Multidimensional Quantification*

— with the Applications to Analysis of Social Phenomena —

By Chikio HAYASHI

(Received February 20, 1954)

The present paper is a continuation of the papers [1], [2], previously published, in which we treated some methods of quantification of qualitative data in multidimensional analysis and especially the use of quantification of qualitative patterns to secure the maximum success rate of prediction of phenomena from the statistical point of view. The important problem in multidimensional analysis is to devise the methods of quantification of complex phenomena (intercorrelated behaviour patterns of units in dynamic environments) and then the methods of classification of them. Quantification means that the patterns are categorized and given numerical values in order that they may be able to be treated as several indices, and classification means a prediction of phenomena. The aim of multidimensional quantification is to make numerical representation of intercorrelated patterns synthetically to maximize the efficiency of classification (success rate of prediction). Quantification does not mean to find numerical values but to give them to the patterns from the operational point of view in the proper sense. In the present paper, the methods of quantification of qualitative patterns will be considered in case where an outside variable (realized by the outside criterion) is given in the form of qualitative classification. In this case it is most important that we must devise the methods to fulfil the property of validity. Let us take a universe of n elements, each of which has, as a label, behaviour patterns categorized by a survey method and is classified into only one class by the definite outside criterion of (this is an outside variable). Here the outside criterion must be based on the absolute scale and must not change according to what elements of universe are classified (judged). That is to say, $J(O_i)=J(O_j)=$constant independent of i, j; $i \neq j$, $i, j=1, 2, \cdots, n$ where $J(O_i)$ represents symbolically the frame of criterion for the i-th element O_i when it is classified (judged). It is our aim to predict to which class the element will belong in future which has a definite behaviour pattern at present, by the method of quantification using the past data.

---

* The theoretical parts of the present paper are the same as those in [4], and its aim is to describe their applications. The theoretical descriptions are repeated here for the convenience of understanding of the meanings of applications.

1. Case 1: The case where elements are classified in $S\ (\geqq 3)$ strata by an outside criterion which is unidimensional, that is to say, the so-called law of transitivity holds in the field where elements are judged by only one norm. Whenever $S=2$, this method is applicable. Each element has a response pattern in R items which have several sub-categories and the label of the stratum to which it belongs. Response pattern is represented by marked sub-categories in the items. The essential point of this method is the same as in [2, § 3]. Let $\{C_{11}, C_{12}, \ldots, C_{1K_1}\}$, $\{C_{21}, C_{22}, \ldots, C_{2K_2})$, $\ldots$, $\{C_{R_1}, C_{R_2}, \ldots, C_{RK_R}\}$ be sub-categories in all items. Let us consider a way of giving a numerical value x_{lm} to the m-th sub-category in the l-th item, C_{lm}, from the mathematico-statistical point of view. It is sometimes desirable that quantitative data are expressed in the forms of responce patterns mentioned above. Response patterns of n elements are, for example, as follows.

Response patterns (behaviour patterns) of elements

By outside criterion, outside variable ↓ Number of stratum	Item: 1 … R / Sub category \ Element	C_{11}	C_{12}	$\cdots$	C_{1K_1}	....	C_{R1}	C_{R2}	$\cdots$	C_{RK_R}
1	1	V				....	V			
	2		V			....		V		
	⋮					⋮				
	n_1		V			....		V		
⋮	⋮					⋮				
S	1				V	....		V		
	⋮					⋮				
	n_S				V	....				V

V sign means the check in response of elements.

n_t is the number of elements belonging to the t-th stratum, where $n=\sum_{t=1}^{s} n_t$. Let $\{X_{1(i)}, X_{2(i)}, \ldots, X_{R(i)}\}$ be the respone patterns of element i, where $X_{j(i)}$

denotes the sub-category of the j-th item that element i marks. Now we use the score $a_i = x_{1(i)} + x_{2(i)} + \cdots x_{R(i)}$ as the score of element i where $x_{j(i)}$ is the numerical value given to the sub-category of the j-th item which element i marks. The linear form is considered to be appropriate, according to the outside criterion being unidimensional and the idea of the first approximation. This linear form is not so restricted, from the point of view of making the data (having not linear relations) linear by quantifying qualitative patterns.

We have $\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(a_i - \bar{a})^2$ as the total variance with respect to elements, where $a = \frac{1}{n}\sum_{i=1}^{n} a_i$. Now we want to quantify the sub-categories (items) so as to maximize the effect of stratification, that is, so as to maximize the correlation ratio $\eta^2 = \frac{\sigma_b^2}{\sigma^2}$, where σ_b^2 is the variance between strata. This is a reasonable method of quantification, because η^2 is a measure of discriminative power of the items, i.e., a measure of efficiency of the classification (success rate of prediction).

If η^2 is large in the result of the quantification, we can treat quantitatively the behaviour patterns by using x_{lm} (or a). Thus we can introduce a metric into qualitative patterns and define the distances between qualitative patterns (data) using the obtained values x_{lm}, and, so to speak, obtain the functional form of them which is valid in the above sense. To obtain x_{lm} which maximize η^2, let us introduce the following definitions.

Set

$$\delta_i(jk) = \begin{cases} 1, & \text{if element } i \text{ marks the } k\text{-th sub-category in the } j\text{-th item,} \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$a_i = \sum_{j=1}^{R}\sum_{k=1}^{K_j} \delta_i(jk) x_{jk}$$

And

$$\bar{a} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{R}\sum_{k=1}^{K_j} \delta_i(jk) x_{jk},$$

$$\sigma^2 = \frac{1}{n}\left(\sum_{i=1}^{n}\sum_{j=1}^{R}\sum_{k=1}^{K_j} \delta_i(jk) x_{jk}\right)^2 - \bar{a}^2 = \frac{1}{n}\left(\sum_{j=1}^{R}\sum_{k=1}^{K_j} n_{jk} x_{jk}^2 + \sum_{l}{}'\sum_{m}{}'\sum_{j=1}^{R}\sum_{k=1}^{K_j} f_{jk}(lm) x_{jk} x_{lm}\right) - \bar{a}^2$$

where

$$n_{ik} = \sum_{i=1}^{n} \delta_i(jk), \qquad f_{jk}(lm) = \sum_{i=1}^{n} \delta_i(jk)\delta_i(lm),$$

($f_{jk}(lm)$ represents the correlated pattern between responses in items of each element),

$\sum_l' \sum_m' \sum_{j=1}^{R} \sum_{k=1}^{K_j}$ covers all range of l, m, j, k, except when $l=j$, $m=k$ hold simultaneously. Further

$$\sigma_b^2 = \sum_{t=1}^{s} (\bar{a}_t - \bar{a})^2 \frac{n_t}{n}$$

where

$$\bar{a}_t = \frac{1}{n} \sum_{j=1}^{R} \sum_{k=1}^{K_j} g^t(jk) x_{jk},$$

$$g^t(jk) = \sum_{i(t)=1}^{n_t} \delta_{i(t)}(jk), \qquad n_{jk} = \sum_{t=1}^{s} g^t(jk),$$

