On the Prediction of Phenomena from Qualitative Data and the Quantification of Qualitative Data from the Mathematico-Statistical Point of View

By Chikio HAYASHI

(Received March 5, 1952

§ 1. Introduction

The quantification of qualitative data obtained by measurements and observations is an important method in both social and natural science research [3]. However, that is of course doubtful from the point of view of methodology, if the method of quantification is not reasonable and optional.

If the method is resonable from the mathematico-statistical point of view, quantification of qualitative data is very useful for our purpose, for example, not only in analysing and predicting of phenomena but also in programming or designing of surveys. This is a different branch of analysing methods (in wider sense) of phenomena from the theory of so-called statistical estimation or testing hypothesis or statistical inference which is also a branch of them, though we sometimes use the conception of theory of probability and statistics. Especially this gives how to solve the problems in question. It gives the method of formulating our chaotic universe (this word is used in the sense of sampling theory) into a clear form from the mathematico-statistical point of view and the method of surveying, analysing and predicting, which is theoretically reasonable and effective for our purpose, by introducing metrics into our universe.

Quantification should be done so that we may be able to know reasonably just what we want to know. So we must devise the methods of quantification to fulfil the following properties; validity, reliability, objectivity, reproducibility, consistency, and adequacy, that is to say, effectiveness in a nutshell, which are often discussed in psycometric theory. General processes, in which we give these properties to the methods of quantification, are mathematico-statistical, but are not discussed here.

The essential point of quantification method from the mathematico-statistical point of consists in view quantifying, for our concrete and clear purpose, that is, to obtain a scientifically useful guide for our action, the qualitative statistical data on the basis of our measurements and observations without unreasonable and optional hypotheses.

In this sense, quantification has not absolute meaning but relative, functional, and operational meaning to our purpose. The numerical values given to qualitative statistical data by mathematico-statistical quantification methods vary with our purpose. So the numerical values given to the same qualitative data are also naturally variable with our purpose. In quantifying qualitative data, the leading idea of practical operation of quantification is to pursue "effectiveness" from the mathematico-statistical point of view.

To achieve this purpose, in the course of our quantification, we should always be aware of "rational behaviour" that seeks to obtaine the so-called "optimum" in which we can utilize to the maximum extent, what already is and therefore is available, while, endeavoring to rationally deal with and control what is unknown in the process. In some cases this results in the maximization of correlation ratio and in the maximization of success rate of prediction. And in other cases, the idea of the so-called "minimum-maximum" (maximum-minimum) with the aim of securing the most rational safety as the criteria of voluntary actions, may be used [2, 10].

This paper is a continued report from the previous paper [2], that discusses on the problem of classification by quantification method from the point of maximizing the success rate of prediction of social phenomena in the sense of theory of probability.

In this paper the following methods of quantification will be discussed below; (1) the methods of quantification of qualitative statistical data obtained by our measurements and observations when an outside criterion is given and so the property of validity is necessary in it; (2) when the property of reproducibility is necessary in it, that is to say, the patterns of behaviour must be represented by some numerical values; (3) when the effective grouping is required.

The general methods of quantification are, roughly speaking, classified into the followings,

(1) Where the hypothesis of latent structure which is used as the pillar of theory, is set, in reality, for example, the theory of so-called paired comparison.

(2) Where the latent structure is not considered in reality.

(a) Where an outside criterion is given: the quantification of behaviour patternings (represented by qualitative data or rational behaviour) to satisfy the outside criterion: in this case the property of validity is indispensable: in the problem of prediction of phenomena, this method is often used.

(b) Where an outside criterion is not given: in this case it is essential to make an index (numerical value) representing, in some sense, behaviour patternings, so the property of reproducibility is indispensable: for example, the method of sca'e analysis and quantification of attitude, etc. belong to this category.

(1) is often treated by L. L. Thurstone. The method of the previous paper belongs to (2) (a), and that of this paper belongs to (2) (a), (b).

§ 2. Quantification Method of Qualitative Data when an Outside Criterion Represented by a Numerical Value is Given

We draw random sample of size n from a population. We use the data of these n persons.

A numerical value is given to each person as an outside criterion, that is obtained by another survey system. Now suppose that the questionnaires consisting of R items, each of which has several sub-categories in it respectively, are given to sample, under the instruction of checking in only one sub-category in each item which he thinks to be so in it.

Then each person has a numerical value as an outside criterion, which we call an outside variable or external variable, and a behaviour pattern represented by his responses in the form of item-category reaction. Then the problem arises that the relation between the outside variable and behaviour patterns is to be quantified, that is to say, the unknown outside variable is to be estimated from a known behaviour pattern of a person, using a quantitative formula obtained from the analysis of the sample data we had previously. An approaches which aims at estimating the outside variable from a behaviour pattern that is quantified by so-called scale analysis method [11], has been seen but it is meaningless [4]. The essential point of scale analysis aiming at reproducibility is different [4]. When an outside variable is given, the property of validity is indispensable in quantification. This problem has quite different features from scale analysis. Now, how to quantify behaviour patterns, that is to say, to give a numerical value to each sub-category of each item and synthetize responses in the items so as to be able to estimate, with high confidence level, the outside variable from the data concerning with behaviour patterns, must be considered. To estimate the outside variable from behaviour patterns with high confidence level, they must be quantified and synthetized. In this case, all items need not be scalable. Now let A be an outside variable and I_s be the s-th item, $s=1,2,\cdots\cdots,R$ (R is the number of items).

Let A_i be the numerical value which i-person has as the outside criterion. Each item has respectively the following sub-categories; k_j being the number sub-categories in the j-th item $(j=1, 2, \cdots\cdots, R)$

$$I_1,\ \{c_{11}, c_{12}, \cdots\cdots, c_{1k_1}\}$$
$$I_2,\ \{c_{21}, c_{22}, \cdots\cdots, c_{1k_2}\}$$
$$\vdots$$
$$I_R,\ \{c_{R1}, c_{R2}, \cdots\cdots, c_{Rk_R}\}$$

Let $X_{s(i)}$ be one of the sub-categories, $c_{s1}, c_{s2}, \cdots\cdots, c_{sk_s}$, in the s-th item, which i-person checks in the s-th item. Let i-person's response pattern (behaviour pattern) be $X_{1(i)}$ in I_1, $X_{2(i)}$ in I_2, $\cdots\cdots$, $X_{s(i)}$ in I_s, $\cdots\cdots$, $X_{R(i)}$ in I_R.

The response patterns (behaviour patterns) of persons of size n are shown below, for example.

item	I_1				I_2				....	I_R		
sub-categories outside variable of sample	c_{11}	c_{12}	......	c_{1k_1}	c_{21}	......	$c_{2(k_2-1)}$	c_{2k_2}	....	c_{R1}	......	c_{Rk_R}
A_1		V			V				....	V		
A_2	V							V	....	V		
⋮	⋮	⋮	⋮	⋮	⋮	⋮			⋮	⋮	⋮	⋮
A_i				V			V		....			V
⋮	⋮	⋮	⋮	⋮	⋮	⋮			⋮	⋮	⋮	⋮
A_n	V							V	....			V

V is the sign of response

For example i-person has A_i as an outside variable and checks in the sub-category c_{1k_1} in I_1, $c_{2(k_2-1)}$ in I_2, $\cdots\cdots$ and c_{Rk_R} in I_R, his response pattern being shown to be V-pattern in the above table.

