第1部

統計数理研究輯報

第 8 號

數量化における計算法について I

昭和 27 年 2 月

統計數理研究所

東京都世田ケ谷區三軒茶屋町10

この輯報は実際問題について準備の段階から計画，実施，処理に到る間に必要な統計数理的考え方，技術を述べたものである。ねらいは実際に役に立つ報告ということである。

之は其の性質からいつて，統計数理の研究者だけでなく，調査，分析等広くこのような実証的な仕事にたずさわる人々の參考となるようにと願つて刊行するものである。

發行所	東京都世田谷區三軒茶屋町十 統計數理研究所
編集責任者	林　知己夫
印刷所	東京都文京區高田豐川町十三 駐文社印刷所 古田義雄

数量化における計算法について

I

前　書　き

これは定性的なものをある立場から数量化する問題，即ち所謂数量化問題を取扱う際に現はれてくる計算の実例について記述し実際に数量化を行う人々のための参考に資するものである。

勿論数量化の方法は色々と考へ得るであらうが，最小自乗法の助けを借りる場合が多いので結局計算としては，

1). 連立一次方程式の数値解

2). 條件つきの最大最小問題

（ラグランジュの係数）

3). 特性根の数値解

を度々行うことになるために，ここに記したものも多少は利用出来ると思はれる。

数値計算に当つては桁数をどこまでとるかが問題である。最後の結果にどれほどの精度が必要かと云うことを考へ，計算の手順の多少によつて二桁から四桁余計にとつて計算すればまず大抵大丈夫である。　残念なことには数値計算の誤差を論ずるもので実際に充分利用出来るものは見当らない。　又数量化の計算では，近似計算，逐次近似法が便利な場合，又，それでなければ計算が出来ない場合がある。　その時ははじめ桁数を少なく（二，三桁）して計算を行ひグラフを数多く作つて，早く，計算の大体の結

果，見通しを握み，だんだん桁数をまして，必要の精度まで上げて行くことが肝要である。

数表と参考書

Barlow の数表

Peter の計算表

Gauss の対数表

Fisher の統計数値表

Jahnke, Emde の函数表

丸山文行　計算法（統計数理研究所統計技術員養成所）

Frazer, Duncan, Collar　Elementary Matrices

柴垣和三雄　実用解析

目　　次

序

ある立場から調査された定性的（質的）なものを数量の形としてとらへる事の極めて大切であることはあらためて言ふ迄もない。

数として取扱ふことによつて，複雑であるものを単純化して取扱ふことが出来，かくして勘案と言ふプロセスを合理的且つ能率的に行ふことが可能となつてくる。 ここで問題になるのは単純化と言ふことである。

唯任意に単純化して取扱ふのであつては，全く無意味であつて科学性がないと言はねばならない。 如何なる立場から科学的に単純化して行くかが問題になるのである。

定性的なものをあらゆる角度から，実態概念的に数量化することはまづ不可能と言へるであらう。 一つの明皙な立場，しかも理論的な面ばかりでなく，きわめて具体的な面を明らかにした立場，即ち我々が一つのある行為をおこさうとする，それに有用な指針を得たいとするやうな立場から数量化は行はれねばならない。

ねらた一つの面に着目し，それをもつとも能率よく示す様にある立場から調査された定性的な性質（通常これは *item*, *category*, 型の形であらはされる）を単純化して数量化してゆくのである。

ねらつた面も異なれば，あたへられる数量もおのづから異ってくるのである。

定性的な豊富な面をもつ性格のすべてがそれなりに一つの目標に集約され，その立場から他のことを捨象して単純化され数量化されてくると言へるのである。 もしもこの結果が現実の目的にたいして有効な知識とならないといふ点で香ばしくない時には調査された定性的のものが工合悪いのであつてあらためて調査しなほさなければならない。 数量化はこの資料からせいぜいの所，つまり最大の利益ある知識をあたへようとするものであると言へる。

以上の様な考へ方は機械的で浅薄であると言ふのは当らない。

我々が一つの行為をなさうとする（行為そのものは思考とことなり一つの方向をもつた単純なものである。複雑な行為は有限の肉体をもつ我々にとつては不可能である）のに用ふる知識はきわめて単純なものである。　もしも知識が複雑なままであれば勘案のプロセスが入り，そのまゝでは行為の指針とはならないのである。　この場合は勘案のプロセスにより単純化して得た知識が（これは単純なものである！）始めて我々の指針となるのである。

このやうな過程を科学的に *optimum*（目的に対してプラグマティックの意味で *best*）と言ふ考へから機能的操作的に取扱ふのが数量化の核心問題なのである。　あることをなすのに行ふ通常の我々の思考は，ふくざつな性質を"行為のための知識"と言ふ一つの次元の上にのせる事にある事を思へば当然首肯し得られる所であらうと思ふ。

ここで数量化の時 *Optimum* と言ふことを述べたが，これは統計数理では重要な考へ方である。　数量化されて得られた知識が実際にきわめて（最も）有用であり，妥当性のあると言ふこと同時に又説得性のあること，理論的にもつとも妥当性あるものであること，これらが *Optimum* なものと言はれてよからう。

それではある立場から調査された定性的なものが勘案のプロセスを通じてなされることがらの数量化が具体的にいかに行はるべきであらうか，これらの事は単なる抽象や思辯では片がつかぬ問題であり，個々の具体的な事象につき統計数理的に考へられてゆかねばならない。

これらの実際については邦文では（外国文は下記のものの中に参考文献としてあげられてゐる）

林　知己夫　　統計数理的数量化の問題及び補遺

（統計数理研究所，講究録第6巻第1,2,3,11号）

林　知己夫　　予想の的中につき（講究録第７巻第３号）

青山博次郎
松下嘉米男
林　知己夫
水野　坦
　　社会現象の統計数量（朝倉書店）

林　知己夫　　社会計量における統計数理（日本数学会，数学，第３巻第３号）

等の中に記述されてある。

さて数量化を実際に行ふには普通相当の数値計算と言ふものが必要となつてくる。　これを除外しては数量化は考へられないのである。　しかもこの計算はきまりきつたものは少くそのプロセスも簡単に行かぬものが多く，多くの工夫技巧と努力とを必要とするものである。　計算は，実際的には近似計算によるものも多いのである。　新しい数量化の理論が出来れば，新しい計算法を工夫しなければならない。　一律の公式にしたがへぬ所が面白い所でもあり，むづかしい所でもある。

数量は與へられてあるものではなく，我々のある目的に対して我々が與へるものである事を了解しなければならない。　即ち目的に対して行為の立場からみて有用な知識をあたへる機能的，操作的のものであらねばならない。

§２　以下に於て前記論文にしめしてある数量化の方法に関してなされた計算法の工夫，計算の手順が書かれてある。

ここにしめしてある数量化の方法に用ひられるのみでなく，一般に同種問題の計算法の参考にならうかと思ふ。

前述の様に数量化には複雑な計算を必要とするものが多いが，米国においては計算機構が発達してゐるため，（或はむしろこの種研究が計算機構の発達を促したとも言へる），これを用ひて数量化の計算が行へる。

したがつて理論も，計算法の面倒さをさう顧慮せずにつくれるものと思はれる。 しかし我が国においては，機構がなく手による計算によらねばならぬ実状であるから，理論も"手計算で可能な範囲内の計算"で片のつくやうなものを工夫しなければならず，（しかも Optimum な立場を放棄せずに！）ここに新な数量化研究の方向が進められねばならない。

§ 2. 重相関係数と回帰平面

これは直接に数量化とは関係はないが，その2が回帰平面の一つの変形であり，又統計特に数量化に於ては行列式の計算，連立方程式の解法が屡々必要になつてくるので，こゝでは連立方程式の解法を中心として，重相関係数，回帰平面を実例から述べてみる。

よく知られてゐるやうに重相関係数回帰平面の計算は各単相関係数を計算し，それからつくられる行列式を計算することに帰着する。 然しこの行列式を計算することは，行列の次数が小なる時は行列式の運算で簡単にもとまるが，次数が4以上にもなると，行列式そのままの計算は楽ではない。 これをもとめるにはむしろ一次聯立方程式にたちもどり，逐次的に解き根をもとめ逆にこれから行列式の値をもとめる方が検算の点からも計算の手続きの点からもよい。 したがつてここではこの方法によつて，重相関係数を求める方法を示してみよう。

即ち， n 個のものから成る集団があり，それに属する各々は x_0, x_1, x_2, ‥‥‥, x_k の $k+1$ 個の性質を持つてゐるとき x_i と x_j の単相関係数 r_{ij} について，行列式を

$$\begin{vmatrix} 1 & r_{01} & r_{02} & \cdots & r_{0k} \\ r_{01} & 1 & r_{21} & \cdots & \vdots \\ r_{02} & r_{21} & 1 & & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{0k} & r_{k1} & & & 1 \end{vmatrix} = R$$