$\delta_{i(t)}(jk)$ means $\delta_i(jk)$ which element i belonging to the t-th stratum has.
Thus we have

$$\eta^2 = \frac{\sigma_b^2}{\sigma^2}$$

To maximize η^2 with respect to x_{uv} $(u=1, 2, \ldots, R,\ v=1, 2, \ldots, K_u)$, put $\frac{\partial \eta^2}{\partial x_{uv}} = 0$ that is, $\frac{\partial \sigma_b^2}{\partial x_{uv}} = \eta^2 \frac{\partial \sigma^2}{\partial x_{uv}}$. In calculating this, we can assume $\bar{a}=0$ without loss of generality, and, as will be easily shown, we have

$$\frac{\partial \sigma_b^2}{\partial x_{uv}} = \frac{2}{n} \sum_{j=1}^{R} \sum_{k=1}^{K_j} h_{uv}(jk) x_{jk}$$

$$\frac{\partial \sigma^2}{\partial x_{uv}} = \frac{2}{n} \left(\sum_{l=1}^{R} \sum_{m=1}^{K_l} f_{uv}(lm) x_{lm} \right)$$

where

$$h_{uv}(jk) = \sum_{t=1}^{s} \frac{g^t(jk) g^t(uv)}{n_t}$$

Then

$$\sum_{j=1}^{R} \sum_{k=1}^{K_j} h_{uv}(jk) x_{jk} = \eta^2 \sum_{l=1}^{R} \sum_{m=1}^{K_l} f_{uv}(lm) x_{lm} \quad (u=1, 2, \ldots, R,\ v=1, 2, \ldots, K_u)$$

Let H be the matrix $(h_{uv}(jk))$, F be the matrix $(f_{uv}(lm))$, X be vector (x_{jk}). The above equation is written as follows.

$$HX = \eta^2 FX$$

It is our problem to solve this under the conditions $\sum_{k=1}^{K_j} n_{jk} x_{jk} = 0$, $(j=1, 2, \ldots, R)$, and to require the largest maximum value (this is the largest value) of η^2 which is not equal to 1 and the corresponding vector X to it. It is easily proved that $\eta^2 (0 \leqq \eta^2 \leqq 1)$ satisfying the above equations exists and the value we require exists, for the quadratic forms from symmetric matrices H, F are positive definite. We can obtain the required x_{lm} by the successive approximation method.

The method of classification (prediction) by quantified behaviour patterns and its efficiency are described in [2, § 3], because it is considered that the stratum means the label of outcome of the element in future.

It is generally proved that the success rate of prediction is a monotone increasing function of η, ([6]). Thus the maximization of η^2 has the valid sense.

The methods of calculation to solve the equations in case 1 and the following case 2 have been devised by Mr. H. Akaike and the complicated calculations of examples have been done by Miss M. Taguma, Miss S. Takakura and Miss Y. Saegusa.

[**Example**]

(i) Decision of status.

This is a problem to decide the status of the resident in a city. Status must be considered in the society (social field) where he leads a life. In many cases the status judged by us is different from the status judged by the residents. The latter is meaningful. It is valid to decide the status of a resident by the judgement of some experts resident in the city. But they do not know all residents at the same degree and can not judge the stata of residents in the valid sense. Thus we adopt the sub-sampling method, regarding the small area as a primary sampling unit. Experts know the residents (samples) in the small area at the same degree and can judge their stata.

It is our purpose to decide (estimate) the stata of all residents by analyzing the structure of judgement of experts. This procedure means that we estimate the stata which would be obtained if the residents were judged by the same criteria as those of the experts. For this purpose we inquire into the following four items which have six sub-categories respectively; income, occupation, appearance of house, and the experience as a staff. Each sample has a label of status judged by experts which is classified into six strata and the results to the above inquiries. The former is the outside variable and the latter is the response pattern in the above items. It is our problem to represent the outside variable by the response pattern at the high confidence level. In this case the relation $J(O_i)=J(O_j)$ holds.

We want to require the numerical values of sub-categories in items to maximize the success rate of decision (estimation) of status by the above theory and the above data. In calculating, six strata and six sub-categories are reduced to three strata and three sub-categories, respectively, in order to increase the coefficient of reliability.

Response Patterns in Total

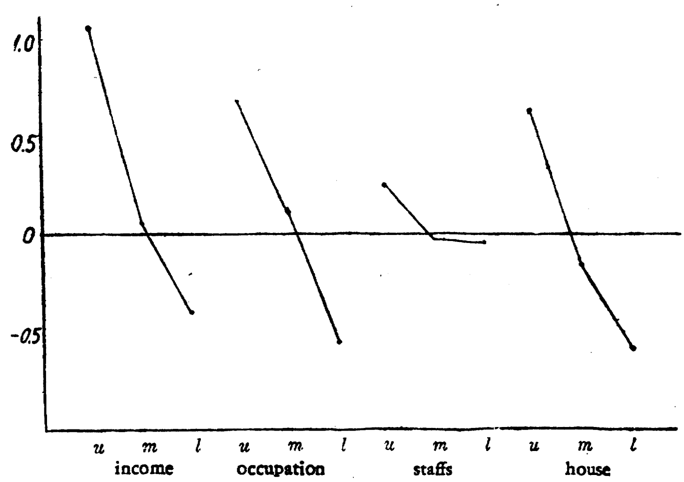
		Income			Occupation			Staff			House			Status		
		u	*m*	*l*	*u*	*m*	*l*	*u*	*m*	*l*	*u*	*m*	*l*	*u*	*m*	*l*
Income	*u*	15	0	0	5	9	1	2	11	2	9	5	1	9	6	0
	m		54	0	13	31	10	7	21	26	18	30	6	7	41	6
	l			47	1	20	26	3	14	30	6	31	10	3	19	25
Occupation	*u*				19	0	0	7	8	4	13	6	0	10	9	0
	u					60	0	3	28	29	17	37	6	8	43	9
	l						37	2	10	25	3	23	11	1	14	22
Staff	*u*							12	0	0	7	5	0	5	6	1
	m								46	0	17	24	5	11	25	10
	l									58	9	37	12	3	35	20
House	*u*										33	0	0	14	19	0
	m											66	0	5	40	21
	l												17	0	7	10
Status	*u*													19	0	0
	m														66	0
	l															31

u···upper; *m*···middle; *l*···lower.