In this case where we estimate the outside variable from the response patterns, we wish to quantify the response patterns so as to maximize the correlation coefficient ρ between A and $X_1+X_2+\cdots\cdots+X_R$, symbolically writing. This is the first approximation of estimation, because we need not necessarily use the correlations coefficient between A and $X_1+X_2+\cdots\cdots+X_R$ and we may well use the coefficient between A and $(X_1+X_2+\cdots\cdots+X_R)^{\alpha}$ or etc. But it is meaningful from the mathematico-statistical point of view to use the linear form

$X_1+X_2+\cdots\cdots+X_R$, as the first approximation. It leads to the notion of multiple correlation between variable A and responses in items. This linear form is not so restricted, considering from the point of view of making the patterns linear by quantifying qualitative data represented in the form of response patterns. Maximizing the correlation coefficient has the meaning to maximize the confidence of estimation, that is, minimize the variance of the estimation from the theory of regression estimate in sampling theory. Now suppose that numerical value x_{lm} is given to the m-th sub-category of the l-th item, c_{lm}, $(m=1, 2, \cdots\cdots, k_l;\ l=1, 2, \cdots\cdots, R)$. The total number of x_{lm} is equal to the sum of number of all sub-categories, $k_1+k_2+\cdots\cdots+k_R$.

Let n_{lm} be frequency of responses in the m-th sub-category in the l-th item, that is the frequency of persons who checked in the category c_{lm}. So

$$n=\sum_{k=1}^{k_j} n_{jk} \quad (j=1, 2, \cdots\cdots, R),$$

since each person certainly checks only in one sub-category in each item. Now let be $n_{jk}\geqq 2$. The response pattern of i-person, $(X_{1(i)}, X_{2(i)}, \cdots\cdots, X_{R(i)})$, is to be synthesized in the form of

$$\alpha_i=X_{1(i)}+X_{2(i)}+\cdots\cdots+X_{R(i)},$$

α_i having a numerical value, because of being given a numerical value $x_{s(i)}$ to $X_{s(i)}$, which we call a numerical score of i-person.

The sample correlation coefficient ρ between A and $X_1+\cdots\cdots+X_R$ is written as followings,

$$\rho(A, X_1+\cdots\cdots+X_R)=\frac{\frac{1}{n}\sum_{i=1}^{n}(A_i-\bar{A})(\alpha_i-\bar{\alpha})}{\sigma_A\sigma_\alpha}$$

where
$$\bar{A}=\frac{1}{n}\sum_{i=1}^{n}A_i, \quad \sigma_A{}^2=\frac{1}{n}\sum_{i=1}^{n}(A_i-\bar{A})^2,$$
$$\bar{\alpha}=\frac{1}{n}\sum_{i=1}^{n}\alpha_i, \quad \sigma_\alpha{}^2=\frac{1}{n}\sum_{i=1}^{n}(\alpha_i-\bar{\alpha})^2.$$

It is our purpose to determine x_{lm} $(m=1, 2, \cdots\cdots, k_l,\ l=1, 2, \cdots\cdots, R)$ so as to maximize the correlation coefficient ρ. ρ is invariant under a shift of origin. Hence the origin is arbitrary, and we may select it a priori so that $\bar{A}=0$, $\bar{\alpha}=0$, and the mean value in each item are zero. Then we can put ρ

$$\rho=\frac{\frac{1}{n}\sum_{i=1}^{n}A_i\alpha_i}{\sigma_A\sigma_\alpha}$$

where
$$\sigma_A{}^2=\frac{1}{n}\sum_{i=1}^{n}A_i{}^2, \quad \sigma_\alpha{}^2=\frac{1}{n}\sum_{i=1}^{n}\alpha^2{}_i.$$

To maximize ρ with respect to x_{lm}, the conditions

$$\frac{\partial \rho}{\partial x_{lw}}=0, \quad (m=1,2,\cdots\cdots,k_l,\ l=1,2,\cdots\cdots,R)$$

are necessary.

Before we calculate this, some notations will be introduced. Let

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

So the following relations hold,

$$\sum_{k=1}^{k_j}\delta_i(jk)=1$$

$$\begin{cases}\delta_i(jk)\,\delta_i(jk')=0 & (k\neq k')\\ \qquad '' \qquad\quad =1 & (k=k')\end{cases}$$

$$\sum_{i=1}^{n}\delta_i(jk)=n_{jk} \qquad \text{(for all } j,k\text{)},$$

$$\sum_{k=1}^{k_j}\sum_{i=1}^{n}\delta_i(jk)=n \qquad \text{(for all } j\text{)}.$$

Let $f_{lm}(j,k)$ be $\qquad \sum_{i=1}^{n}\delta_i(lm)\,\delta_i(jk)=f_{lm}(jk)$.

This is equal to the number of those who check in the m-th sub-category in the l-th item and moreover k-th sub-category in the j-th item, and represents, in a sense, correlated relations between the l-th item and the j-th item, and means the correlation patterns between items.

So the followings hold,

$$f_{lm}(jk)=f_{jk}(lm)$$

$$\sum_{m=1}^{k_l}f_{lm}(jk)=n_{jk} \qquad \text{(for all } l,j\text{)}$$

$$\sum_{k=1}^{k_j}f_{lm}(jk)=n_{lm} \qquad \text{(for all } l,m\text{)}$$

$$\sum_{m=1}^{k_l}\sum_{k=1}^{k_j}f_{lm}(jk)=n \qquad \text{(for all } l,j\text{)}$$

$$f_{lm}(jk)=0 \qquad j=l,\ k\neq m.$$

So $\frac{\partial \rho}{\partial x_{lm}}=0$ is written as below.

$$\frac{1}{n}\frac{\partial}{\partial x_{lm}}\sum_{i=1}^{n}A_i\alpha_i-\sigma_A\rho\frac{\partial \sigma_\alpha}{\partial x_{lm}}=0 \quad (*)$$

$$\frac{1}{n}\frac{\partial}{\partial x_{lm}}\sum_{i=1}^{n}A_i\alpha_i=\frac{1}{n}\sum_{i=1}^{n}A_i\delta_i(lm)$$

$$\frac{\partial \sigma_\alpha}{\partial x_{lm}}=\frac{1}{n\sigma_\alpha}\left(x_{lm}n_{lm}+\sum_{j=1}^{R}{}'\sum_{k=1}^{k_j}{}'\,x_{jk}f_{lm}(jk)\right)$$

where $\sum'\sum'$ covers all range of j and k except the case where $j=l$, $k=m$ holds simultaneously.