又その i 行 j 列の余因数を R_{ij} で表せば $x_1, \cdots\cdots, x_k$ から x_0 への回帰平面の方程式は

$$x_0 = -\frac{\sigma_0}{\sigma_1}\frac{R_{01}}{R_{00}}(x_1-\bar{x}_1) - \frac{\sigma_0}{\sigma_2}\frac{R_{02}}{R_{00}}(x_2-\bar{x}_2) - \cdots\cdots - \frac{\sigma_0}{\sigma_k}\frac{R_{0k}}{R_{00}}(x_k-\bar{x}_k)$$

又その重相関係　は

$$r = \sqrt{1-\frac{R}{R_{00}}}$$

（但し $\bar{x}_i$ は x_i の平均，σ_i^2 は x_i の分散とする）
となる。　そこで数値計算の場合は，$\bar{x}_i$，σ_i，r_{ij} の計算はよいとして；R 及び R_{ij} の計算が k が大きくなると非常に厄介になってくる。　そのために行列式の値を次の様な手順で計算すると間違ひをすることがなくて都合がよい。

つまり，$\frac{R}{R_{00}}$，$\frac{R_{0i}}{R_{00}}$ を求めるには

$$\begin{cases} x_0 + r_{01}x_1 + r_{02}x_2 + \cdots\cdots + r_{0k}x_k = 1 \\ r_{01}x_0 + x_1 + r_{12}x_2 + \cdots\cdots + r_{1k}x_k = 0 \\ r_{02}x_0 + r_{12}x_1 + x_2 + \cdots\cdots + r_{2k}x_k = 0 \\ \\ r_{0k}x_0 + r_{1k}x_1 + \cdots\cdots\cdots + x_k = 0 \end{cases}$$

又は，これを書きなほして

$$\begin{pmatrix} 1 & r_{01} & r_{02} & \cdots & r_{0k} \\ r_{01} & 1 & r_{12} & & \vdots \\ r_{02} & r_{12} & 1 & & \vdots \\ \vdots & & & \ddots & \vdots \\ r_{0k} & \cdots & \cdots & \cdots & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_k \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbb{R}\mathbb{X} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

から x_i を求めれば

$$x_i = \frac{R_{0i}}{R}$$

となるから，両辺を r_{0i} でわつて $\frac{R_{0i}}{R}$ がもとまる。

$\frac{R_{0i}}{R_{00}}$ も $\frac{R_{0i}}{R} \Big/ \frac{R_{0c}}{R}$ とすれば求めることが出来る。又 R_{ij} を求めるには右辺の（1, 0, 0, ---- 0）の縦ヴエクトルを（0, 0, ----, 0, 1, 0, ----, 0）と 1 の位置を適当に於きかへることによつて，同様に計算すればよい。これは偏相関係数を出す時に必要である。次に連立方程式の計算であるが，まず次にその簡単な実例を示すことにする。

東京都に於ける一日の平均濕度と出火件数とについて相関係数を計算したところ次のやうになつた。

		出火件数	その日の濕度	前日の濕度	2日前の濕度
x_0	出火件数	1	−0.512	−0.419	−0.303
x_1	その日の濕度	−0.512	1	0.540	0.353
x_2	前日の濕度	−0.419	0.540	1	0.540
x_3	2日前の濕度	−0.303	0.353	0.540	1

	平均	標準偏差
出火件数	4.718	2.994
平均湿度	74.575	12.600

これを使つて前三日間の湿度を知つて，その日の出火件数を推定する一次式，つまり前三日間の湿度と出火件数の回帰平面及びその重相関係数を計算する。

手　順　第　1

1,000 000	−,512 000	−,419 000	−,303 000	1,000 000
	1,000 000	,540 000	,353 000	,000 000
		1,000 000	,540 000	,000 000
			1,000 000	,000 000

［註］

まず充分大きな紙（縦の長さに注意）に上の様な係数の行列を記入する　このとき，行の間隔は以後の計算のために適当にとっておかなければならない。

1や0でも必要な桁数だけ0をつけておく。

重相関係数を三桁まで正しく出すために，計算は六桁くらいとる。
この行列は対象であるから，行列の半分だけ書けばよい。

手順第2

①	②	③	④		チェック欄
1,000 000 →	-.512 000 →	-.419 000 →	-.303 000 →	1,000 000 →	① .766 000
	1,000 000 →	.540 000 →	.353 000 →	.000 000 →	② 1,381 000
		1,000 000 →	.540 000 →	.000 000 →	③ 1,661 000
			1,000 000 →	.000 000 →	④ 1,590 000

［註］

ここでは計算の間違いをチェックする欄を右はしに作る。

行列の係数を全部書けばそのまま各行の係数を横に加へてチェック欄を作ればよいのであるが，この場合は手間をはぶいて半分だけしか書いてないから下半分があるつもりで上の様に矢印の通りに加へてこれを作る。

手順第3の1

1,000 000	−.512 000	−.419 000	−.303 000	1.000 000	.766 000
	.262 144	.214 528	.155 136	−.512 000	−.392 192
	1,000 000	.540 000	.353 000	.000 000	1.381 000
		1,000 000	.540 000	.000 000	1.661 000
			1,000 000	.000 000	1.590 000
	.737 856	.325 472	.197 864	.512 000	1.773 192

[註] ここからいよいよ計算に入る。まず x_0 を消去するのがこの第3の手段である。まず第一行の左端の係数で第一行をわる。いまこれを第一行 (2) としておく。実際の場合はこの逆数をかけた方がよい。しかしこの場合はそれが1であるので，そのままでよい。これは第4の手順と見くらべるとわかる。 第二行，第一列の係数を第一行 (2)（この場合は第一行と同じ）の各係数にかけて，それに対応する第二行の係数からそれぞれひくと x_0 を含まない式が一つ得られることになる。しかし対象行列では第二行第一列の係数と第一行第二列の係数は同じであるからここでは −0.512 を第一列にかけて第二行からひく。かけたものを第二行の上に書きひいたものはずっと下に書く。これが第四手順のもとになる。

そこでこの計算が正しいかどうかを次にしらべる。それには，0.737856 から 0.512000 までを加へ，これがチェック欄の 1.773192 と等しくなるかどうかをしらべる。しかし小数点以下六位目で1とか2とかちがうことがあるが，これは四捨五入のために出来たので氣にする必要は勿論ない。

手順第3の2

1,000,000	-512,000	-419,000	-303,000	1,000,000	,766000
	,262144	,214528	,155136	-,512000	-,392192
	1,000000	,540000	,353000	,000000	1,381000
		,175561	,126957	-,419000	-,320954
		1,000000	,540000	,000000	1,661000
			,091809	-,303000	-,232098
			1,000000	,000000	1,590000

,737856	,325472	,197864	,512000	1,773192
	,824439	,413043	,419000	1,981954
		,908191	,303000	1,822098

[註]

第3の2を各行で行へば上のようになる。チェックのしかたは第2でチェック欄を作るときのやうに加へ，その加へた結果とチェック欄とが等しくなれば計算はまちがつていない。
これで x_0 を含まない式が三つ出来たことになるので，以下は，これを何度もくりかへすにすぎない。

手　順　第　4

1,000 000	−,512 000	−,419 000	−,303 000	1,000 000	,766 000
	,262 144	,214 528	,155 136	−,512 000	−,392 192
	1,000 000	,540 000	,353 000	,000 000	1,381000
		,175 561	,126957	,419 000	−,320954
		1,000000	,540 000	,000000	1,661000
			,091809	−,303000	−,232098
			1,000000	,000000	1,590000

1.	441 105	268 161	693 902	2,403 168
,737856	,325.472	,197864	,512 000	1,773 192
	,143 567	087279	,225846	,782164
	:,824 439	,413043	,419000	1,981954
		,908191	,303000	1,822098
	,680872	,325764	,193 154	1.199790

[註]　第3で得られた行列の第一行を第一行第一列の係数 0,737856 でわり，それを第一行の上に書く。これが前に述べた第一行(2)である。
次に第一行(2)に第二行第一列の係数つまり，第一行第二列の係数 0.325472 をかけて，第二行からひく。以下は第2第3の手順とまつたく同じで，終には x_4 だけを含む式になる。　これを順次に代入して行けば x_3, x_2, x_1, x_0 を求めることが出来る。

完成したもの

	(−288143)	(−097854)	(−031756)	(1,417753)	
1,000000	−,512000	−,419000	−303000	1,000000	766000
	262144	214528	155136	−512000	−392192
	1,000000	540000	.353000	.000000	1.381000
		175561	126957	−419000	−,320954
		1,000000	.540000	.000000	1.661000
			.091809	−.303000	−232098
			1.000000	.000000	1.590000

	(103017)	(028105)	(.562780)	
1.000000	.441105	.268161	.693902	2.403168
737856	.325472	.197864	.512000	1.773192
	.143567	087279	.225846	.782164
	.824439	413043	.419000	1.981954
		053059	137298	475500
		908191	.303000	1.822098