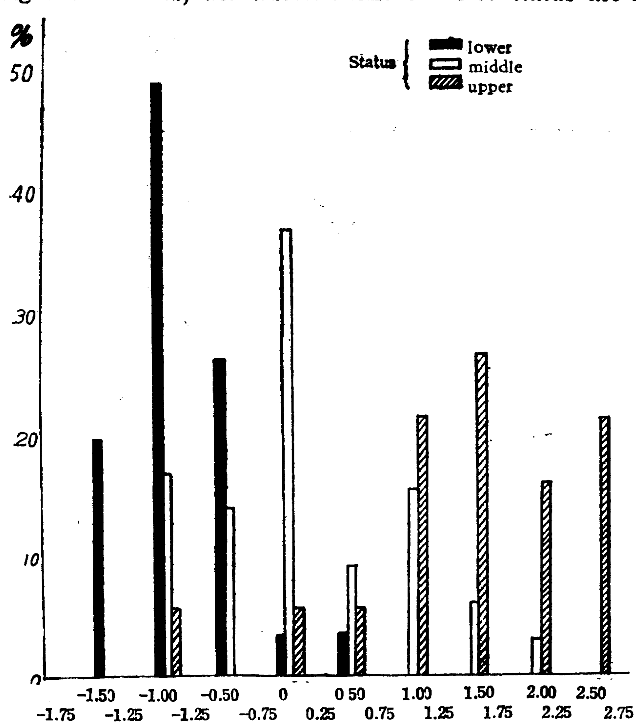
Thus, $\eta=0.71$, and the numerical values are as below.

Income			Occupation			Staff			House		
u	*m*	*l*	*u*	*m*	*l*	*u*	*m*	*l*	*u*	*m*	*l*
1.066	0.056	−0.404	0.690	0.120	−0.550	0.265	−0.025	−0.035	0.640	−0.170	−0.590

The order of effectiveness ($x_{\text{upper}}-x_{\text{lower}}$) is; income (1.47), occupation (1.24), appearance of a house (1.22).



Using these values, the distributions of three status are described.



Mean values	
lower	−0.97
middle	+0.06
upper	+1.35

The dividing points by which one can judge to what status a resident belong using his numerical value (score, α_i), are obtained by the max-min theory.

	Dividing points
upper	$x \geqq 0.95$
middle	$0.95 > x \geqq -0.74$
lower	$-0.74 > x$

The success rate of judgement is 0.707. Next let us require the dividing points by the method of [3, § 2], approximating the histograms by smooth curves.

	Dividing points
upper	$x \geqq 1.2$
middle	$1.2 > x \geqq -0.49$
lower	$-0.49 > x$

The success rate is 0.716. The comparison of judgements by this method with our judgements are as below.

Sample	Judgement by this method	Our judgement (3 persons)			Sample	Judgement by this method	Our judgement (3 persons)		
		I	S_l	S_b			I	S_l	S_b
1	m	l	l	u	13	l	l	m	l
2	m	m	m	m	14	l	l	l	l
3	u	m	u	u	15	m	m	l	m
4	l	l	l	m	16	l	l	l	l
5	u	l	l	m	17	u	m	u	u
6	m	m	u	m	18	m	m	l	m
7	u	u	u	u	19	m	m	u	m
8	m	l	m	l	20	u	m	u	m
9	m	l	m	m	21	m	l	m	l
10	l	l	m	l	22	m	m	m	u
11	m	l	m	m	23	m	l	u	m
12	m	u	m	m	24	m	u	l	m

u···upper; m···middle; l···lower.

The above relations are interesting.

(ii) The following questionnaire is given.

There is an opinion: If there were a great politician, we would confidently leave affairs of the state to him for the reconstruction of Japan without arguing with each other. Do you agree with it?

1. Yes 2. On some occasions 3. No

This response is an outside variable. We consider to what extent the outside variable will be decided by the following five characteristics; residence (urban, rural), sex (male, female), age (20～24, 25～29, 30～39, 40～49, 50 over), political party to support (conservative, progressive, no, others), schooling (no, primary school, middle school, high school, college over).

Each sample has a label of the outside variable (i.e., the response to the questionnaire) and a response pattern in the above five items. $J(O_i)=J(O_j)$ holds evidently. The sample has been obtained by the stratified random sub-sampling method from the universe of the Japanese, in April, 1953. The size is about 2000.

The cross tabulations are the following.

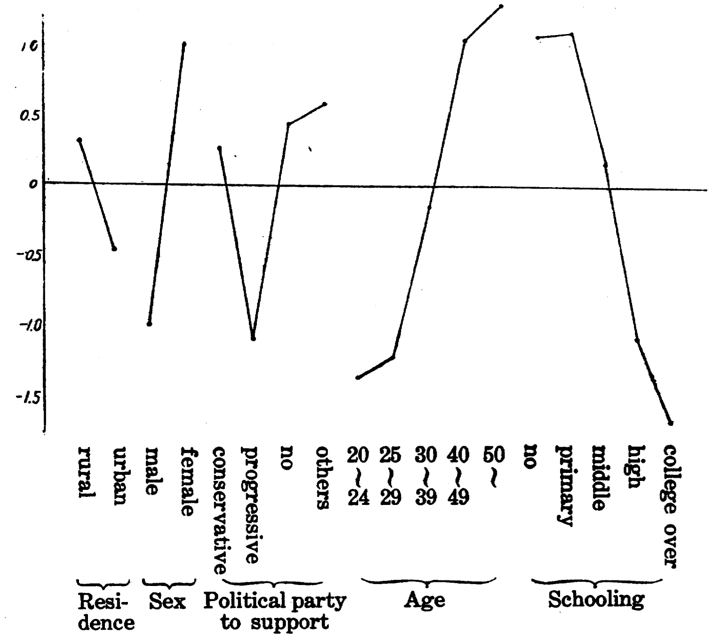
		Residence		Sex		Political party to support				Age					Schooling					Responce		
		rural	urbran	male	female	conservative	progressive	no	others	20～24	25～29	30～39	40～49	50～	no	primary	middle	high	college over	yes	medium	no
Residence	rural	1222	0	607	615	532	246	253	191	234	189	265	224	310	67	381	497	232	45	635	114	473
	urban		803	387	416	314	255	144	90	159	126	193	148	177	29	151	243	283	97	321	82	400
Sex	male			994	0	456	281	180	77	189	163	195	178	269	25	220	415	228	106	374	110	510
	female				1031	390	220	217	204	204	152	263	194	218	71	312	325	287	36	582	86	363
Political party to support	conservative					846	0	0	0	121	136	186	171	232	24	239	308	218	57	421	81	344
	progressive						501	0	0	144	94	128	80	55	7	74	198	167	55	159	43	299
	no							397	0	84	57	92	58	106	22	120	150	87	18	206	43	148
	others								281	44	28	52	63	94	43	99	84	43	12	170	29	82
Age	20～24									393	0	0	0	0	1	41	164	153	34	125	39	229
	25～29										315	0	0	0	1	48	136	103	27	108	30	177
	30～39											458	0	0	8	109	181	132	28	212	39	207
	40～49												372	0	11	129	127	78	27	209	42	121
	50～													487	75	205	132	49	26	302	46	139

		Residence		Sex		Political party to support				Age					Schooling					Responce		
		rural	urbran	male	female	conservative	progressive	no	others	20∼24	25∼29	30∼39	40∼49	50∼	no	primary	middle	high	college over	yes	medium	no
Schooling	no														96	0	0	0	0	68	10	18
	primary															532	0	0	0	339	42	151
	middle																740	0	0	342	70	328
	high																	515	0	169	66	280
	college over																		142	38	8	96
Responce	yes	635	321	374	582	421	159	206	170	125	108	212	209	302	68	339	342	169	38			
	medium	114	82	110	86	81	43	43	29	39	30	39	42	46	10	42	70	66	8			
	no	473	400	510	363	344	299	148	82	229	177	207	121	139	18	151	328	280	96			

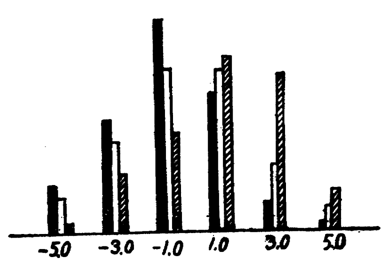
The values given by the above method are as follows. $\eta=0.383$.