So (*) is

$$\sum_{i=1}^{n}A_i\delta_i(lm)=\frac{\sigma_A\rho}{\sigma_\alpha}x_{lm}n_{lm}+\frac{\sigma_A\rho}{\sigma_\alpha}\sum_j{}'\sum_k{}'\,x_{jk}f_{lm}(jk) \quad (\overset{*}{*})$$

$$(m=1,2,\cdots\cdots,k_l,\ l=1,2,\cdots\cdots,R)$$

The conditions mentioned above, are the followings

$$\left.\begin{array}{l} \overline{A}=\dfrac{1}{n}\sum\limits_{i=1}^{n} A_i=0 \\ \overline{\alpha}=\dfrac{1}{n}\sum\limits_{m}^{k_l}\sum\limits_{j}^{R}\sum\limits_{k}^{k_j} x_{jk}f_{lm}(jk)=\dfrac{1}{n}\sum\limits_{j}\sum\limits_{k} x_{jk}n_{jk}=0 \\ \sum\limits_{k=1}^{k_j} x_{jk}n_{jk}=0 \qquad \text{(for all } j) \end{array}\right\} \begin{pmatrix}*\\ *\\ *\end{pmatrix}$$

To require all x_{lm}, it is necessary to solve $\binom{*}{*}$ under the condition $\begin{pmatrix}*\\ *\\ *\end{pmatrix}$.

We may put $\dfrac{\sigma_A\rho}{\sigma_\alpha}$ a constant, for example, equal to 1, because the left hand of $\binom{*}{*}$ is independent of x_{lm} and $\dfrac{\sigma_A\rho}{\sigma_\alpha}$ is common to the all. It is easy to solve $\binom{*}{*}$ if $k_1+k_2+\cdots\cdots+k_k$ is relatively small. The solutions is that we require. After it, the estimation of outside variable from response patterns is done by the ordinary method of regression plane.

Now let $\dfrac{\sigma_A\rho}{\sigma_\alpha}\dfrac{1}{n_{lm}}\sum\limits_{j}'\sum\limits_{k}' x_{jk}f_{lm}(jk)$ be constant independently of m under the condition of l being definite. Then it is shown that the relation

$$\sum_{j}'\sum_{k}' x_{jk}f_{lm}(jk)=0$$

holds under the condition $\begin{pmatrix}*\\ *\\ *\end{pmatrix}$.

Then we obtain $\qquad x_{lm}=\dfrac{1}{n_{lm}}\sum\limits_{i=1}^{n} A_i\delta_i(lm).$

In this case, it is very simple. If the l-th item and the other items are statistically independent, the relation $\dfrac{1}{n_{lm}}\dfrac{\sigma_A\rho}{\sigma_\alpha}\sum\limits_{j}'\sum\limits_{k}' f_{lm}(jk)=\text{const.}$ holds, of course, in population. Then it is shown that the above relation holds if ρ, σ_A, σ_B are population correlation coefficient and variances. We may think that it is the first approximation that to maximize the correlation coefficient between a quantitative variate and items (the sub-categories in items) are given to the sub-category numerical values which are proportional to the mean value of the numerical values (outside variable) of the persons who checked in the sub-category in the item [6].

Here we considered quantification using the data of random sample. The confidence interval of the estimation of population correlation coefficient will be calculated, complicated it may be, from the ρ we obtained, using by the theory of sampling.

It tells me that the variance or mean square error, $\sigma_a{}^2$, of $a=f(b)$ is approximately

$$\sigma_a{}^2 \fallingdotseq f'(\bar{b})\,\sigma_b{}^2$$

where a and b are random variables,

f is a differentiable function,

$\sigma_b{}^2$ is variance of b,

$\bar{b}$ is mean of b.

The method will be considered that maximizes the correlation coefficient in populate in some sense and narrows the confidence interval of the estimate under a definite confidence level, but it will be different from that mentioned above which is a first step towards it.

This method will be often used in analysing social phenomena. In some cases this is used in estimating the effect of stratification in sampling surveys. Strata are often constructed by using the qualitative data which are considered to be highly correlated with an attribute (represented by real number) about which we wish to require some informations in population.

In two-stage sampling where only one primary sampling unit in a stratum is at random sampled, the variance between primary sampling units in a stratum can not be estimated reasonably. If n primary sampling units in the whole are surveyed, we obtain their information about the attribute and qualitative data used for stratification. Then we can quantify the qualitative data by the method mentioned so as to maximize the correlation coefficient between the attribute and qualitative data, so we can estimate the attributes of all primary sampling units and the between variance by evaluating the value of the obtained correlation coefficient.

[Example]

In Literacy survey [5], the reading and writing ability of a person was quantified by literacy test items. This ability is represented by a score. We regard this as an outside variable. As an example, let us consider that we estimate the ability score (outside variable) of a person from his schooling and age (response pattern). As the result of survey, each person has a score, and the degree of schooling and age, which are obtained in the form of sub-category.

Let the first item be schooling, which has six sub-categories,

c_{11}; not entering a school

c_{12}; leaving elementary school half-way

c_{13}; finishing elementary school

c_{14}; finishing higher elementary school

c_{15}; finishing middle school

c_{16}; over higher school

let the second item by age which has five sub-categories

c_{21}; 15～19

c_{22}; 20～39

c_{23}; 40～49

c_{24}; 50~59

c_{25}; 60~over

Random sample of size 1000 in all over the country was used in this case.

Response patterns of size 1000 and scores (outside variable) are given. These relations are tabulated as below.

Frequency distribution between the three items

Score Schooling	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	Total
c_{11}	15	3	—	1	1	1	—	—	—	—	1	—	—	—	—	1	—	—	23
c_{12}	6	—	3	5	2	1	4	1	1	1	1	2	—	1	2	2	1	—	33
c_{13}	3	2	6	7	10	13	5	13	8	14	12	10	10	12	24	25	28	19	221
c_{14}	—	1	—	2	—	—	1	2	6	6	10	17	19	28	41	80	116	140	469
c_{15}	—	—	—	—	—	—	—	—	1	1	—	—	1	3	9	12	41	112	180
c_{16}	—	—	—	—	—	—	—	—	—	—	—	—	—	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

	Score Age	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	Total
c_{21}	10	—	—	1	2	—	—	—	3	4	7	8	12	6	13	21	27	43	32	179
c_{22}	20~30	3	1	2	4	3	5	3	5	6	7	7	5	16	14	38	62	99	218	498
c_{23}	40	2	2	1	6	3	5	6	3	4	5	4	7	5	9	12	25	24	55	178
c_{24}	50	13	—	3	1	6	5	—	4	1	3	4	3	3	5	5	11	22	21	110
c_{25}	60	6	3	2	2	1	—	1	1	1	—	1	2	—	4	2	1	5	5	35
	Total	24	6	9	15	13	15	10	16	16	22	24	29	30	45	78	126	193	329	1000

	c_{11}	c_{12}	c_{13}	c_{14}	c_{15}	c_{16}	Total
c_{21}	—	—	9	122	24	24	179
c_{22}	—	7	84	247	124	36	498
c_{23}	5	13	70	60	22	8	178
c_{24}	10	10	44	34	9	3	110
c_{25}	8	3	14	6	1	3	35
Total	23	33	221	469	180	74	1000

From these, x_{lm} are obtained. (This complicated calculation was done by my colleague, Mr. Ishida D. Masatsugu in the Institute of Statistical Mathematics)

As the results, we obtain $\rho=0.73$.

$$\begin{array}{ll}
x_{11}=-11.640 & x_{21}=-1.207 \\
x_{12}=-7.873 & x_{22}=+0.382 \\
x_{13}=-2.959 & x_{23}=+0.591 \\
x_{14}=+1.175 & x_{24}=-0.276 \\
x_{15}=+2.287 & x_{25}=-1.404 \\
x_{16}=+2.960 &
\end{array}$$

§ 3. Some Quantification Methods of Attitude; especially the case where content universe is not scalable which consists of items to measure attitude towards something

Suppose that attitude is measured by the questionnaire items from content universe, which are assumed to fulfil the property of validity in content, and the items have several sub-categories respectively as those of scale analysis type.