	(050144)	(.233542)	
1.000000	478451	283686	1.762137
.680872	.325764	.193154	1.199790
	155862	092415	574041
	.855132	.165762	1.346598

1.000000	(104805)	
.699270	.073287	.772557

i	0		1		2		3
$\frac{R_{0i}}{R}$	1.417753		.562780		.233542		.104505
R/R_{00}	.705341	$\frac{R_{0i}}{R_{00}}$	.396952		.164727		.073923
$1-R/R_{00}$	.294659	$\frac{\sigma_0}{\sigma_i}$	.237619		〃		〃
γ	.5428	$\frac{\sigma_0 R_{0i}}{\sigma_i R_{00}}$	.094323		.039127		.017538

$$x_0 = -(0.094x_1 + 0.039x_2 + 0.018x_3) + 15.979$$

§ 3. 重相関係数を最大にする数量化の方法

（重相関比による予測の方法）

定性的な豫測因子から定量的な豫後を能率よく豫測するには定性的な豫測因子をいかに能率よく数量化するかの問題である。（この理論については講究録第6巻第11号参照）これには両者の相関係数を最大にする立場から因子を数量化する方法をのべる。

N人の人について，よみかきの試験を行ひ同時に各人の学歴と年令を調査したところ，次のやうな結果を得た。

	学歴＼得点	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	計
x_1	ナシ	15	3		1	1	1					1					1			23
x_2	中退	6		3	5	2	1	4	1	1	1	1	2		1	2	2	1		33
x_3	小	3	2	6	7	10	13	5	13	8	14	12	10	10	12	24	25	28	19	221
x_4	高小		1		2			1	2	6	6	10	17	19	28	41	80	116	140	469
x_5	中									1	1			1	3	9	12	41	112	180
x_6	以上														1	2	6	7	58	74
	total	24	6	9	15	13	15	10	16	16	22	24	29	30	45	78	126	193	329	1000

$$\eta = 0.708$$

	年令＼得点	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	計
y_1	10台			1	2				3	4	7	8	12	6	13	21	27	43	32	179
y_2	20-30台	3	1	2	4	3	5	3	5	6	7	7	5	16	14	38	62	99	218	498
y_3	40台	2	2	1	6	3	5	6	3	4	5	4	7	5	9	12	25	24	55	178
y_4	50台	13		3	1	6	5		4	1	3	4	3	3	5	5	11	22	21	110
y_5	60以上	6	3	2	2	1		1	1	1		1	2		4	2	1	5	3	35
	Total	24	6	9	15	13	15	10	16	16	22	24	29	30	45	78	126	193	329	1000

$$\eta = 0.369$$

年令＼学歴	x_1	x_2	x_3	x_4	x_5	x_6	計
y_1			9	122	24	24	179
y_2		7	84	247	124	36	498
y_3	5	13	70	60	22	8	178
y_4	10	10	44	34	9	3	110
y_5	8	3	14	6	1	3	35
Total	23	33	221	469	180	74	1000

これらの表から，学歴と年令が試験の成績にかなりの関係があることがわかる。それならば，ある人について年令と学歴がわかっておる時，この人がもしその試験を行へば何点をとるかを推定出来ないであらうか。今その合理的な予測方法を考えてみることにする。統計的な予測をするには学歴と年令を数量 (x, y) になおして，試験の成績 Z について

$$Z = f(xy)$$

と函数の型にするのが常套である。このとき函数の型は出来るだけ簡単で，しかも予測の精度のよいものでなければならない。

一番簡単な函数の型は精度の計算から云つても一次式であるので，ここでは一次式をとり上げることにする。

そこで今学歴 x_i，年令 y_j，点数 Z_k の人が f_{ijk} 人いたとする。そのとき

$$Z = x + y$$

と函数の型をきめて，

$$\sum_i \sum_j \sum_k f_{ijk} \{ Z_k - (x_i + y_j) \}^2 = A$$

が最小になるやう，x, y の数量を定める。

そのために

$$\frac{\partial A}{\partial x_i} = 0, \qquad \frac{\partial A}{\partial y_j} = 0$$ とおけば

$$f_{i\cdot}\, x_i + \sum_j f_{ij}\, y_j = f_{i\cdot}\, M_{i\cdot}$$

$$\sum_i f_{ij}\, x_i + f_{\cdot j}\, y_j = f_{\cdot j}\, M_{\cdot j}$$

但

$$f_{ij} = \sum_k f_{ijk}$$

$$f_{i\cdot} = \sum_j \sum_k f_{ijk}$$

$$f_{\cdot j} = \sum_i \sum_k f_{ijk}$$

$$M_{i\cdot} = \frac{1}{f_{i\cdot}} \sum_j \sum_k f_{ijk}\, Z_{ijk}$$

$$M_{\cdot j} = \frac{1}{f_{\cdot j}} \sum_i \sum_k f_{ijk}\, Z_{ijk}$$

又

$$M_{ij} = \frac{1}{f_{ij}} \sum_k f_{ijk}\, Z_{ijk}$$ （後出）

$$M = \frac{1}{N} \sum_i \sum_j \sum_k f_{ijk}\, Z_{ijk}$$ （〃〃）

この式を書きなおせば

$$\begin{pmatrix} f_{1\cdot} & 0 & 0 & f_{11} & f_{12} & f_{13} \\ 0 & f_{2\cdot} & 0 & f_{21} & f_{22} & f_{23} \\ 0 & 0 & f_{3\cdot} & f_{31} & f_{32} & f_{33} \\ f_{11} & f_{21} & f_{31} & f_{\cdot 1} & 0 & 0 \\ f_{12} & f_{22} & f_{32} & 0 & f_{\cdot 2} & 0 \\ f_{13} & f_{23} & f_{33} & 0 & 0 & f_{\cdot 3} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} f_{1\cdot} M_{1\cdot} \\ f_{2\cdot} M_{2\cdot} \\ f_{3\cdot} M_{3\cdot} \\ f_{\cdot 1} M_{\cdot 1} \\ f_{\cdot 2} M_{\cdot 2} \\ f_{\cdot 3} M_{\cdot 3} \end{pmatrix}$$

但し　$i = 1, 2, 3$

　　　$j = 1, 2, 3$

となり，これから x_i と y_j を求めればよい。

この式の型から，一般に x_i y_j の比が決定される。 そしてかうして定めた x_i と y_j を用いたとき Z と $(x+y)$ との相関係数 ρ は

$$\rho^2 = \frac{\frac{1}{N}\Sigma f_{ij} M_{ij}^2 - M^2}{\sigma_Z^2}$$

で與えられる。 前出の表からこれを計算してみると

y_1	y_2	y_3	y_4	y_5	x_1	x_2	x_3	x_4	x_5	x_6	
1	0	0	0	0	0	0	05028	68156	13408	13408	15.01
0	1	0	0	0	0	01406	16867	49598	24900	07229	16.00
0	0	1	0	0	02809	07303	39326	33708	12360	04494	14.20
0	0	0	1	0	09091	09091	40001	30908	08182	02727	12.26
0	0	0	0	1	22857	08571	40001	17143	02857	08571	9.46
0	0	21739	43478	34783	1	0	0	0	0	0	2.74
0	21212	39394	30303	09091	0	1	0	0	0	0	7.09
04072	38009	31674	19910	06335	0	0	1	0	0	0	12.04
26013	52666	12793	07249	01279	0	0	0	1	0	0	15.96
13333	68889	12222	05000	00556	0	0	0	0	1	0	17.30
32432	48649	10811	04054	04054	0	0	0	0	0	1	17.61
7.0	8.0	7.0	6.0	5.0	1.0	3.0	6.0	8.0	9.0	9.0	
6.84	8.09	7.38	6.04	4.17	-2.94	0.26	4.88	8.48	9.59	10.24	
6.326	7.839	7.841	6.867	5.552	-3.874	-0.247	4.600	8.614	9.724	10.394	
6.210	7.783	7.945	7.053	5.881	-4.099	-0.363	4.536	8.643	9.754	10.426	
6.155	7.771	7.970	7.097	5.960	-4.151	-0.391	4.519	8.648	9.760	10.433	
6.181	7.770	7.978	7.109	5.979	-4.165	-0.399	4.514	8.648	9.759	10.432	
6.182	7.771	7.980	7.113	5.985	-4.169	-0.402	4.512	8.646	9.758	10.431	