Residence		Sex		Political party to support			
rural	urban	male	female	conservative	progressive	no	others
0.305	−0.465	−1.018	0.982	0.270	−1.120	0.430	0.580

Age				
20∼24	25∼29	30∼39	40∼49	50∼
−1.392	−1.242	−0.162	1.038	1.288
Schooling				
no	primary	middle	high	college over
1.074	1.094	0.154	−1.086	−1.686



The distributions of scores in the strata by the outside variable are



	no	on some occasions	yes
mean	−1.03	−0.23	0.91

The dividing points are obtained if we neglect the strata of "on some occasions"

Responce	Dividing point
yes	$x \geqq 0.20$
no	$x < 0.20$

The success rate of judgement is 0.690. The success rate of predicting to which stratum he belongs (to which category he responses in the questionnaire item) is 0.613 in the whole.

2. Case 2: The case where elements are classified into S strata by an outside criterion which is not unidimensional. Each element has a response pattern in R items and has a label of the stratum, which is called an outside variable. We use the same symbols as in 1. In this case we consider the R-dimensional Euclidean space, whose coordinates correspond to R items. Each element will be represented as a point in this space if sub-categories in items are to be quantified. Now we want to quantify the sub-categories C_{lm} so as to maximize the effect of stratification. Here it is not at all effective to use the method of case 1. As the total variance σ^2, we take the generalized variance which is considered to be proportionate to the square of the volume of the so-called ellipsoid of concentration. The validity of using this generalized variance is secured by Chebyshev's inequality of R dimensions, because the ellipsoid which contains a definite probability and has the minimum volume, is given by the so-called ellipsoid of concentration, and the volume is proportionate to the generalized variance. The generalized variance is $\sigma^2=|\rho_{jl}\sigma_j\sigma_l|$, where $\rho_{jl}\sigma_j\sigma_l$ is the covariance between the j-th and the l-th item (dimension) when the sub-categories are quantified and $|\cdots|$ expresses a determinant, element of which is $(\rho_{jl}\sigma_j\sigma_l \;\; j, l=1, 2, \ldots, R.)$. As the within variance, we take $\sigma_t^2=|\sigma_{jl}(t)\sigma_j(t)\sigma_l(t)|$, where $\rho_{jl}(t)\sigma_j(t)\sigma_l(t)$ is the covariance between the j-th and the l-th item (dimention) in the t-th stratum, which is deeply related, in the above sense, to the ellipsoid of concentration in the t-th stratum. Thus we take $\mu=1-\frac{\sum_{t=1}^{s} p_t\sigma_t^2}{\sigma^2}$ as a measure of the efficiency of stratification, i.e., a measure of the discriminative power of items which will be an index related to the efficiency of classification (success rate of prediction), by quantifying behaviour patterns. Here p_t is a weight assigned to the t-th stratum, and $\sum_{t=1}^{s} p_t=1$; especially, we put $p_t=\frac{n_t}{n}$, where n_t is the size of the t-th stratum, and we have clearly $\sum_{t=1}^{s} n_t=n$. $\sum p_t\sigma_t^2$ is considered to be a sort of within variances in the whole. If $\sigma_t^2=0$, we have $\mu=1$, and if $\sigma_t^2=\sigma^2$, $\mu=0$. Besides this, we can take several indices as a measure. One example of these will be shown later on.

Now we require x_{lm} given to C_{lm} to maximize μ. Put $\frac{\partial\mu}{\partial x_{uv}}=0$, $u=1, 2, \ldots, R$, $v=1, 2, \ldots, K_u$. Then we obtain

$$\sum_{t=1}^{s} p_t\frac{\partial\sigma_t^2}{\partial x_{uv}}=(1-\mu)\frac{\partial\sigma^2}{\partial x_{uv}} \quad \cdots(1).$$

It is our problem to solve these equations and to obtain the largest maximum value of $\mu(\neq 1)$ (which is the largest value of μ) and the corresponding vectors. Let $\bar{x}_j$ be the total mean of the j-th item, $\bar{x}_j=\frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{K_j}\delta_i(jk)x_{jk}=\frac{1}{n}\sum_{k}^{K_j}n_{jk}x_{jk}$, σ_j^2 the total variance of the j-th item: $\sigma_j^2=\frac{1}{n}\sum_{k=1}^{K_j}n_{jk}x_{jk}^2-\bar{x}_j^2$, $\bar{x}_j(t)$ the mean of the j-th item in the t-th stratum: $\bar{x}_j(t)=\frac{1}{n_t}\sum_{k=1}^{K_j}g^t(jk)x_{jk}$, and $\sigma_j(t)^2$ the variance of the j-th item in the t-th stratum: $\sigma_j(t)^2=\frac{1}{n_t}\sum_{k=1}^{K_j}g^t(jk)x_{jk}^2-\bar{x}_j(t)^2$. Without loss of generality, we can assume $\bar{x}_j=0$, $j=1,2,\ldots,R$. Then $\sigma_j^2=\frac{1}{n}\sum_{k=1}^{K_j}n_{jk}x_{jk}^2$. Let $\rho_{jl}\sigma_j\sigma_l$ be the total covariance between j and l, $\rho_{jl}\sigma_j\sigma_l=\frac{1}{n}\sum_{k}^{K_j}\sum_{m}^{K_l}f_{jk}(lm)x_{jk}x_{lm}$; $\rho_{jl}(t)\sigma_j(t)\sigma_l(t)$ be the covariance in the t-th stratum, $\rho_{jl}(t)\sigma_j(t)\sigma_l(t)=\frac{1}{n_t}\sum_{k}^{K_j}\sum_{m}^{K_l}f_{jk}^t(lm)\,x_{jk}x_{lm}-\bar{x}_j(t)\bar{x}_l(t)$

where

$$f_{jk}^t(lm)=\sum_{i(t)=1}^{n_t}\delta_{i(t)}(jk)\delta_{i(t)}(lm), \quad \sum_{t=1}^{s}f_{jk}^t(lm)=f_{jk}(lm).$$

Then

$$\sigma^2=\left|\frac{1}{n}\sum_{k}^{K_j}\sum_{m}^{K_l}f_{ik}(lm)\,x_{jk}x_{lm}\right| \qquad \ldots\ldots\ldots\ldots\ldots\ldots (2)$$

$$\sigma_t^2=\left|\frac{1}{n_t}\sum_{k}^{K_j}\sum_{m}^{K_l}f_{jk}^t(lm)\,x_{jk}x_{lm}-\bar{x}_j(t)\bar{x}_l(t)\right| \qquad \ldots\ldots\ldots (3)$$

From (1), (2), (3), we can obtain the numerical values x_{lm} ($l=1,2,\ldots,R$, $m=1,2,\ldots,K_l$) which maximize μ.