For example, we have R dichotomous items (questionnaires), each of which has only two sub-categories, yes or no. If R items are scalable as the whole, quantification of items (sub-categories) is performed by the idea of L. Guttman [11]. If items are scalable, that is, each response pattern has a definite rank order in the whole, each item (sub-categories) and each response pattern are reasonably quantified and moreover intensity of response pattern (attitude) is also quantified and so on. The fundamental idea of his theory is that content, intensity and etc. (several phases of attitude) are to be contained in response patterns and so the quantification method can reveal these features, if items are scalable. If items are not scalable, sub-groups of items which are scalable, are constructed from the whole. Thus we have scalable sub-groups of items. But we can not make an attitude score toward something by synthetizing these sub-groups as they are, though this is our purpose.

Then we must adopt the different stand point. We shall give a method of quantifying non-scalable or quasi-scalable items below.

If items are not scalable, it is desirable to consider at least three phases, "content," "weight" and "intensity" though they highly correlate with each other. "Weight" means the degree of importance of each item in determining the attitude towards something as a whole. This is the conception not to be considered in the theory of scale analysis. "Weight" will be obtained by using the fact that dichotomous items are not scalable in our case. The statistico-mathematical method will be discussed, to quantify dichotomous items on these three phases by using the response patterns in the R items. Before they are generally discussed, we given some illustrations. As an example, the survey* that we have performed will be discribed as below. This is an attitude survey towards French or American culture, to see whether a person is favourable to French culture or American culture. The questionnaires consists of R dichotomous items, (in our case R being six). As an example, an item will be shown.

> " There is a proposition as the following. If you agree with it, ckeck "yes", if you do not agree, check "no".
>
> Culture must have creative power. But the more important factor to determine the value of culture is considered to be that it has depth of tradition. From this point of view, we think that French culture has higher value than American culture.
>
> (1) yes (2) no "

Such R biased, in some sense, questionnaires items as this, are set, which have the same form and are different in dimensions. It is considered that they expect favourable responses to French culture. The reason why these biased questionnaires should be, used will be shown later on. From the result of pretest, we knew that these R items were not scalable in the rigorous sense, but were quasi-scalable (the degree of reproducibility is about 70%) and they had $P_i(i=1, 2, \cdots\cdots, R)$, where P_i was the ratio of the number of positive (yes) responses in the i-th item to the total, and $\max(P) \doteqdot 0.8$, $\min(P) \doteqdot 0.3$.

So the whole survey was performed following the order mentioned below so as to know three phases i.e. content, weight, and intensity. It has three types of test.

(i) R dichotomous items (questionnaires), having the form mentioned above, are set to each of sample of size n (n-persons).

Each person must check only one sub-category, "yes" or "no", in each item.

(ii) After the above test, the sheets filled with responses were presented by the sample.

Then the test of the next form is set. Giving the following types of response patterns previously determined, let each person choice and check in only one type to which he thinks that his proper response pattern-attitude-(represented by the responses in the test (i)) is the nearest.

[*] This survey has been planned and performed by the section of social psychology in the committee of Mathematical-behavioristics. The members of this section are C. Hayashi, H. Ikeuchi, Y. Mizuhara, K. Sano, S. Ōshio.

Type of response pattern

Items Type of response patterns	I_1	I_2	I_3	....	I_{R-1}	I_R
1	+	+	+	....	+	+
2	+	+	+	....	+	−
3	+	+	+	....	−	−
⋮	⋮	⋮	⋮	⋮	⋮	⋮
$R+1$	−	−	−	....	−	−

+ sign shows positive (yes) response, and
− sign shows negative (no) response

Let s_j be the ratios of the number of positive responses in the j-th item to the total, which has been obtained in pretest. The order of items has been changed with the order of magnitude of s_j. Let $s_1>s_2>\cdots\cdots>s_R$. In this table, type 1 of response pattern means the type of persons who check in positive sub-categories in all items and type 2 of response pattern means the type of persons who check in positive sub-categories in the first $(R-1)$ items and in negative sub-category in the last item, and so on.

But in the test of (i), the order of items is randomized. Thus we compare the type of response pattern which each person choiced in (ii), with the response pattern he checked freely in (i).

If items are perfectly scalable, it is naturally considered that the choiced type coincides with the pattern he checked freely, so far as the items fulfil the property of reliability. If not scalable, the former does not coincide with the latter. So, by comparing the former with the latter with respect to every person, we consider to quantify the dichotomous items from the mathematico-statistical point of view.

(iii) After few days, biased questionnaire items to the inverse direction are set. It is considered that the items expect favourable responses to American culture. As an example, the item corresponding to the item mentioned above, will be shown.

> "There is a proportion as the following. If you agree with it, check "yes" if you do not agree, check "no".
>
> Culture must have depth of tradition. But the more important factor to determin the value of culture is considered to be that it has creative

power. From this point of view, we think that American culture has higher value than French culture.

(1) yes (2) no ”

These R items are given to the same sample. And they are required to check in only one sub-category in each item. This result is also compared with the results previously obtained and used to measure intensity in attitude. The concrete method will be described later.

Thus each person has been subjected to three kinds of test:

(i) biased questionnaire items which expect favourable responses to something

(ii) choice of the given response patterns

(iii) biased questionnaire items to the inverse direction which expect unfavourable response to the same thing.

By comparing these responses, we quantify three phases, i.e. content, weight, and intensity. The essential point of our idea is that we make use of the patterns of consistency and inconsistency of responses in (i), (ii), (iii), and quantify the three phases.

We shall general y discuss these problems as below.

(1) Quantification of content, mainly.

As items are not scalable, it is meaningless to require content score purely. Content score with weight must be obtainted. The following method will be appropriate for our purpose from this point of view. In this case we use the test (i) and (ii). If we use the responses in (ii) as the criterion of stratification of n persons, the responses in (i) are considered to be attributes of the elements (persons) belonging to the stratum. Then we have $R+1$ strata, the characteristics of which are represented by response in (ii) (see (ii)).

The characteristics of strata being represented by responses in (ii) Number of strata	Item	Item	....	Item	Item
	1	2	....	$R-1$	R
1	+	+	....	+	+
2	+	+	....	+	−
⋮	⋮	⋮	⋮	⋮	⋮
$R+1$	−	−	....	−	−
+, − signs the same meaning as that previously mentioned at (ii) of this paragraph.					

To which stratum each person belongs is determined by his response in test (ii). For example, he belongs to the 2nd stratum, if his response pattern in (ii) is $(++\cdots\cdots+-)$.

(a) R items have dichotomous sub-categories respectively. Let $\{c_{11}, c_{12}\}$, $\{c_{21}, c_{22}\}, \cdots\cdots, \{c_{R1}, c_{R}\}$ be sub-categories in items. Let us consider that we give a numerical value (content score) x_{lm} to the m-th sub-category in the l-th itém, c_{lm}, from the mathematico-statistical point of view. Response patterns of n persons are, for example, as below.