この数量を用いて，それぞれの相関係を計算してみると，

$$\rho_{xz} = 0.709$$

$$\rho_{yz} = 0.155$$

$$\rho_{xy} = -0.007$$

Zと $x+y$ の重相関係数は，

$$\rho = \sqrt{1 - \frac{\begin{vmatrix} 1 & \rho_{xz} & \rho_{yz} \\ \rho_{xz} & 1 & \rho_{xy} \\ \rho_{yz} & \rho_{xy} & 1 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_{xy} \\ \rho_{xy} & 1 \end{vmatrix}}}$$

$$= 0.726$$

となる。

しかしこの方法は非常に手数がかかるので（結局数量を與へようとする要因の数の元数の聯立一次方程式をとかなければならない），次の様な単便法を用うる。つまり，要因の数量として，その要因の中含まれるものの平均値，即ち，中学卒業者の数量としては，中学卒業者の平均得点 17.3 を與えて，同様に最小自乗法を用いて（結局重相関）予測式を作ればよい。　この方法は前述の方法の非常によい近似となつてゐる。　因みにこの方法で前記の例について重相関係数を計算すれば 0.712 となる。

この方法が第一次近似であることは前記の理論によつて容易に示される。　この時は数量があらかじめ與へられてしまふので計算は全く簡単になり，§1 にのべた重相関係数の計算法によつて求められてしまふ（§1 の応用問題となつてしまふ）

§ 4.　GUTTMAN の PAIRED　COMPARISON

N人の審査員がn人の美人についてその順位をきめようとする場合，美人数nが多くなると，これを一時にみることは出来ないし，又各審査員がn人の順位を決め得たとしても，N人の結果をどうまとめればよいかと云ふことが問題になつてくる。

$Guttman$はこの問題を解決するために，美人を二人づつ組にして撰び出し，どちらがきれいかを審査員がきめて，これを次に述べるような方法でまとめあげることを考えた。

美人が O_1，O_2，O_3，O_4，O_5，O_6 の６人，審査員が４人居る。n，N は何人でも同様である。　審査の勝ち負けのしるしに

$$e_{ijk}$$

と云う符号を用いる。　これはj子さんとk子さんをi君が審査した結果を表すものである．　今i君がj子さんの方がk子さんよりもきれいであるとしたならば

$$e_{ijk} = 1, \qquad e_{ikj} = 0$$

k子さんの方がきれいだとなれば

$$e_{ijk} = 0, \qquad e_{ikj} = 1$$

と決める。　このとき，i君は必ずどちらがきれいかを決めなければならない。　即ち　どちらも，同じくらいかとか，わからないとか云うやうなことは云ってはならないとしておくと上からも分るように

$$e_{ijk} + e_{ikj} = 1$$

となる。　つまり勝てば１点を得，負ければ０となる。

次に

f_{ij} をi君の審査によるj子さんの勝点
g_{ij} をi君の審査によるj子さんの負点

とすれば

$$f_{ij} = \sum_k e_{ijk}, \quad g_{ij} = \sum_k e_{ikj}$$

となる。kは相手である。　又

$$f_{ij} + g_{ij} = n - 1 = 6 - 1 = 5$$

となることは勿論である。

さてこの審査の結果が次のやうになつた。

$n = 6$

e_{ijk}

$N = 4$

[第1審査員審査表]

$i = 1$

j \ k	O_1	O_2	O_3	O_4	O_5	O_6	f_{ij}
O_1		1	0	1	0	1	3
O_2	0		0	0	0	1	1
O_3	1	1		1	1	1	5
O_4	0	1	0		0	0	1
O_5	1	1	0	1		1	4
O_6	0	0	0	1	0		1
g_{ik}	2	4	0	4	1	4	15 = F

[第2審査員審査表]

$i = 2$

j \ k	O_1	O_2	O_3	O_4	O_5	O_6	f_{ij}
O_1		1	0	1	0	0	2
O_2	0		0	1	1	0	2
O_3	1	1		0	1	0	3
O_4	0	0	1		0	1	2
O_5	1	0	0	1		1	3
O_6	1	1	1	0	0		3
g_{ik}	3	3	2	1	2	2	15

[第3審査員審査表]

$i = 3$

j \ k	O_1	O_2	O_3	O_4	O_5	O_6	f_{ij}
O_1		1	0	1	0	1	3
O_2	0		0	0	0	0	0
O_3	1	1		1	0	1	4
O_4	0	1	0		0	0	1
O_5	1	1	1	1		1	5
O_6	0	1	0	1	0		2
g_{ik}	2	5	1	4	0	3	15

第4審査員審査表

j \ k	O_1	O_2	O_3	O_4	O_5	O_6	f_{ij}
O_1		1	0	1	0	0	2
O_2	0		0	1	1	1	3
O_3	1	1		1	1	1	5
O_4	0	0	0		0	0	0
O_5	1	0	0	1		1	3
O_6	1	0	0	1			2
g_{ik}	3	2	0	5	2	3	15

この関係は6大学野球リーグ戦の四シーズンの戦績とみればわかりがよいと思う。

各審査員の判定を下した回数をF（シーズンの試合の回数）とすれば

$$F = \binom{n}{2} = \frac{n(n-1)}{2} = \sum_j f_{ij}$$

この場合は

$$F = 15$$

となる。

Cを一人の美人の審査をうけた回数（チームの四シーズンの全試合回数）とすれば

$$C = (n-1)N \equiv \sum_i (f_{ij} + g_{ij})$$

この場合

$$C = 20$$

Cを全審査回数（四シーズンの全試合回数）とれば

$$C = n(n-1)N$$

この場合は

$$C = 120$$

さてこれだけの準備をして6人の美人の美しさの度合いを表わす値，x_1, x_2, x_3, x_4, x_5, x_6 をきめるのが次の問題である。

そのために次の諸量をまずきめておく。

t_i ―― i君の與えた勝点についてのxの平均

u_i ―― 同じく負点についてのxの平均

$$t_i \equiv \frac{1}{F} \sum_k x_k f_{ik}$$

$$u_i \equiv \frac{1}{F} \sum_k x_k g_{ik}$$

y_i ―― i君の與えた勝点についてのxの分散度

$$y_i \equiv \sum_k (x_k - t_i)^2 f_{ik} \equiv \sum_k x_k^2 f_{ik} - t_i^2 F$$

$$Z_i \equiv \sum_k (x_k - u_i)^2 g_{ik} \equiv \sum_k x_k^2 g_{ik} - u_i^2 F$$

V —— xの全審査にわたつての平均

W —— xの全審査にわたつての分散度

$$V = \frac{1}{C} \sum_k x_k C = \frac{1}{n} \sum_k x_k$$

$$W = \sum_k (x_k - V)^2 C = C \sum_k x_k^2 - V^2 C$$

R —— 審査員間のxの値の分散度

S —— 各審査員間の分散度の総和

$$R = \sum_i \left[(t_i - V)^2 + (u_i - V)^2 \right] F = F \sum_i (t_i^2 + u_i^2) - V^2 C$$

$$S = \sum_i (y_i - Z_i) = W - R$$

E —— 相関比

$$E^2 = 1 - \frac{S}{W} = \frac{R}{W}$$

そこでこの E^2 が最大になるように $x_1, \cdots\cdots x_\ell$, をきめる。この意味は各審査員についての判定の分散を最小にして審査員間の判定の分散を最大にするやうxの値をきめることである。

さらに云いかえれば，各審査員の好みをはつきり表し，しかも審査員の個人差を出来るだけ少なくするやうにxの値をきめることになる。　野球で云えば各シーズンの強さが最もよく表はされるようにXの値をきめるわけである。

そのためEの極値をもつ時の $x_1, x_2, \cdots\cdots, x_6$ の値を探すのであるから，E^2 を $x_1, x_2, \cdots\cdots x_6$ で微分して0と置きこれから $x_1, x_2, \cdots\cdots x_6$ をきめればよい。

式の形から明らかなように一般性を失わずに

$$V = 0$$

とすることが出来る。 さうすれば

$$R = \sum_i (t_i^2 + u_i^2)$$

$$W = C \sum_k x_k^2$$

となり

$$\frac{\partial E^2}{\partial x_j} = 0$$

と置けば

$$\frac{\partial R}{\partial x_j} = E^2 \frac{\partial W}{\partial x_j}$$

そこで

$$\frac{\partial R}{\partial x_j} = \frac{2}{F} \sum_k x_k \sum_i (f_{ij} f_{ik} + g_{ij} g_{ik})$$

$$\frac{\partial W}{\partial x_j} = 2C_j$$

である。

今 $$H_{ik} \fallingdotseq \frac{1}{CF} \Sigma_i (f_{ij} f_{ik} + g_{ij} g_{ik})$$

$$\sum_k x_k H_{ik} = E^2 x_j$$

つまり

$$\mathbb{X} \mathbb{H} = E \mathbb{X}$$

と書ける。 但

$\mathbb{X}$ は $(x_1, x_2, \cdots\cdots, x_6)$ なるベクトル

$\mathbb{H}$ は H_{ik} を要素としたマトリックスである.