Special Case: Let $R=3$.

$$\sigma^2=\begin{vmatrix}\sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3\\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3\\ \rho_{13}\sigma_1\sigma_3 & \rho_{23}\sigma_2\sigma_3 & \sigma_3^2\end{vmatrix}$$

Set

$$\tau_{\beta\gamma}^2=\sigma_\beta^2\sigma_\gamma^2-\rho_{\beta\gamma}^2\sigma_\beta^2\sigma_\gamma^2 \qquad (\beta,\gamma=1,2,3),$$

$$\begin{aligned}\gamma_{\alpha,ku}=&S_k(\alpha\gamma)S_u(\alpha\gamma)\sigma_\beta^2+S_k(\alpha\beta)S_u(\alpha\beta)\sigma_\gamma^2\\ &-\{S_u(\alpha\beta)S_k(\alpha\gamma)+S_u(\alpha\gamma)S_k(\alpha\beta)\}\rho_{\beta\gamma}\sigma_\beta\sigma_\gamma,\\ &\alpha\neq\beta,\quad \alpha\neq\gamma,\quad \beta\neq\gamma.\end{aligned}$$

$$S_k(\delta\varepsilon)=\frac{1}{n}\sum_m^{K_\varepsilon} f_{\varepsilon k}(\varepsilon m)\, x_{\varepsilon m}\,, \qquad \delta,\ \varepsilon=1, 2, 3.$$

$$p_{\alpha u}=\frac{n_{\alpha u}}{n},$$

and

$$\tau_{\beta\gamma}(t)^2=\sigma_\beta(t)^2\sigma_\gamma(t)^2-\rho_{\beta\gamma}(t)^2\sigma_\beta(t)^2\sigma_\gamma(t)^2\,,$$

$$\gamma^t_{\alpha,ku}=S^t_k(\alpha\gamma)S^t_u(\alpha\gamma)\sigma_\beta(t)^2+S^t_k(\alpha\beta)S^t_u(\alpha\beta)\sigma_\gamma(t)^2$$
$$-\{S^t_u(\alpha\beta)S^t_k(\alpha\gamma)+S^t_u(\alpha\gamma)S^t_k(\alpha\beta)\}\,\rho_{\beta\gamma}(t)\sigma_\beta(t)\sigma_\gamma(t)$$

$$S^t_k(\delta\ \varepsilon)=\frac{1}{n_t}\sum_{m=1}^{K_\varepsilon} f^t_{\delta k}(\varepsilon\ m)\, x_{\varepsilon m}-\frac{g^t(\delta k)}{n_t}\,\bar{x}_\varepsilon(t)$$

$$p^t_{\alpha k}=\frac{g^t(\alpha k)}{n_t}, \quad \alpha,\ \beta,\ \gamma=1,\ 2,\ 3, \quad \alpha\neq\beta,\ \alpha\neq\gamma,\ \beta\neq\gamma.$$

Then, in the long run, we obtain

$$\sum_{k=1}^{K_1} \Lambda_{1,ku}x_{1k}=\mu\sum_{k=1}^{K_1}\Phi_{1,ku}x_{1k} \qquad (u=1, 2, \ldots, K_1) \quad (*)$$

$$\sum_{m=1}^{K_2} \Lambda_{2,mv}x_{2m}=\mu\sum_{m=1}^{K_2}\Phi_{2,mv}x_{2m} \qquad (v=1, 2, \ldots, K_2) \quad (\overset{*}{*})$$

$$\sum_{l}^{K_3} \Lambda_{3,lw}x_{3l}=\mu\sum_{l=1}^{K_3}\Phi_{3,lw}x_{3l} \qquad (w=1, 2, \ldots, K_3) \quad (\overset{*}{\underset{*}{*}})$$

where

$$\Lambda_{\alpha,ku}=\bar{\gamma}_{\alpha,ku}+\varphi_{\alpha,ku}-\gamma_{\alpha,ku}+\delta_{ku}(p_{\alpha u}\tau^2_{\beta\gamma}-\psi_{\alpha u})$$

$$\Phi_{\alpha,ku}=\delta_{ku}p_{\alpha u}\tau^2_{\beta\gamma}-\gamma_{\alpha,ku}$$

$$\varphi_{\alpha,ku}=\sum_{t=1}^{s} p_t p^t_{\alpha u}p^t_{\alpha k}\tau_{\beta\gamma}(t)^2\,,$$

$$\bar{\gamma}_{\alpha,ku}=\sum_{t=1}^{s} p_t\gamma^t_{\alpha,ku}\,,$$

$$\psi_{\alpha u}=\sum_{t=1}^{s} p_t p^t_{\alpha u}\tau_{\beta\gamma}(t)^2$$

$$\alpha=1, 2, 3 \qquad \alpha\neq\beta, \quad \beta\neq\gamma, \quad \alpha\neq\gamma$$

δ_{ku} is Kronecker's symbol ($\delta_{ku}=1$, $k=u$; $=0$, $k\neq u$). $\Lambda_{1,ku}$ is symmetric and quadratic form from symmetric matrix ($\Phi_{\alpha,ku}$) is positive definite (equal to variance) in a sense. Therefore it is shown that the required μ exists.

Numerical values x_{lm} which are obtained in the valid sense (taking appropriate unit), by this method, have the same meanings and function as in case **1.** Moreover we can give the distances between the strata explicitly in the senses which will allow useful interpretation. The above equations can be solved by a successive approximation, "step by step in turn" method. Thus we can generally obtain the required results. As another measure, we can also take $\lambda=1-\frac{\sigma_w^2}{\sigma^2}$, where $\sigma_w^2=|\sum p_t\rho_{ji}(t)\sigma_j(t)\sigma_i(t)|$, and σ^2 is generalized variance, which is suggested by Mr. Akaike. The method of solving the equations is the same as above. Generally it turns out to be better that we adopt, as the first approximation, the values of x_{lm} obtained by solving the following equations, for example,

$$\sum_{k}^{K_a}\sum_{t=1}^{s}\frac{g^t(ak)g^t(au)}{n_t}x_{ak}=\kappa^2 n_{au}x_{au}, \qquad (a=1, 2, 3)$$

which means the quantification in case where each dimension is treated separately. The method of prediction is similar to case 1. The essential point of this method consists in making no use of the sum of the response patterns (additivity). This is also applicable to the case where the method of case 1 is not effective even when the outside criterion is unidimensional or $S=2$. Mr. Midzuno has devised an interesting method [5] in the similar problems.