Number of strata	Item	1		2		………………	R	
	Sub-category / Person	c_{11}	c_{12}	c_{21}	c_{22}	………………	c_{R1}	c_{R2}
1	1	V		V		………………	V	
	2	V		V		………………		V
	⋮					⋮		
	n_1	V			V	………………	V	
2	1	V		V		………………		V
	2	V		V		………………	V	
	⋮					⋮		
	n_2	V			V	………………		V
⋮	⋮					⋮		
$R+1$	1		V		V	………………		V
	2		V		V	………………		V
	⋮					⋮		
	n_{R+1}		V	V		………………		V

V sign means the check in response of i-person.

n_t is the number of persons belonging to the t-th stratum, where $n=\sum_{t=1}^{R+1} n_t$. Let $\{X_{1(i)}, X_{2(i)}, \cdots\cdots, X_{R(i)}\}$ be the response pattern of i-person, where $X_{j(i)}$ means the sub-category of the j-th item that i-person checks in. Now we use the score $\alpha_i = x_{1(i)} + x_{2(i)} + \cdots\cdots x_{R(i)}$ as the score of i-person, where $x_{j(i)}$ is the numerical value given to the sub-category in the j-th item he checks in. The linear form is considered to be appropriate, as we discussed in § 2, according to the idea of the first approximation.

So we have $\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(\alpha_i - \bar{\alpha})$ as the total variance with respect to perpsons, where $\bar{\alpha} = \frac{1}{n}\sum_{i=1}^{n}\alpha_i$. Now we consider that we quantify the sub-categories (items) so as to maximize the effect of stratification, that is, so as to maximize the correlation ration $\eta^2 = \frac{\sigma_b{}^2}{\sigma^2}$, where $\sigma_b{}^2$ is the variance between strata. This is reasonable method of quantification, because η^2 is a measure of discrimative power of items from the fact that η^2 is the larger, the more homogeneous the patterns of perrons winthin each stratum are, and η^2 is the smaller, the more heterogeneous the patterns of persons within each stratum, and because it makes use of our data most effectively to maximize the η^2.

In some sense, we can not expect more information. If the response patterns within strata are heterogeneous, that is, homogeneous between strate, the tendency will be seen that the numerical value x_{l1} given to the first sub-category in the l-th item and x_{l2} $(l=1, 2, \cdots\cdots, R)$ become equal and items have no discriminative power, and so the purpose of attitude survey is lost, and α_i $(i=1, 2, \cdots\cdots, n)$ has no meaning.

If η^2 is large in the result of quantification, we can treat quantitatively the attitude by using x_{lm} (or α). By the score α, the attitude of the person having α, is interpreted to be more favourable (to French culture in our case) or more unfavourable with high confidence level.

It is considered to be valid, from the construction of the tests (i), (ii) and from the results of pretest, to use the response pattern of (ii) as the criterion of stratification. So as to require x_{lm} to maximize η^2, let us introduce the following definition. Let

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

Then $$\sum_{k=1}^{2}\delta_i(jk)=1, \qquad \sum_{j=1}^{R}\sum_{k=1}^{2}\delta_i(jk)=R, \qquad \alpha_i=\sum_{j=1}^{R}\sum_{k=1}^{2}\delta_i(jk)\,x_{jk}$$

So $$\bar{\alpha}=\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{R}\sum_{k=1}^{2}\delta_i(jk)\,x_{jk}, \qquad \sigma^2=\frac{1}{n}\sum_{i=1}^{n}\left(\sum_{=1}^{R}\sum_{k=}\delta_i(jk\ x_{jk})\right)^2-\bar{\alpha}^2$$

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

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

which represents correlation pattern between responses in items of each person, $\sum_l{}' \sum_m{}' \sum_{j=1}^{R} \sum_{k=1}$ covers all range of $l, m, j, k,$ except $l=j$, $m=k$ holds simultaneously.

$$\sigma_b{}^2=\sum_{t=1}^{R+1}(\bar{\alpha}_t-\bar{\alpha})^2\frac{n_t}{n}$$

where $$\bar{\alpha}_t=\frac{1}{n_t}\sum_{i[t]=1}^{n_t}\alpha_{i[t]}=\frac{1}{n_t}\sum_{j=1}^{R}\sum_{k=1}^{2}x_{jk}g_t(jk),$$

$$g_t(jk)=\sum_{i[t]=1}^{n_t}\delta_{i[t]}(jk), \qquad n_{jk}=\sum_{t=1}^{R+1}g_t(jk), \qquad n_t=\frac{1}{R}\sum_{j=1}^{R}\sum_{k=1}^{2}g_t(jk)$$

$\delta_{i[t]}(jk)$ means $\delta_i(jk)$ which the i-person 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 $x_{uv}(u=1, 2, \cdots\cdots, R,\ v=1, 2)$, $\frac{\partial\eta^2}{\partial x_{uv}}=0$, that is, $\frac{\partial\sigma_b{}^2}{\partial x_{uv}}=\eta^2\frac{\partial\sigma^2}{\partial x_{uv}}$. Calculating this, using $\bar{\alpha}=0$ without loss of generally as easily shown,

$$\frac{\partial\sigma^2}{\partial x_{uv}}=\frac{2}{n}\sum_{j=1}^{R}\sum_{k=1}^{2}\left(\sum_{t=1}^{R+1}\frac{g_t(jk)g_t(uv)}{n_t}\right)x_{jk}=\frac{2}{n}\sum_{j=1}^{R}\sum_{k=1}^{2}h_{uv}(jk)\,x_{jk}$$

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

where $$h_{uv}(jk)=\sum_{t=1}^{R+1}\frac{g_t(jk)g_t(uv)}{n_t}.$$

Then $$\sum_{j=1}^{R}\sum_{k=1}^{2}h_{uu}(jk)\,x_{jk}=\eta^2\sum_{l=1}^{R}\sum_{m=1}^{2}x_{lm}f_{uv}(lm) \qquad (u=1, 2, \cdots\cdots, R,\ v=1, 2)$$

Let the matrix $(h_{uv}(jk))$ be H, the matrix $(f_{uv}(lm))$ be F, which highly correlates with correlation patterns of responses of each person in the whole, vector x_{jk} be X. The above equation is written as follows.

$$HX=\eta^2FX \quad (*)$$

It is our problem to solve this under the conditions $\sum_{k=1}^{2}n_{jk}x_{jk}=0$, $(j=1, 2, \cdots\cdots, R)$. And to require the maximum value of η^2 being not equal to 1 and the corresponding vector X to it.

From the different standpoint, we may be able to quantify x_{lm} so as to maximize the value $\gamma=-\frac{\sigma_w{}^2}{\sigma_T{}^2}$, where $\sigma_T{}^2$ is the total variance with respect to x_{lm}

$$\sigma_T{}^2=\frac{1}{Rn}\sum_{j=1}^{R}\sum_{k=1}^{2}n_{jk}x_{jk}{}^2-\bar{x}^2, \qquad \bar{x}=\frac{1}{Rn}\sum_{j=1}^{R}\sum_{k=1}^{2}n_{jk}x_{jk},$$

and $\sigma_w{}^2$ is the wihtin variance, $\sigma_w{}^2=\sigma^2-\sigma_b{}^2$. Then we have

$$\sum_{j=1}^{R}\sum_{k=1}^{2}\{h_{uv}(jk)-f_{uv}(jk)\}x_{jk}=\frac{\gamma}{R}n_{uv}x_{uv} \qquad (u=1,2,\cdots\cdots,R,\ v=1,2).$$

Or we may also be able to quantify x_{im} so as to the value $\mu^2=1-\frac{\sigma_b{}^2}{\sigma_T{}^2}$ as the above idea. This is the similar form to (b).