これから $\mathbb{X}$ を求めればよい。 我々の場合のマトリックス $\mathbb{H}$

を作つてみれば次のやうになる。

f_{ij}

	1	2	3	4	5	6
	3	1	5	1	4	1
i	2	2	3	2	3	3
	3	0	4	1	5	2
	2	3	5	0	3	2

g_{ij}

	1	2	3	4	5	6
	2	4	0	4	1	4
i	3	3	2	3	2	2
	2	5	1	4	0	3
	3	2	0	5	2	3

iについて加える

$$\Sigma(f_{i1}f_{i1}+g_{i1}g_{i1}) = 9+4+9+4+4+9+4+9 = 52$$

$$\Sigma(f_{i1}f_{i2}+g_{i1}g_{i2}) = 3+4+0+6+8+9+10+6 = 46$$

$$\Sigma(f_{i1}f_{i3}+g_{i1}g_{i3}) = 15+6+12+10+0+6+2+0 = 51$$

$$\Sigma(f_{i1}f_{i4}+g_{i1}g_{i4}) = 3+4+3+0+8+9+8+15 = 50$$

$$\Sigma(f_{i1}f_{i5}+g_{i1}g_{i5}) = 12+6+15+6+2+6+0+6 = 53$$

$$\Sigma(f_{i1}f_{i6}+g_{i1}g_{i6}) = 3+6+6+4+8+6+6+9 = 48$$

$$\Sigma(f_{i2}f_{i2}+g_{i2}g_{i2}) = 1+4+0+9+16+9+25+4 = 68$$

$$\Sigma(f_{i2}f_{i3}+g_{i2}g_{i3}) = 5+6+0+15+0+6+5+0 = 37$$

$$\Sigma(f_{i2}f_{i4}+g_{i2}g_{i4}) = 1+4+0+0+16+9+20+10 = 60$$

$$\Sigma(f_{i2}f_{i5}+g_{i2}g_{i5}) = 4+6+0+9+4+6+0+4 = 33$$

$$\Sigma(f_{i2}f_{i6}+g_{i2}g_{i6}) = 1+6+0+6+16+6+15+6 = 56$$

$$\Sigma(f_{i3}f_{i3}+g_{i3}g_{i3}) = 25+9+16+25+0+4+1+0 = 80$$

$$\Sigma(f_{i3}f_{i4}+g_{i3}g_{i4}) = 5+6+4+0+0+6+4+0 = 25$$

$$\Sigma(f_{i3}f_{i5}+g_{i3}g_{i5}) = 20+9+20+15+0+4+0+0 = 68$$

$$\Sigma(f_{i3}f_{i6}+g_{i3}g_{i6}) = 5+9+8+10+0+4+3+0 = 39$$

$$\Sigma(f_{i4}f_{i4}+g_{i4}g_{i4}) = 1+4+1+0+10+9+16+25 = 72$$

$$\Sigma\,(f_{i4}f_{i5} + g_{i4}g_{i5}) = 4+6+5+0+4+6+0+10 = 35$$

$$\Sigma\,(f_{i4}f_{i6} + g_{i4}g_{i6}) = 1+6+2+0+16+6+12+15 = 58$$

$$\Sigma\,(f_{i5}f_{i5} + g_{i5}g_{i5}) = 16+9+25+9+1+4+0+4 = 68$$

$$\Sigma\,(f_{i5}f_{i6} + g_{i5}g_{i6}) = 4+9+10+6+4+4+0+6 = 43$$

$$\Sigma\,(f_{i6}f_{i6} + g_{i6}g_{i6}) = 1+9+4+4+16+4+9+9 = 56$$

$$CF = 300$$

$$
H = \begin{bmatrix}
\frac{52}{300} & \frac{46}{300} & \frac{51}{300} & \frac{50}{300} & \frac{53}{300} & \frac{48}{300} \\
\frac{46}{300} & \frac{68}{300} & \frac{37}{300} & \frac{60}{300} & \frac{33}{300} & \frac{56}{300} \\
\frac{51}{300} & \frac{37}{300} & \frac{80}{300} & \frac{25}{300} & \frac{68}{300} & \frac{39}{300} \\
\frac{50}{300} & \frac{60}{300} & \frac{25}{300} & \frac{72}{300} & \frac{35}{300} & \frac{58}{300} \\
\frac{53}{300} & \frac{33}{300} & \frac{68}{300} & \frac{35}{300} & \frac{68}{300} & \frac{43}{300} \\
\frac{48}{300} & \frac{56}{300} & \frac{39}{300} & \frac{58}{300} & \frac{43}{300} & \frac{56}{300}
\end{bmatrix}
= \begin{bmatrix}
0.17333 & 0.15333 & 0.17000 & 0.16667 & 0.17667 & 0.16000 \\
0{,}15333 & 0{,}22667 & 0{,}12333 & 0{,}20000 & 0{,}11000 & 0{,}18667 \\
0{,}17000 & 0{,}12333 & 0{,}26667 & 0{,}08333 & 0{,}22667 & 0{,}13000 \\
0{,}16667 & 0{,}20000 & 0{,}8333 & 0{,}24000 & 0{,}11667 & 0{,}19333 \\
0{,}17667 & 0{,}11000 & 0{,}22667 & 0{,}11667 & 0{,}22667 & 0{,}14333 \\
0{,}16000 & 0{,}18667 & 0{,}13000 & 0{,}19333 & 0{,}14333 & 0{,}18667
\end{bmatrix}
$$

［このマトリックスは各行　　　の和が1になるから計算のまちがひがすぐわかる］

これを実際にとくために逐次近似法を用いる。それには，x_1，x_2，x_3，x_4，x_5，x_6 の大まかな値をきめなければならない。

そこで O_1，O_2，------ O_6 の大体の順位をきめてみる。

O_j の大体の順位を r_i としたとき x_j の大まかな値として，Guttman は次のやうにとることをすすめてゐる。

$$x_j^{(1)} = 2r_j - (n+1)$$

即ち

j	O_1	O_2	O_3	O_4	O_5	O_6
f_{1j}	3	1	5	1	4	1
f_{2j}	2	2	3	2	3	3
f_{3j}	3	0	4	1	5	2
f_{4j}	2	3	5	0	3	2
$\frac{1}{t}$と（勝点）	10	6	17	4	15	8
大体の順位	3	5	1	6	2	4
$x_j^{(1)}$	-1	3	-5	5	-3	1

それで二番の x の値を次のやうに計算する。

$$E^2 x_1^{(2)} = -1 \times 0.17333 + 3 \times 0.15333 - 5 \times 0.17000 + 5 \times 0.16667 - 3 \times 0.17667 + 1 \times 0.16000$$
$$= -0.10000$$

$$E^2 x_2^{(2)} = -1 \times 0.15333 + 3 \times 0.22667 \cdots + 1 \times 0.18667$$
$$= 0.76670$$

$$E^2 x_6^{(2)} = -1 \times 0.16000 + 3 \times 0.18667 \cdots\cdots$$
$$= 0.47334$$

— — — — — — — — — — — — — —

それから x_1, x_2 -------- x_6 の値は比だけが問題なのであるから，x_6 を 1 にしてしまふと，

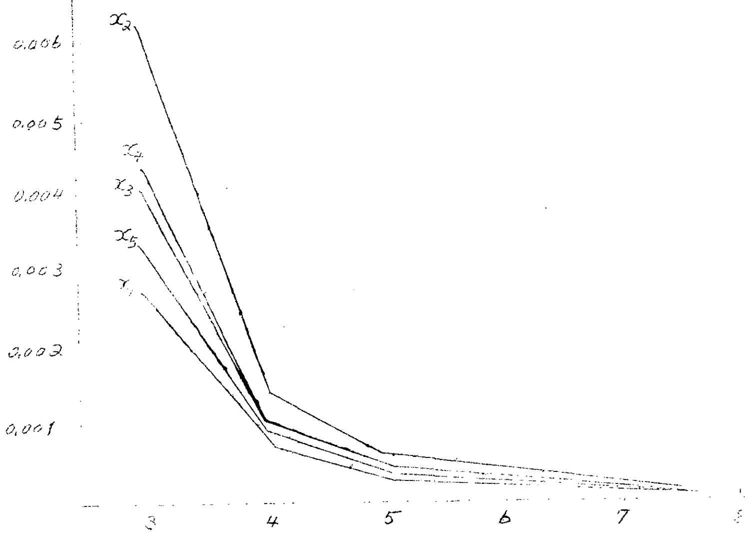
$$x_1^{(2)} = -0.21126$$
$$x_2^{(2)} = 1.61977$$
$$x_3^{(2)} = -2.67613$$
$$\vdots$$
$$x_6^{(2)} = 1{,}00000$$

