Z_2=1$, $Z_3=2$ として x_i, y_j を求めると次のようになる。

$$\begin{cases} 3x_1 + y_1 + 2y_2 = 1 \\ 4x_2 + 4y_2 = 2 \\ 3x_3 + 2y_2 + y_3 = 5 \\ y_1 + x_1 = 0 \\ 8y_2 + 2x_1 + 4x_2 + 2x_3 = 7 \\ y_3 + x_3 = 1 \end{cases}$$

これより

$$\begin{cases} x_1 = x_2 = -\frac{1}{2} - k \\ x_3 = 1 - k \\ y_1 = \frac{1}{2} + k \\ y_2 = 1 + k \\ y_3 = k \end{cases}$$

$k=1$ とすると $x_1 = x_2 = -1.5$; $x_3 = 0$, $y_1 = 1.5$, $y_2 = 2$, $y_3 = 1$

このとき

$$\rho = 0.8556$$

となる。

§ 4. 両方の item を同時に数量化する場合.

$i=1$ のときは既に発表した。[*3]　　$i>1$ の場合は trivial な解以外には確定的なものはない。

方法は§3，§4を結合したもので

$$f_\rho^{(i)} z_\rho^{(i)} + \sum_{j\neq i}\sum_u f_{\rho u}^{(ij)} z_u^{(j)} = \sum_s f_{\rho s}^{(i)} y_s \tag{26}$$

$$\mu f_s y_s = \sum_{i'}\sum_{\rho'} f_{\rho' s}^{(i')} z_{\rho'}^{(i')} \tag{27}$$

から得られる。

(27) を (26) に代入すると i, ρ に関して

$$z_\rho^{(i)}\left(f_\rho^{(i)} - \frac{1}{\mu}\sum_s \frac{f_{\rho s}^{(i)^2}}{f_s}\right) + \sum_{j\neq i}\sum_u f_{\rho u}^{(ij)} z_u^{(j)} - \frac{1}{\mu}\sum_s \frac{f_{\rho s}^{(i)}}{f_s}\sum_{i'}\sum_{\substack{\rho' \\ (i',\rho')\neq(i,\rho)}} f_{\rho' s}^{(i')} z_{\rho'}^{(i')} = 0 \tag{28}$$

が得られる。　この式も係数の行列は対称であるがその値は0である。従ってどれか一つの $z_\rho^{(i)}=k$ とおいて解くと $z_\rho^{(i)} \equiv const.$ ($\rho=1,2,\cdots,k_i$) が得られるに過ぎない。この関係は前例を用いると具体的に看取できる。

例，　§2.の例で，x_i, y_j, z_k が共に数量化されるとき(28)式は

$$3x_1+y_1+2y_2 = \frac{1}{4\mu}(5x_1+6x_2+x_3+2y_1+9y_2+y_3)$$

$$4x_2+4y_2 = \frac{1}{4\mu}(6x_1+8x_2+2x_3+2y_1+12y_2+2y_3)$$

$$3x_3+2y_2+y_3 = \frac{1}{4\mu}(x_1+2x_2+9x_3+11y_2+y_3)$$

$$y_1+x_1 = \frac{1}{4\mu}(2x_1+2x_2+y_1+3y_2)$$

$$8y_2+2x_1+4x_2+2x_3 = \frac{1}{4\mu}(9x_1+12x_2+11x_3+3y_1+26y_2+3y_3)$$

$$y_3+x_3 = \frac{1}{4\mu}(x_1+2x_2+x_3+3y_2+y_3)$$

$4\mu=\xi$ とおいて $(x_1\ x_2\ x_3\ y_1\ y_2\ y_3)$ の係数の行列式を Ξ とおけば

$$\Xi = 0$$

従って $y_3=k$ とおいて最初の五式より $x_1, \cdots, y_2$ を解けば

$$x_1 = x_2 = x_3 = -k, \quad y_1 = y_2 = y_3 = k$$

となる。 この場合は category が凡て消失して了う。

そこで $y_2=k_1$, $y_3=k_2$ とおいて最初の四式より $x_1, \cdots y_1$ を求めると

$$\frac{k_1}{1-3\xi+\xi^2} = K_1, \quad \frac{k_2}{1-3\xi+\xi^2} = K_2$$

とおきかえるとき

$$\begin{cases} x_1 = x_2 = -(3\xi^2-8\xi+2)K_1+(\xi-1)K_2 \\ x_3 = -(2\xi^2-8\xi+3)K_1-\xi(\xi-1)K_2 \\ y_1 = (3\xi^2-8\xi-1)K_1-(\xi-4)K_2 \\ y_2 = K_1 \\ y_3 = K_2 \end{cases}$$

とおくことが出來る。従って (28) より

$$\begin{cases} z_1 = -(9\xi-24+\frac{6}{\xi})K_1+3K_2 \\ z_2 = -(11\xi-32+\frac{6}{\xi})K_1-(\xi-4+\frac{2}{\xi})K_2 \\ z_3 = -8(\xi-4+\frac{1}{\xi})K_1+4(1-\xi)K_2 \end{cases}$$

K_1, K_2, ξ は任意にとることが出來るが，重相関係数 ρ を定めたとき K_1, K_2, ξ は一定の関係をもって幾組もの解が得られる。従つて他に條件を加えない限りどのようにでも x_i, y_j, z_k がえらべる。 例えば $K_1=0$, $K_2=1$ のとき $\xi=1$ ならば $\rho=0.3272$, $\xi=2$ ならば $\rho=0.5557$, $\xi=3$ ならば $\rho=0.8434$ 等々が得られる。

註 ※1. 林知己夫，石田正次：数量化の問題について，統数研附屬養成所講義 (1951)

※2. 青山博次郎，西平重喜：質問形式についての研究．統計数理研究輯報 No.9，1952.

※3. 同：全国小中学校教育課程調査について（その二） 統計数理研究輯報 No.5，1951.