(b) Let us quantify the items from the different point of view from (a). x_{im} is tle same as (a). In this case, we give a numerical value y_t to the t-th stratum. The idea of scale analysis [11] leads to this method. The quantification of x_{im} is as followings. Let $\sigma_b{}^2$ be between variance, which is in near form to case (a), but in this case we do not think the numerical value α_i given tp i-person and the corresponding value to i-person in the linear form. The value given to i-person is determined by the value ν given to the stratum to which he belongs.

Let $\sigma_T{}^2$ be total variance, then

$$\sigma_T{}^2=\frac{1}{S}\sum_{j=1}^{R}\sum_{k=1}^{2}n_{jk}x_{jk}{}^2-\bar{x}^2, \qquad \bar{x}=\frac{1}{S}\sum_{j=1}^{R}\sum_{k=1}^{2}n_{jk}x_{jk},$$

where $$S=\sum_{j=1}^{R}\sum_{k=1}^{2}n_{jk}=Rn,$$

n_{jk} has the same meaning as case (a).

$$\sigma_b{}^2=\sum_{t=1}^{R+1}\frac{S_t}{S}\left(\sum_{k=1}^{R}\sum_{j=1}^{2}\frac{x_{jk}g_t(jk)}{S_t}\right)^2-\bar{x}^2,$$

where $S_t=Rn_t$, n_t has the same meaning as case (a).

To maximize $\eta^2=\frac{\sigma_b{}^2}{\sigma_T{}^2}$ with respect to x_{uv}

$$\frac{\partial\eta^2}{\partial x_{uv}}=0 \qquad (u=1,2,\cdots\cdots,R,\ v=1,2).$$

Thus we obtain with $\bar{x}=0$,

$$\frac{1}{R}\sum_{j=1}^{R}\sum_{k=1}^{2}h_{uv}(jk)\,x_{jk}=\eta^2 n_{uv}x_{uv}, \qquad (u=1,2,\cdots\cdots,R,\ v=1,2)$$

that is

$$HX=\eta^2 RAX \qquad (*)$$

where matrix A is diagonal

$$\begin{pmatrix} n_{11} & & 0 \\ & n_{12} \ddots & \\ 0 & & n_{RR} \end{pmatrix}$$

This is the similar form to the last proposition of (a) except a multiplier.

Next, let us consider the quantification of trata. The idea is tie same. Let σ'^2 be total variance with respect to y,

$\sigma_b'^2$ be variance between sub-categories (items) with respect to y. Then

$$\sigma'^2=\sum_{t=1}^{R+1}\frac{S_t}{S}y_t^2-\bar{y}^2=\sum_{t=1}^{R+1}\frac{n_t}{n}y_t^2-\bar{y}^2,\qquad \bar{y}=\frac{1}{S}\sum_{t=1}^{R+}S_ty_t$$

$$\sigma_b'^2=\sum_{j=1}^{R}\sum_{k=1}\bar{y}_{jk}^2\frac{n_{jk}}{Rn}-\bar{y}^2,\qquad \bar{y}_{jk}=\frac{1}{n_{jk}}\sum_{t=1}^{R+1}y_tg_t(jk),\qquad \eta'^2=\frac{\sigma_b'^2}{\sigma'^2}.$$

Let us give a numerical value y_t to the t-stratum so as to maxmize η'^2 with respect to y_t, $(t=1, 2, \cdots\cdots, R+1)$. This means that we make the discriminative power of items as strong as possible, so it is our purpose. Thus to maximize $\dot{\eta}'^2$ with respect to y_w,

$$\frac{\partial\eta'^2}{\partial y_w}=0\quad(w=1, 2, \cdots\cdots, R+1),\qquad \frac{\partial\sigma_b'^2}{\partial\sigma_w}=\eta'^2\frac{\partial\sigma'^2}{\partial y_w}$$

This is written as below, without loss of generality we taken $\bar{y}=0$.

$$\sum_{t=1}^{R+1}J_{wt}y_t=\eta'^2n_wy_w,\qquad (w=1, 2, \cdots\cdots, R+1),\qquad \left(\begin{smallmatrix}*\\ *\\ *\end{smallmatrix}\right)$$

where
$$J_{wt}=\sum_{j=1}^{R}\sum_{k=1}^{2}\frac{g_w(jk)g_t(jk)}{Rn_{jk}}.$$

If we put $z_t=\sqrt{n_t}\,y_t$, $\left(\begin{smallmatrix}*\\ *\\ *\end{smallmatrix}\right)$ is as below

$$\sum_{t=1}^{R+1}K_{wt}z_t=\eta'^2z_w,\qquad \text{where}\quad K_{wt}=\frac{1}{\sqrt{n_tn_w}}\sum_{1=1}^{R}\sum_{k=j}^{2}\frac{g_w(jk)g_t(jk)}{Rn_{jk}}$$

Matrix $K(K_{wt})$ is symetric and positive definite. The latent roots in the characteristic equation $|K-\lambda E|=0$, where E is unit matrix, are real and positive. So we can require the maximum value of η'^2 except $\eta'=1$ which is not appropriate to our purpose, in the first step of the well-known approximation method [7, 9] and also the corresponding vector our required solution. If we use the transformation $\sqrt{u_{uv}}\,x_{uv}=x_{uv}'$ in $\left(\begin{smallmatrix}*\\ *\\ *\end{smallmatrix}\right)$, too we can obtain the same relation as this case, and require the values by the approximation method. η^2 or η'^2 is a measure of effectiveness of quantifications. It is easily shown [11] that the method to maximize the correlation coefficient ρ of (x, y), where xs are numerical values given to sub-categories and ys are to strata, is equialent to the method above mentioned. The fundamental idea which leads the method of §2, §3 (1) (a), (b), is to change the frequency distributions in various classifications into numerical values on the basis of mathematico-statistical conceptions.

(c) Bisides (a) and (b), several similar methods are considered. For example, instead of strata, some response types (some combinations of responses) are able to be used. In this case, it is sufficient to repeat the same operations, putting some types instead of strata.