それを次々にくりかえしてゆくと次の表になり遂には $x_j^{(i)}$ と $x_j^{(i+1)}$ とが同じ値になつてしまう。

これが，もとめるものである。又この次第に近づいてゆくさまをグラフに書いてみた。

-1.0	3.0	-5.0	5.0	-3.0	1.0
-0.10000 -0.21126	0.76670 1.61977	-1.26672 -2.67613	1.06000 2.23941	-0.9335 -1.97184	0.47334 1.00000
-0.05832 -0.22635	0.42236 1.63925	-0.68014 -2.63953	0.51147 2.19856	-0.50180 -1.97193	0.25766 1.00000
-0.05855 -0.22896	0.42108 1.64555	-0.67397 -2.63551	0.56109 2.19421	-0.50512 -1.97193	0.25572 1.00000
-0.05868 -0.22950	0.42108 1.64690	-0.67368 -2.63480	0.56082 2.19344	-0.50524 -1.97604	0.25568 1.00000
-0.05870 -0.22958	0.42116 1.64722	-0.67365 -2.63470	0.56078 2.19329	-0.50529 -1.97622	0.25568 1.00000
-0.05871 -0.22961	0.42118 1.64723	-0.67366 -2.63463	0.56078 2.19323	-0.505271 -1.9762524	0.25569 1.00000
-0.0587088 -0.2296142	0.4211744 1.6472426	-0.6736455 -2.6346748	0.5607728 2.1932217	-0.5052971 -1.9762524	0.2556845 1.00000
-0.0587091 -0.2296148	0.4211760 1.6472443	-0.1736472 -2.6346742	0.5607742 2.1932219	-0.5052988 -1.9762536	0.2556852 1.000000

近似の回数



我々のもとめる値は

x_1	x_2	x_3	x_4	x_5	x_6
-0.230	1.643	-2.635	2.193	-1.976	1.000

で相関比Eは

$$E^2 = 0.2557$$

$$E = 0.5$$

となる。このxの値は小さければ小さいほど美人であると言ふ事になる。即ち美人順にならべれば O_3, O_5, O_1, O_6, O_2, O_4 となつて O_3 がミスの王座につく。

§ 5. HAYASHIのGROUPINGの問題

近頃流行の *Sociometry* の一つの数量化の方法である。

Sociometry がマトリックスの形であらはされてゐるとき，これを *grouping* する方法を考へるとき用ひられる一つの数量化方法（講究録第6号，第11号参照）に関する計算法をのべてみる。

N人の人間を組分けするとき，N人相互の好みを考えに入れてこれを行うにはどうすればよいか。

それにはまず，N人各自について相手の好みの程度を

非常にすきなら	2点
まあすきなら	1点
すきでもきらいでもなければ	0点
きらいなら	−1点
非常にきらいなら	−2点

（この点数の與え方はどうでもよい）

で点数をつけさせる。　今，N人の人間を $O_1, O_2, \cdots\cdots, O_n$ で表し，その点数を e_{ij} とする。これは O_i 君が O_j 君を好む度合いが，e_{ij}（2, 1, 0, −1, −2）であると云うことを示してゐる。

さうすると，N人全体については次の表が得られる。

	O_1	O_2	O_3	O_4	O_5
O_1		e_{12}	e_{13}	e_{14}	e_{15}
O_2	e_{21}		e_{23}	e_{24}	e_{25}
O_3	e_{31}	e_{32}		e_{33}	e_{34}
O_4	e_{41}	e_{42}	e_{43}		e_{45}
O_5	e_{51}	e_{52}	e_{53}	e_{54}	

但しこの時Nは5

この表について，好みの度が大きいものがなるべく近くにくるように O_i に数量 x_i を興えることにする。

それには，

$\sum_i x_i^2 = 1$ の條件のもとに，

$\sum_i \sum_j e_{ij}(x_i - x_j)^2$ が最小になるようにすればよい。

よって，

$$\sum_i \sum_j e_{ij}(x_i - x_j)^2 - \lambda\left(\sum_i x_i^2 - 1\right) = A$$

とすれば

$$\frac{\partial A}{\partial x_i} = 0$$

として，x_i を求めることになる。

つまり

$$\sum_j e_{ij}(x_i - x_j) - \sum_j e_{ji}(x_j - x_i) - \lambda x_i = 0$$

$$\left\{\sum_j (e_{ij} + e_{ji}) - \lambda\right\} x_i - \sum_j (e_{ij} + e_{ji}) x_j = 0$$

この式から，x_i を計算する．例えば今 e_{ij} の表を

	O_1	O_2	O_3	O_4
O_1	＼	2	−1	−2
O_2	1	＼	1	−1
O_3	−1	0	＼	1
O_4	−1	−1	1	＼

とする。

これから $e_{ij} + e_{ji}$ の表を作れば

	O_1	O_2	O_3	O_4	Σ
O_1		3	−2	−3	−2
O_2	3		1	−2	2
O_3	−2	1		2	1
O_4	−3	−2	2		−3

となるから計算式は

$$\begin{pmatrix} 2-\lambda & 3 & -2 & -3 \\ 3 & -2-\lambda & 1 & -2 \\ -2 & 1 & -1-\lambda & 2 \\ -3 & -2 & 2 & 3-\lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = 0$$

で、これから λ と x_i を求めればよい。

2	3	-2	-3
3	-2	1	-2
-2	1	-1	2
-3	-2	2	3
2	1	0	-1
10	6	-5	-11
1.0	0.6	-0.5	-1.1
8.1	3.5	-3.1	-8.5
1.00	0.43	-0.38	-1.05
7.20	3.86	-3.29	-7.77
1.00	0.54	-0.46	-1.08
7.78	3.62	-3.16	-8.21
1.000	0.465	-0.406	-1.059
7.384	3.782	-3.247	-7.919
1.000	0.512	-0.440	-1.072
7.632	3.680	-3.192	-8.120
1	0.482	-0.418	-1.064
1.0	0.497	-0.430	-1.069
7.558	0.3714	-3.211	-8.061
1.0	0.491	-0.425	-1.067
1.0	0.494	-0.428	-1.067
7.542	0.372	-3.214	-8.048
1.0	0.493	-0.426	-1.067

即ち，

x_1	1.000
x_2	0.493
x_3	−0.426
x_4	−1.067

となりこの数量が我々の與えた意味だけに於いて O_i の近さを表してゐることになる。

その2，その3は共に特有根と特有ベクトル（*Eigenwert*，*Eigenvektor*）の数値解法になるが，それには *R. A. Frazer*，*W. J. Duncan* の *Elementary Matrices* をみるとよく書いてある。

§ 6. PAROIE PREDICTION における数量化に用ひられた計算

原論文： C. HAYASHI ： Approach for Quantifying Qualitative Data from the Mathematico-Statistical Point of View. Ann. Math. Stat. Vol. II No. 1

この計算法は，丸山文行氏が考へ，後に石田が幾分手を加へたものである。

これはA, B 二つのグループがある定性的な同一種類の標識に関してある分布をもつてゐるとき，A, B 両グループが平均的に言ってなるべくへだたる様に，即ち A, B 両グループの分散が共に等しく一定の條件の下に平均値の差が最大になる様に定性的な標識を数量化する方法をのべる。

ここでむづかしいのは拘束條件が二つ，即ち A, B 両グループの分散一定と言ふ二つの條件が與へられてゐる所である。

つまりラグランシュの方法によるとき未定常数が２つある所である。

数値計算の問題としては結局，

n 個のカテゴリーからなる分布の比率 p_i, q_i （但し $i = 1, \cdots, n$ （n は４以上）： $\sum_{i=1}^{n} p_i = 1$, $\sum_{i=1}^{n} q_i = 1$ ）が與へられたとき

$$\sigma_p^2 = \Sigma p_i (x_i - s)^2 = 1$$

$$\sigma_q^2 = \Sigma q_i (x_i - t)^2 = 1$$

但し $s = \Sigma p_i x_i, \ t = \Sigma q_i x_i$

の條件のもとに，

$s - t$ が最大になる様にカテゴリーの点数 x_i をきめること、に帰着する。

そのためにラグランジユの乗数法を用いて，

$$A = s - t - \frac{\lambda}{2}(\sigma_p^2 - 1) - \frac{\mu}{2}(\sigma_q^2 - 1)$$

と置き

$$\frac{\partial A}{\partial x_i} = 0$$

から x_i を求める。

即

$$(p_i - q_i) = \lambda(p_i x_i - p_i s) + \mu(q_i x_i - q_i s) \cdots\cdots \quad ①$$

が得られる。

$$x_i = \frac{(p_i - q_i) + (\lambda p_i s + \mu q_i t)}{\lambda p_i + \mu q_i} \cdots\cdots \quad ②$$

又の式の両辺に x_i を乗じて，i について加へれば，

$$s - t = \lambda + \mu \quad \cdots\cdots \quad ③$$

ここで λ と μ に関するもう一つの簡単な恒等式が得られれば，この問題は容易に片がつくのであるが，適当なものがみつからないために，手数が厄介になる。そのために，λ と μ の比を一定として平均の差が最大になる x_i をきめ，分散の比を計算してこれが1になるやうに λ と μ の比を動かして x_i を求めて行