[**Example**] The aim of this research is to clarify the relation between home education and personality of child. This survey has been carried by Mr. Mizuhara at the Ochanomizu Joshi University.

Personality of child is classified into four types, by teachers' observations: I-type, adaptable and group-played; II-type, adaptable and isolate; III-type, recalcitrant and group-played; IV-type, recalcitrant and isolate. This is an outside variable in our case. Evidently these four types are not unidimensional. Response patterns are given by marking the items of questionnaires made to obtain the informations of home education. We want to get the numerical values given to sub-categories of the above items to maximize the discriminative power among the personality types, that is equivalent to maximization of success rate of predicting the personality type of a child by the informations concerning the home education given to it. We want to get the numerical values from this stand point. So they must be interpreted from this point. As an index of home education, the following three dimensions have been taken into consideration: democratic-autocratic, lenient-strict, affectionate-non-affectionate. The three items corresponding to the above three dimensions have three sub-categories respectively: democratic, medium, autocratic; lenient, medium, strict; affectionate, medium, non-affectionate.

We consider three dimensional space. Each of three orthogonal axes corresponds to one dimension. Each sample has a label of personality type as an outside variable and a response pattern in the three items. The cross tabulations of a sample are as follows.

		Democratic (*X*)			Lenient (*Y*)			Affectionate (*Z*)		
		demo-cratic	medium	auto-cratic	lenient	medium	strict	affec-tionate	medium	non-affec-tionate
Democratic (*X*)	dem.	31	0	0	9	15	7	12	14	5
	med.		30	0	6	12	12	5	15	10
	aut.			19	6	6	7	3	7	9
Lenient (*Y*)	leni.				21	0	0	7	10	4
	medi.					33	0	9	13	11
	stri.						26	4	13	9
Affectionate (*Z*)	affec.							20	0	0
	med.								36	0
	non-affect.									24

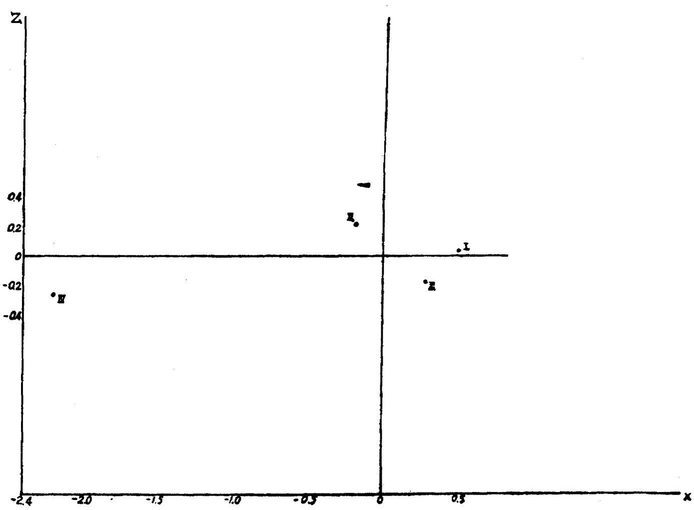
Thus using the theory of case 2, $x_{lm}s$ are obtained. In this case the numerical value given to the sub-category of "medium" in each dimension is assumed to be unity which does not mean any loss of generality. This is considered to be valid in interpretation, because the relations between qualitative ratings (qualitative response patterns) and numerical values given to them are important. Besides this, it is also meaningful to assume that the numerical values given to "medium" are equall to unity and the values to "autocratic" (strict, non-affectionate) are equall to zero. The numerical values are as follows, $\mu=0.902$.

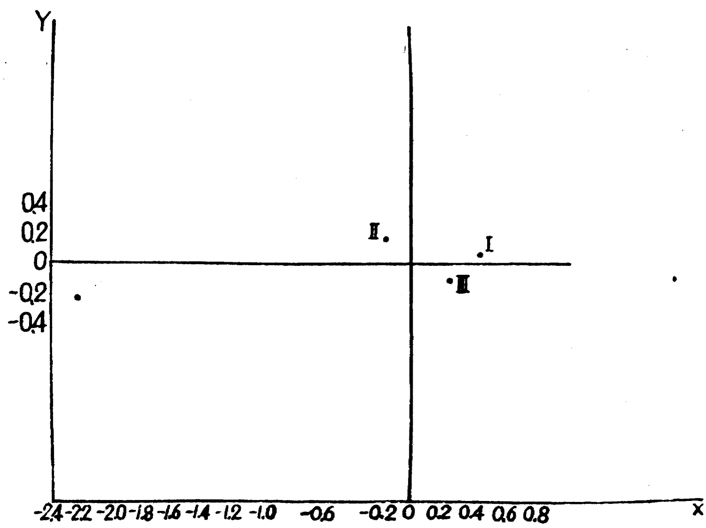
Democratic (*X*)			Lenient (*Y*)		
democratic	medium	autocratic	lenient	medium	strict
1.688	1	−4.333	1.053	1	−2.120

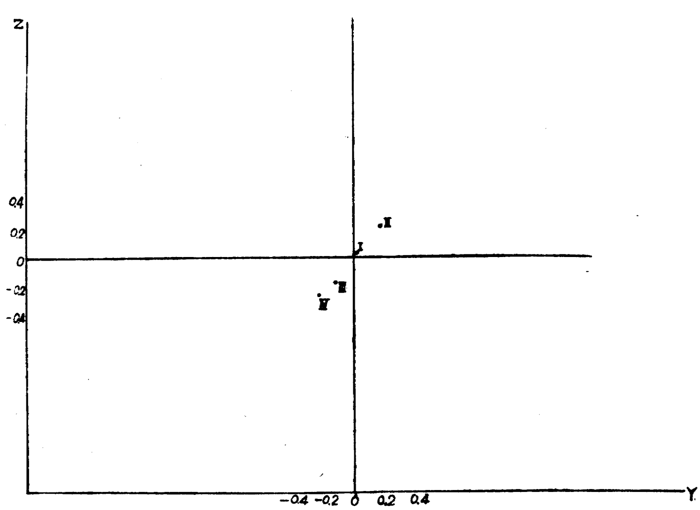
Affectionate (*Z*)		
affectionate	medium	non-effectionate
0.032	1	−1.527