(d) To measure the effectiveness, in the case of (a) the following methods will be considered. Then, the idea of the previous paper is used. Suppose that x_{lm} $(l=1, 2, \cdots\cdots, R,\ m=1, 2)$ are given by the method (a) and the cores given to each person, $\alpha_i (i=1, 2, \cdots\cdots, u)$ are determined. Thus we can estimate the density function $f_t(\alpha)$ of α in the t-th stratum, in practical sense that density function is obtained approximatly though the distribution is discrete. (See Fig. 1)

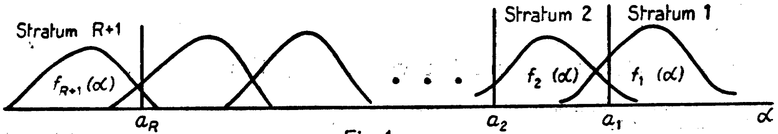

Fig. 1

Let k_t be $\frac{n_t}{n}$. Let $a_1, a_2, \cdots\cdots, a_R$ be cutting points (see [2]). It is shown in [2] that the classification using $a_1, a_2, \cdots\cdots, a_R$ has P as the success rate,

$$P=k_1 \int_{a_1}^{\infty} f_1(\alpha)\, d\alpha + k_2 \int_{a_2}^{a_1} f_2(\alpha)\, d\alpha + \cdots\cdots + k_{R+1} \int_{-\infty}^{a_R} f_{R+1}(\alpha)\, d\alpha$$

To maximize P with respect to $a_1, a_2, \cdots\cdots, a_R$,

$$\frac{\partial P}{\partial a_i}=0 \qquad (i=1, 2, \cdots\cdots, R)$$

Let the values obtained from these equation be $a_1^0, a_2^0, \cdots\cdots, a_R^0$, where

$$k_i f_i (a_i^0) = k_{.+1} f_{i+1} (a_i^0) \qquad (i=1, 2, \cdots\cdots, R)$$

hold simultaneously.

Then the maximum value of success rata is

$$P_{\max}=k_1 \int_{a_1^0}^{\infty} f_1(\alpha)\, d\alpha + k_2 \int_{a_2^0}^{a_1^0} f_2(\alpha)\, d\alpha + \cdots\cdots + k_{R+1} \int_{-\infty}^{a_R^0} f_{R+1}(\alpha)\, d\alpha.$$

This $P_{\max}$ is considered to be a measure of discriminative power of items.

(2) Especially quantification of weight.

Weight is the conception not to be considered in scale analysis, and represents the degree of importance that each item contributes in determining the attitude (behaviour) towards something as a whole. First we must be aware of that the weight of an item is not measured by itself, but in relations to others. It is difficult to quantify it purely, and the mixture of weight with content is often obtained. In some cases, where interpretation requires it, this, is, on the contrary, appropriate. Here the methods to quantify it as purely as possible and also those to quantify it with content are described below. In this case, responses (i) and (ii) are also compared.

(a) the method of using the degree of consistency in responses (i) and (ii).

By comparing the patterns of (i) with the patteıns checked in (ii), we know the number of consistent items (number of matches of signs in items).

For example, if

$$\left\{\begin{array}{ll}\text{response type of (ii) that a person checked in,} & + + + + + - \\ \text{his response in (i)} & + - + + + - \end{array}\right\};$$

the consistent number (number of matches) is 5, and if

$$\left\{\begin{array}{ll}\text{response type of (ii) that a person checked in,} & + + - - - - \\ \text{his response in (i)} & + - - - + - \end{array}\right\}$$

the number is 4. Using this number we stratify the persons.

Characterististic / Stratum number	The number of consistent items of (i) with (ii)
0	0
1	1
2	2
⋮	
R	R

Thus we count the frequency of consitent items (number of matches of signs) in each stratum with respect to every item. This is shown as below.

Frequency of consistent items (number of matches of signs) in each stratum

Item / Number of stratum	1	2	3	…………………………	R	Total
0	0	0	0	…………………………	0	0
1	n_{11}	n_{12}	n_{13}	…………………………	n_{1R}	$n_{1\cdot}$
2	n_{21}	n_{22}	n_{23}	…………………………	$n_{\ R}$	$n_{\ .}$
⋮						
R	n_{R1}	n_{R1}	n_{R3}		n_{RR}	$n_{R\cdot}$
Total	$n_{\cdot 1}$	$n_{\cdot 2}$	$n_{\cdot 3}$		$n_{\cdot R}$	$\underline{n}$

Next we calculate the frequency distribution of items in each stratum. Let $p_{ij}=n_{ij}/n_{i\cdot}$, $L_{\cdot j}/\underline{n}$. Then we obtain the following table,

Item / Number of stratum	1	2	....	R	Total
1	p_{11}	p_{12}		p_{R}	1
2	p_{21}	p_{22}		p_{R}	1
⋮	⋮	⋮		⋮	⋮
R	p_{R1}	p_{R2}		p_{RR}	1
Total	$L._1$	$L._2$		$L._R$	1

By comparing these distribution patterns, we can find the weight of each item. If we have $p_{1j} \geqq p_{2j} \geqq \cdots\cdots \geqq p_{Rj}$ in the j-th item, and $L._j$ is large, the weight of the j-th item is considered to be heavy. If, the relation $p_{1j} < p_{2j} < \cdots\cdots < p_{Rj}$ holds and $L._j$ is small, the weight is considered to be light. As a measure to known this, J_j and L_j are considered,

$$\text{(i)} \qquad J_j = \frac{\sum_{i=1}^{R} \sum_{k>i} (p_{ij} - p_{kj})}{\frac{1}{2} R(R-1)}, \qquad (j = 1, 2, \cdots\cdots, R)$$

in which the patterns, of p_{ij} near to each end point are attached importance to. J_j is the larger, the weight is the more heavy. If p_{ij} is a random variable (p_{ij} is correlated with p_{ik}), mean of J_j, $E(J_j)$ and the variance of J_j is easily calculated, because J_j are random variables. So, in the statistical sense, comparisons of J_j are possible. These results must be interpreted.

$$\text{(ii)} \qquad L_j = L._j = n._j/n \qquad j = (1, 2, \cdots\cdots, R)$$

This means the frequency of consistency in the j-th item. L_j is the larger, the weight of the j-th item is the more heavy. If L_j is a random variable, the mean and variance are easily calculated.

These two measures have different phases. The weight of the j-th item may be well mensured by the weight vector (J_j, L_j). The J_j is reasonably interpreted. The frequency in the old number of strata contains the content. The frequency in the young number of strata contain much more tne meaning of weight. The difference between those has important meaning in the sense of weight. The alsolute value of frequency has also the meaning of weight in a sense, though content and others are mixtured. So it is considered to be reasonable to interprete the weight of each item by (J_j, L_j).

(b) The idea is near to (a). We observe the patterns of consistency and in consistency of items in each stratum. We think that a persnn's response hè checked freely in (i) is drawn towards his response in (ii), since the power of consistent items is more powerful than that of inconsistent items.

This is described as below. For example, if

$$\left\{\begin{array}{ll}\text{the type of response of (ii) of a person} & +\ +\ +\ +\ +\ - \\ \text{his response in (i)} & +\ -\ +\ +\ +\ -\end{array}\right\}$$

we write $(1, 3, 4, 5, 6) < (2)$ symbolically, where $1, 2, 3, 4, 5, 6$ are the number (order) of item, and if

$$\left\{\begin{array}{ll}\text{the type of response of (ii) of a person} & +\ +\ -\ -\ -\ - \\ \text{his response in (i)} & +\ -\ -\ -\ +\ -\end{array}\right\}$$

we write $(1, 3, 4, 6) > (2, 5)$ symbolically.

This means that the item group $(1, 3, 4, 5, 6)$ are more powerful than (2) and the item group $(1, 3, 4, 6)$ are than the item group $(2, 5)$.