く方法をとる。

まず，この問題ですぐわかることは，x_i のうち一つだけはどこに決めてもよいと云うことである。　即ち，分散と平均の差のみを問題にするのであるから，座標の原点はどこに於いてもよいことになる。　そこで今我々は計算の便宜上

$$S = -t$$

と置くことにする。

$$\frac{\mu}{\lambda} = k$$

とすれば ③ 式から

$$\lambda = \frac{2S}{1+k}$$

$$\mu = \frac{2S \cdot k}{1+k}$$

これを ② に代入すれば

$$x_i = \frac{(p_i - q_i)\dfrac{1+k}{2S} + S(p_i - q_i k)}{p_i + q_i k} \quad \cdots\cdots \quad ④$$

この両辺に p_i をかけて，i について加へると，

$$S = \frac{1+k}{2S}\sum \frac{p_i - q_i}{p_i + q_i k} p_i + S \sum \frac{p_i - q_i k}{p_i + q_i k} p_i$$

故に

$$S = \sqrt{\frac{\dfrac{1+k}{2}\sum \dfrac{p_i - q_i}{p_i + q_i k} p_i}{1 - \sum \dfrac{p_i - q_i k}{p_i + q_i k} p_i}} \quad \cdots\cdots \quad ⑤$$

を得る。

そこで，計算の順序は次の様なことになる。

1. 任意に k_1 の値を決める。 この場合 k_1 は通常まず 1 として計算をはじめるが，分布の形から何か見当がつけば，又適当な値にした。 結局この値の決め方がこの計算の勘どころである。 うまい値をみつけると計算は非常にはやく終ることが出来る。

2. ⑤式から p_i, q_i と k_1 とで S_1 を計算する。

特に

$$k_1 = 1$$

と置いたときは

$$S_1 = \sqrt{\frac{\sum \frac{p_i - q_i}{p_i + q_i} p_i}{1 - \sum \frac{p_i - q_i}{p_i + q_i} p_i}}$$

となるから計算は比較的簡単になる。

3. 次に k_1, S_1, p_i, q_i を④ 式に代入して，x_{i1} を計算する。

4. この x_{i1} で $\sigma_{p_1}^2$ $\sigma_{q_1}^2$ を計算し，

$$\frac{\sigma_{p_1}^2}{\sigma_{q_1}^2} = \tau_1$$

を求める。

5. この τ_1 が 1 ならば計算は完了したことになるが，これが 1 に等しくないときは更に k_2 をきめ S_2, τ_2 を計算する。

6. τ と k とのグラフを書きこのグラフの助けをかりて τ が 1 になるやうな k をさがして，x_i を求めればそれが解である。

計算の実例を示せば次の様になる。

計算の便宜上

$$k = \frac{\mu}{\lambda}$$

$$\frac{p_i - kq_i}{p_i + kq_i} S + \frac{p_i - q_i}{p_i + kq_i} \frac{1+k}{2S} = x_i$$

$$\begin{cases} a = \dfrac{p_i - kq_i}{p_i + kq_i} \\ b = \dfrac{p_i - q_i}{p_i + kq_i} \\ c = \dfrac{1+k}{2S} \\ d = b \times (1+k) \end{cases}$$

とおけば

$\frac{\mu}{\lambda} = 1$ のとき

$$S^2 = \frac{\Sigma a p_i}{1 - \Sigma a p_i}$$

$$x_i = a(S + C)$$

$\frac{\mu}{\lambda} \neq 1$ のとき

$$S^2 = \frac{\Sigma d_i p_i}{2(1 - \Sigma a p_i)}$$

$$x_i = aS + bC$$

となる。

p 0.024 0.191 0.262 0.333 0.095 0.095

q 0.418 0.340 0.079 0.042 0.085 0.036

について，上に述べた方法によつて計算を行つてみる。

$k = 1$ のとき

$p-q$	−0.394	−0.149	0.183	0.291	0.010	0.059
$p+q$	0.442	0.531	0.341	0.375	0.180	0.131
a	−0.8914	−0.2806	0.5367	0.7760	0.0556	0.4504
x	−1.8441	−0.5805	1.1103	1.6054	0.1150	0.9318
x^2	3.4007	0.3370	1.2328	2.5773	0.0132	0.8683

$$S^2 = \frac{0.37211}{1-0.37211} = 0.59264$$

$$S = 0.7698$$

$$C = 1.2890$$

$$S + C = 2.0688$$

m_p 0.7698 σ_p^2 0.8184

m_q −0.7697 σ_q^2 1.1817

$$\frac{\sigma_q^2}{\sigma_p^2} = 1.444$$

$k = 10$ のとき

kq	4.180	3.400	0.790	0.420	0.850	0.360
$P-kq$	−4.156	−3.209	−0.528	−0.087	−0.755	−0.265
$P+kq$	4.204	3.591	1.052	0.753	0.945	0.455
a	−0.9886	−0.8936	−0.5019	−0.1155	−0.7989	−0.5824
b	−0.0937	−0.0415	0.1740	0.3865	0.0106	0.1297
d	−1.0307	−0.4565	1.9140	4.2515	0.1166	1.4267
aS	−0.7986	−0.7219	−0.4054	−0.0933	−0.6454	−0.4705
bC	−0.6380	−0.2826	1.1847	2.6315	0.0722	0.8831
x	−1.4366	−1.0045	0.7793	2.5382	−0.5732	0.4126
x^2	2.0638	1.0090	0.6073	6.4425	0.3286	0.1702

$$S^2 = \frac{1.95190}{2(1+0.49559)} = 0.65256$$

$S = 0.8078$

$C = 6.8086$

$m_p \quad 0.8078$

$m_q \quad -0.8077$

$\sigma^2_p \quad 1.9416$

$\sigma^2_q \quad 0.9059$

$$\frac{\sigma_q^2}{\sigma_p^2} = 0.4655$$

$k = 5$ のとき

kq	2.090	1.700	0.395	0.210	0.425	0.180
$p-kq$	-2.066	-1.509	-0.133	0.123	-0.330	-0.085
$p+kq$	2.114	1.891	0.657	0.543	0.520	0.275
a	-0.9773	-0.7980	-0.2024	0.2265	-0.6346	-0.3091
b	-0.1864	-0.0788	0.2785	0.5359	0.0192	0.2145
d	-1.1184	-0.4728	1.6710	3.2154	0.1152	1.2870
as	-0.7653	-0.6249	-0.1585	0.1774	-0.4970	-0.2421
bc	-0.7141	-0.3019	1.0669	2.0530	0.0736	0.8217
x	-1.4794	-0.9266	0.9084	2.2304	-0.4234	0.5796
x^2	2.1886	0.8590	0.8252	4.9747	0.1793	0.3359

$$S^2 = \frac{1.52459}{2(1+0.24313)} = 0.61321$$

$S = 0.7831$

$C = 3.8309$

$m_p \quad 0.7830$

$m_q \quad -0.7832$

$\sigma_p^2 \quad 1.5252$

$\sigma_q^2 \quad 0.8950$

$$\frac{\sigma_q^2}{\sigma_p^2} = 0.5869$$

$k = 1.5$ のとき

kq	0.6270	0.5100	0.1185	0.0630	0.1275	0.0540
$p-kq$	−0.6030	−0.3190	0.1435	0.2700	−0.0325	0.0410
$p+kq$	0.6510	0.7010	0.3805	0.3960	0.2225	0.1490
a	−0.9263	−0.4551	0.3771	0.6818	−0.1461	0.2752
b	−0.6052	−0.2126	0.4809	0.7348	0.0449	0.3960
d	−1.5130	−0.5315	1.2023	1.8370	0.1123	0.9900
as	−0.7051	−0.3464	0.2870	0.5190	−0.1112	0.2095
bc	−0.9938	−0.3491	0.7897	1.2066	0.0737	0.6503
x	−1.6989	−0.6955	1.0767	1.7256	−0.0375	0.8598
x^2	2.8863	0.4837	1.1593	2.9777	0.0014	0.7387

$$S^2 = \frac{0.8936}{2\,(1-0.22895)} = 0.57948$$

$$S = 0.7612$$

$$C = 1.6421$$

m_p 0.7612

m_q −0.7613

σ_p^2 0.9479

σ_q^2 1.0347

$$\frac{\sigma_q^2}{\sigma_p^2} = 1.0915$$

$k = 1.8$ のとき

kq	0.7524	0.6120	0.1422	0.0756	0.1530	0.0648
$p\ kq$	−0.7234	−0.4210	0.1198	0.2574	−0.0580	0.0302
$p\ kq$	0.7764	0.8030	0.4042	0.4086	0.2480	0.1598
a	−0.9382	−0.5243	0.2964	0.6300	−0.2339	0.1890
b	−0.5075	−0.1856	0.4527	0.7122	0.0403	0.3692
d	−1.4210	−0.5197	1.2676	1.9942	0.1128	1.0338
as	−0.7138	−0.3989	0.2255	0.4793	−0.1780	0.1438
bc	−0.9339	−0.3415	0.8331	1.3106	0.0742	0.6794
x	−1.6477	−0.7404	1.0586	1.7899	−0.1038	0.8232
x^2	2.7149	0.5482	1.1206	3.2037	0.0108	0.6777