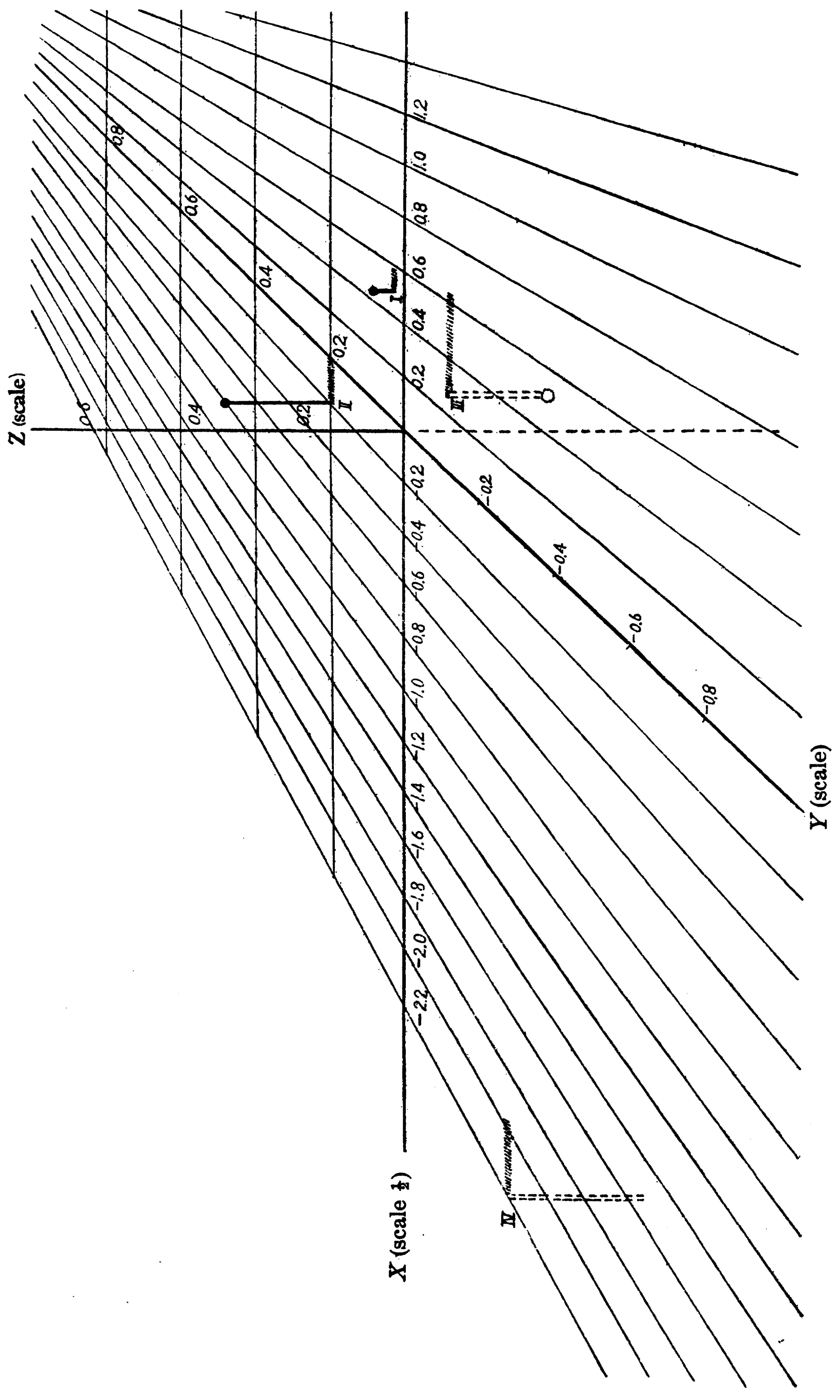
The mean values of the personality types are:

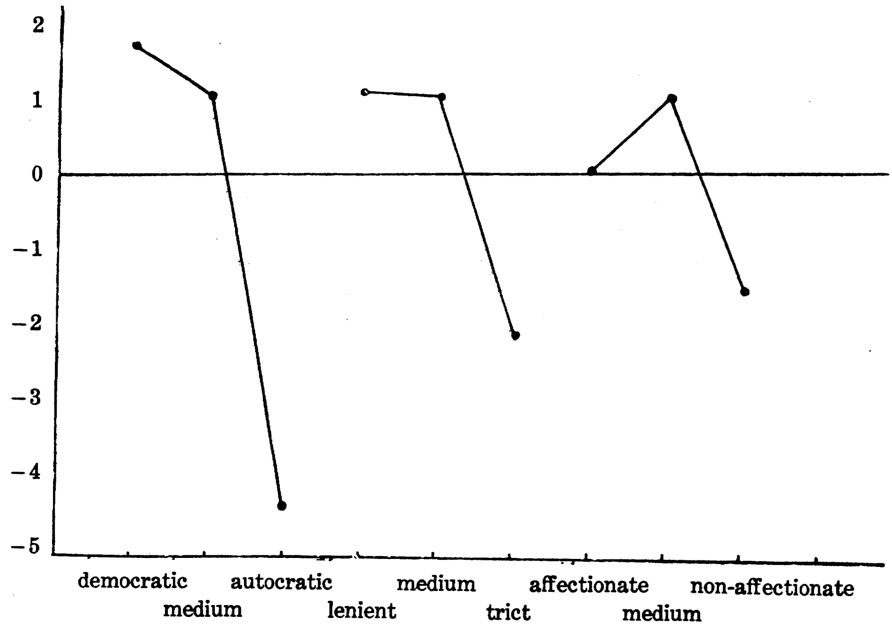
Personality type	Democratic (X)	Lenient (Y)	Affectionate (Z)
I	0.496	0.022	0.033
II	−0.193	0.186	0.211
III	0.281	−0.125	−0.183
IV	−2.200	−0.237	−0.263




Y
0.4
0.2
0
-0.2
-0.4
-2.4 -2.2 -2.0 -1.8 -1.6 -1.4 -1.2 -1.0 -0.6 -0.2 0 0.2 0.4 0.6 0.8
X


Z
0.4
0.2
0
-0.4 -0.2 0 0.2 0.4
Y

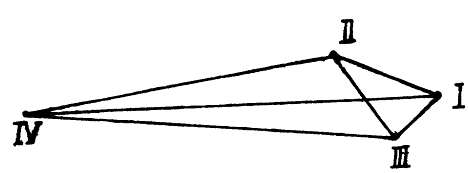

Z (scale)
Y (scale)
X (scale ½)
I
II
III
IV



The distances between means of the personality types are:

	I	II	III	IV
I	0	0.730	0.338	2.725
II		0	0.690	2.106
III			0	2.485
IV				0

In this case, the mean values of the four personality types approximately turns out to be on a plane.



The relations between home education and the personality types will be clear from this angle. The numerical values (i.e., weights in a sense) given to the sub-categories, the mean values and μ have interesting meanings. But we must not forget the idea underlying the above method.

3. Case 3: The compound case. We divide R items into T sub-groups. Suppose that, in each sub-group, the items belonging to it fulfil the property of additivity in the valid sense (see case 1), i.e. they are on a unidimensional scale. In this case we regard each element as a point in the T dimensional Euclidean space if sub-categories of items are quantified and synthesised in sub-groups. We set the T orthogonal coordinate axes that have the same meanings as in case 2. In each dimension, we quantify sub-categories of items under the assumption of linear form by the same idea as in case 1, but the numerical values must be required in correlation with the whole. Elements are classified into S strata by an outside criterion which is not unidimensional.

Let $R_1, R_2, \ldots, R_T$ be the numbers of items in sub-groups, respectively, $\sum_{r=1}^{T} R_r = R$.

Let $\sum_{u_r}^{R_r}\sum_{v_r}\delta_i(u_rv_r)x_{u_rv_r}=a_i(r)$ be the score of the element i in the r-th sub-group. Where $u_r, v_r\ldots$ express the items and sub-categories belonging to the r-th sub-group and the symbols have the same meaning as in cases 1 and 2; the symbols in this section are the same.

Thus
$$\bar{x}=\frac{1}{n}\sum_{i}^{n}a_i(r)=\frac{1}{n}\sum_{u_r}^{R_r}\sum_{v_r}n_{u_rv_r}x_{u_rv_r},$$
$$\sigma_r^2=\frac{1}{n}\sum_{l_r}^{R_r}\sum_{m_r}\sum_{u_r}^{R_r}\sum_{v_r}f_{u_rv_r}(l_rm_r)\,x_{u_rv_r}x_{l_rm_r}-\bar{x}_r^2 \quad \text{in the total,}$$
$$\bar{x}_r(t)=\frac{1}{n_t}\sum_{u_r}^{R_r}\sum_{v_r}g^t(u_rv_r)x_{u_rv_r},$$
$$\sigma_r(t)^2=\frac{1}{n_t}\sum_{l_r}^{R_r}\sum_{m_r}\sum_{u_r}^{R_r}\sum_{v_r}x_{u_rv_r}x_{l_rm_r}f_{u_rv_r}(l_rm_r)-\bar{x}_r(t)^2$$
in the t-th stratum. Without loss of generality, we can assume $\bar{x}_r=0$, $r=1, 2, \ldots, T$. Then $\sigma_r^2=\frac{1}{n}\sum_{lr}^{R_r}\sum_{m_r}\sum_{u_r}^{R_r}\sum_{u_r}f_{u_rv_r}(l_rm_r)x_{u_rv_r}x_{l_rm_r}$. Let $\rho_{rs}\sigma_r\sigma_s$ be the total covariance between the r-th dimension and the s-th dimension, that is
$$\rho_{rs}\sigma_r\sigma_s=\frac{1}{n}\sum_{i}^{n}a_i(r)a_i(s)$$
$$=\frac{1}{n}\sum_{l_s}^{R_s}\sum_{m_s}\sum_{u_r}^{R_r}\sum_{v_r}f_{u_rv_r}(l_sm_s)\,x_{u_rv_r}x_{l_sm_s},$$
Let

$\rho_{rs}(t)\sigma_r(t)\sigma_s(t)$ be the covariance in the t-th stratum,
$$\rho_{rs}(t)\sigma_r(t)\sigma_s(t)=\frac{1}{n_t}\sum_{l_s}^{R_s}\sum_{m_s}\sum_{u_r}^{R_r}\sum_{v_r}f^t_{u_rv_r}(l_sm_s)\,x_{u_rv_r}x_{l_sm_s}-\bar{x}_r(t)\bar{x}_s(t).$$

The behaviours of these numerical values have the functional and operational meanings in intercorrelated patterns and will be interpreted in the valid sense from the point of view of social psychology and statistical mathematics. Thus

$$\sigma^2 = \left| \frac{1}{n} \sum\sum\sum\sum f_{u_r v_r}(l_s m_s)\, x_{u_r v_r} x_{l_s m_s} \right|$$

$$\sigma_t^2 = \left| \frac{1}{n_t} \sum\sum\sum\sum f^t_{u_r v_r}(l_s m_s)\, x_{u_r v_r} x_{l_s m_s} - \bar{x}_r(t)\bar{x}_s(t) \right|$$

From these, we can obtain the numerical values $x_{j_r k_r}$ $r=1, 2, \ldots, T$, $j_r=1, 2, \ldots, R_r$, $k_r=1, 2, \ldots, K_{j_r}$ which maximize μ or λ, these have the same meaning and content as in cases 1 and 2. In special cases, we can explicitly describe the equations which have the same forms as the combination of those of cases 1 and 2. In this case, we can obtain x_{lm} by the same operations using successive approximation "step by step in turn" method which is similar to those of cases 1 and 2. This is possible by using the methods in cases 1 and 2 alternatively. It will be possible to treat very complex phenomena by this method for the purpose of predicting events in the valid sense. For example, suppose that number of items is 9, each of which has the same three sub-categories. So we need to solve the equation containing $9\times 3=27$ unknowns.

In case 1, we must solve the simultaneous linear equations containing 27 unknowns and 1 unknown constant. This is impossible by our calculating machine and may be difficult even by the most improved electronic computer. If we can find that 9 items are divided into 3 sub-groups in which items are on unidimensional scale and which consist of 3 items, respectively, we adopt the method of case 3. In this case we have only to solve the 3 simultaneous linear equations containing 9 unknowns (reduced to 7 generally) and 1 unknown constant step by step on the idea of successive approximation as can be seen in equations (*) $\binom{*}{*}$ $\left(\begin{smallmatrix}*\\*\\*\end{smallmatrix}\right)$ (see case 2) and case 1.

This is easy to solve by our calculating apparatus (simultaneous equation solver) which has been constructed by Dr. T. Sasaki, Director of the Institute, to solve simultaneous linear equations containing 9 unknown variables. Generally by this apparatus and this method, we can quantify the 10×11 unknowns i.e., total sub-categories in items (dimension $10\times$ numbers of total sub-categories in items 11 (reduced to $11-2=9$ generally).

I express my gratitude to Professor Z. Suetuna and my colleague Mr. H. Akaike.

THE INSTITUTE OF STATISTICAL MATHEMATICS

REFERENCES

[1] C. Hayashi: On the Quantification of Qualitative Data from the Mathematico-Statistical Point of View, *Annals of the Institute of Statistical Mathematics*, Vol. II, No. 1, 1950.

[2] C. Hayashi: On the Prediction of Phenomena from Qualitative Data and the Quantification of Qualitative Data from the Mathematico-Statistical Point of View, *Annals of the Institute of Statistical Mathematics*, Vol. III, No. 2, 1952.

[3] C. Hayashi: Measurement and Quantification of Social Attitude (in Japanese), *Proceedings of the Institute of Statistical Mathematics*, Vol. 1, No. 2, 1954.

[4] C. Hayashi: Multidimensional Quantification I, II. *Proceedings of the Japan Academy*, No. 2, No. 3, 1954.

[5] Hiroshi Midzuno, Chikio Hayashi and Hirojiro Aoyama: *Quantification and Prediction — with the prediction of election poll —* (in Japanese), Maruzen Co. Ltd. Tokyo. 1953.

[6] C. Hayashi: Examples of the theory of quantification, *Proceedings of the Institute of Statistical Mathematics*, Vol 2, No. 1, 1954.