$$s^2 = \frac{1.06828}{2(1-0.1605)} = 0.57876$$

$s = 0.7608$

$C = 1.8402$

$m_p \quad 0.7608$

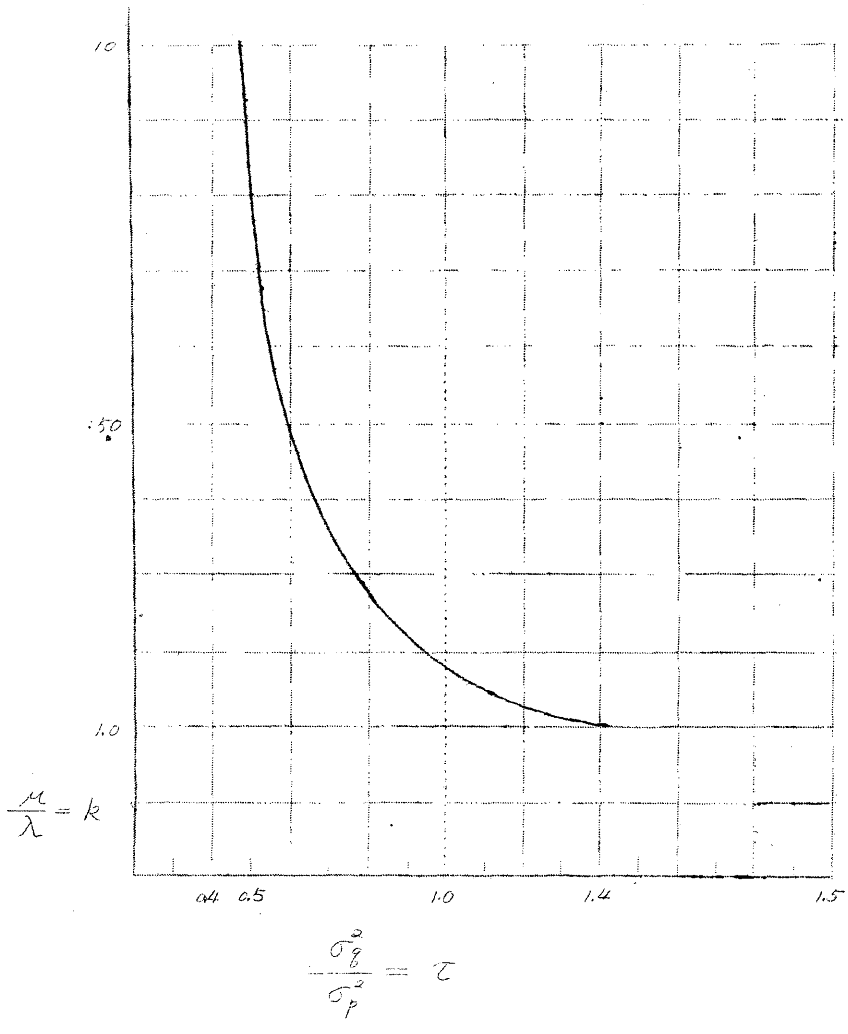
$m_q \quad -0.7609$

$\sigma_p^2 \quad 1.0169$

$\sigma_q^2 \quad 0.9906$

$$\frac{\sigma_q^2}{\sigma_p^2} = 0.9741$$

k と τ のグラフは下図の如くである。



[附　録]

（i） 研究所に新に到着した統計関係の外国雑誌名及び所在する巻数等を参考のためしるしておき，利用の便に供したいと思ふ。

American Journal of Mathematics ……… Vol. 72, No. 1-4
73, 1-4
Annals of Mathematics ……… Vol. 53, No. 1-3
54, 1-3
Annals of Mathematical Statistics ……… Vol. 11, 12
Annals of Engenies ……… Vol. 15, Part 1, 2, 3
Biometrika ……… Vol. 37
Reprints From Biometrika ……… Vol. 36, Parts 3-4
37, 1-2
37, 3-4
38, 1-2
の一部分
Biometrics ……… Vol. 6, No. 1-4
7, 1-3
Bulletin de la Société Mathématique de France ………
Tome 77 Fascicules 1-4
78 1-3,4
Calcutta Statistical Association ……… Vol. 3 No. 12
Cowles Commission for Research in Economics July 1, 1949 - June 30, 1950
Rational decision-making and Economic Behavior

Duke Mathematical Journal .... Vol. 17, No. 1-4
18. 1-4

Econometrica .... Vol. 17, No. 3,4
18, 1-4
19, 1-4

Economica .... Vol. 17, No. 65-68
18, No. 69-71

Indagationes Mathematicae ex actis quibus titulus ....
Vol. 53, Nos 1&2,
Nos. 9&10,
54, No. 1.
55, 2.
56, 3-4

ICAO Monthly Bulletin .... 1951 July, Sept., July-Aug., Oct.

ICAO Statistical Summary .... No. 1-4

Iowa State College .... Research Bulletin 380, 382.
Iowa Engineering Experiment Station Bulletin 166,
Proceedings of the Iowa Academy of Science Vol. 57

Journal of the American Statistical Association ....
Vol. 44, No. 247, 248,
45, 249, 252.
46, 253-256

Journal of the Royal Statistical Society ....
Series A, Vol. CXID
No. I, II, III.

Journal of the Royal Statistical Society ....
Series B, Vol. 12, No. 1-2

Kungl. Lantbruks-Högskolans Annaler ....
Vol. 17, Uppsala 1950

Kungl Tekniska Högskolans Handlingar ....
Nr 27, -30

Mathematical Reviews .... Vol. 11, No. 1-10
12, 1-10

Mitteilungsblatt Für Mathematische
Statistik ..... Jahrg 2, Heft 1-3
3, 2

Nieuw Archief voor Wiskunde ....
Deel 23, Derde Stuk

Ohio Biological Survey Bulletin 42 ..... Vol. 8, No. 2

Pacific Journal of Mathematics ..... Vol. 1, No. 1-4

Proceedings of the Cambridge Philosophical
Society ..... Vol. 45, Part 1-4
46, 1-4
47, 1-4

Publicationes Del Instituto De
Mathematica Y Estadistica
.... Vol. II, No. 1, 2

Rice Institute Pamphlet .... Vol. 38, No. 2, 3

Royal Agric. College & National Agr.
Research Centre ..... M. Sandelius 外3冊

SANKYĀ The Indian Journal of Statistics .... Vol. 9, Part 4.

10, 1-4

11, 1, 2

Reprints of the papers of the members of Statistical Laboratory, St. Andrew's .... Hill, Cambridge England 8冊

Statistics Upsale .... Vol. 2-4

Statistische Vierteljahresschrift ...... Band II, Heft 1.

III, 1-4

IV, 1, 2

Summa Brasiliensis Mathematicae ....

Vol. I, Ano. I, Fasc. 1-14

II, II, 1

II, III, 2, 3

II, IV, 4, 5

V, 6, 7

Summaries of Doctoral Dissertations ... Vol. 57, 1949,

58, 1950

The Bulletin of Mathematical Biophysics ....

Vol. 12, No. 1-4

13, 1-3

The Ohio Journal of Science .... Vol. 50, No. 6

51, 1-6

The University of Michigan .... Reprint 29冊

University of California Publication in Mathematics New Series .... Vol. 1, No. 1-8

2, 1

University of California Publication in
Statistics .... Vol. 1, No. 1-5
University of Illinois .... Reprint 11 冊
Wiskundige Opgaven Met De Oplossingen
Negentiende .... Dell Eerste stuk
Dell Tweede stuk
University of Washington Publications in
Mathematics .... Reprint 7 冊

METHODS OF CALCULATION

IN

QUANTIFICATION PROBLEMS

I

CONTENT

This is an issue of the projected series of reports entitled "The Research Report of the . S. M."
"The Research Report of the . S. M." publishes the reports of researches done in the application of Statistical Mathematics such as initial preparations, study designs, practical procedures and handling of data.

The series aims to be beneficial not only for the theoretical workers, but for research workers who are engaged in the practical problems of surveying, analysis and so on.

Editor	Chikio Hayashi
Published by	The Institute of Statistical Mathematics
	10, Sangenjaya-cho, Setagaya-ku, Tokyo
Printed by	Sobunsha Co.
	13, Takata-toyokawa-cho, Bunkyo-ku, Tokyo

The Research Report of the I.S.M.

Number 8

Methods of Calculation in Quantification Problems I

Feburary 1952

The Institute of Statistical Mathematics
10, Sangenjaya-cho, Setagaya-ku, Tokyo