Then we call the left hand $(>)$ the more powerful group. The frequency distribution of items in this group is shown to be $(n_{\cdot 1}, n_{\cdot}\,, \cdots\cdots, n_{\cdot R})$ or $(L_{\cdot}\,, L_{\cdot}\,, \cdots\cdots, L_{\cdot R})$ (see (a)). We consider the rate Q_j of + signs in the j-th item in the given types of (ii) to the total number of the given types, where $Q_j = \frac{R+1-j}{R+1}$, $j = 1, 2, \cdots\cdots, R$.

The item / Type in (ii)	1	2	…………………	j	………………………………	R
1	+	+	⋮	+	⋮	+
2	+	+	⋮	+	⋮	−
⋮	⋮	⋮	⋮	⋮	⋮	⋮
$R+1$	−	−	⋮	−	⋮	−
rate of + sign	$\frac{R}{R+1}$	$\frac{R-1}{R+1}$	⋮	$\frac{R+1-j}{R+1}$	⋮	$\frac{1}{R+1}$

For example, $R=6$,

$$Q_1 = \frac{6}{7}, \quad Q_2 = \frac{5}{7}, \quad \cdots\cdots, \quad Q_6 = \frac{1}{7}.$$

Next let p_j be the relative frequency of + sign of the j-th item in the responses of n persons in the test (i).

So we consider the following matching theory. Suppose that a person who has − sign in the j-th item in the test (i) choices a type at random, and a person who has + sign in the j-th item, choices a type at random. Then the probability that he choices the type + sign in the j-th item, is Q_j, and the probability that he choices the type of − sign, is $(1-Q_j)$.

So if each person who has a response in (i) choices a type in (ii) at random, it is proved that the expectation E_j of the number of matches of signs in the j-th item is

$$E_j=n(P_jQ_j+(1-P_j)(1-Q_j))$$

and the variance $\sigma_{E_j}{}^2$ is $\sigma_{E_j}{}^2=nQ_j(1-Q_j)$

In the more powerful group, the distribution of consistent items is $(n_{\cdot 1}, n_{\cdot 2}, \cdots\cdots, n_{\cdot L})$.

By comparing these, the weight of each item is measured in some sense. Because the fact that we have more consistent items in our survey than in the case responsed at random, reveals us the existence of power which is interpreted to be weight. Let M_j be this measure

$$M_j=\frac{n_{\cdot j}-E_j}{\sigma_{E_j}} \quad \text{or} \quad M_j=P_r\left\{\frac{|x-E_j|}{\sigma_{E_j}}<\frac{|n_{\cdot j}-E_j|}{\sigma_{E_j}}\right\}$$

where x is a random variable, the mean and variance of which are E_j and σ_{E_j} respectively, the distribution function of which is appropriately determined in every case. M_j is the larger, the frequency of constant items is the higher. So the weight is the more heavy. The magnitudes of M_j can be statistically interpreted. Then the weight vector is $(M_1, M_2, \cdots\cdots, M_K)$ and can be regarded as pure weight. If $n_{\cdot j}$ is a random variable, the mean and variance of M_j in this sense are calculated. Bisides this, we regard the more powerful group and the other proup as two strata and we can quantify the items to maximize the correlation ratio by the simular methods to (1) [1].

The idea underlying the method mentioned above is to find the weight of an item in relation to other items by giving the restrictions to the way of responding, which result in weights of items through the process of considerations, and by comparing the free responses with the restricted responses. By the conception of weights of items, the inconsistency patterns are changed into a type. Restrictions bring us the weights of items. Thus we shall be able to find the weights.

Or more simply, weight x_j of the j-th item can be determined. We consider

$$V=x_1{}^2f_1+\cdots\cdots+x_K{}^2f_K,$$

where f_j is the frequency of inconsistency betwden reponses (i) and (ii) in the j-th item, and $f_j=n-n_{\cdot j}$, n being the total size, $n_{\cdot j}$ having the meaning mentioned above. If we determine x_j to minimize the V under the condition of $\sum_{j=1}^{K} x_j=1$, they are our solution.

These methods are similar to each other. But these numerical values (which have complicated meaning) given to items can not be interpreted simply and must be interpreted from various points of view, on basis of the process of quantification. What methods to use depends on our actual purpose.

(c) Especi l y, quantification of intensity.

In this case, response pattern of (i) and (iii) are used. L. Guttman suggests that the intensity of attitude of those, who have extremely positive or negative responses in content, is the higher if test items are scalable. And these are certified by experiments. From the point of wording in Japanese language, his wording technique can not be used in Japan. Then we consider that we measure intensity by the attitude towards the inverse. So we use biased questionnaires, (i) and (iii). By comparing these response patterns in (i) and (iii), we can measure intensity of attitude. It is considered that the intensity of those who have + sign response in (i) and − sign response in (ii) in each item or − sign response in (i) and + sign response in (iii) in each item, is high, the other is weak.

We may measure the intensity of a person by the number C of items which have the pair + in (i) and − in (iii) in every item and − in (i) and + in (iii). And we stratify the persons by the response of (ii). So we have the mean and variance of C in each stratum. If the behaviour of mean value with strata, is U shaped (see Fig. 2). L. Guttman's proportion will be certified. Instead of using strata specified by (ii) as criterion of content, we may use the numerical score α_i given to i-person by the method (1) (a). Then we quantify the intensity as followings. C varies from O to R. So we give $s_l (l=0, 1, \cdots\cdots, R)$ to the l-th class of C, of which the number C is equal to l. Let us consider the correlation ratio $\eta(\alpha, s)$ concerning with s on the base of α, with respect to persons. If we quantify s_l to maximize $\eta(\alpha, s)$, $(s_0, s_2, \cdots\cdots, s_R)$ will become intensity vector coresponding to our purpose. $\eta(\alpha, s)$ is considered to be a measure of predictable power of intensity from content. The value of η tells us the relation between content and intensity. In our survey, we knew that the C values represented the intensity of attitude well. Those who changed their attitude after group discussions were low in intensity (small C values), and those who did not change their attitude after it were high (large C values). The correlation ratio

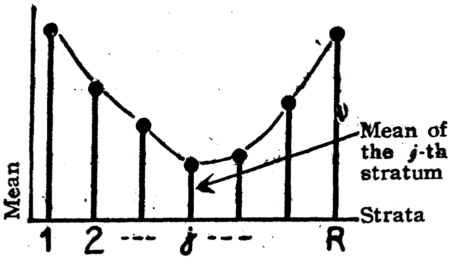

Fig. 2

$\eta(h, C)$ is 0.75, where h is the degree of change of attitude. Intensity seems not to be so simple as the U-shaped theory (Guttman). Those who have extremely positive or negatives responses are not always high in intensity. There are various cases. There are some cases where those who have extremely positive responses are low in intensity if they are small in unmber, because many persons are negative and so the intensity of positive attitude is low. And other cases where those who have extremely positive response are high in intensity if they are small in number, because a few persons are positive and oppose the majority and so the intensity of positive attitude is high. According to the fact, intensity must be measured in every case and the patterns must be interpreted scientifically from various points of view.

The results of above discussions in this section depend on the characteristics of sample (persons). And so, when we consider the quantification of items in pretest, we must use random sample from the population in which we want to obtain some propositions. While we can analyse the various groups' structures or group characteristics by the difference of the numerical values obtained by the above methods between various groups.

§ 4. A Method of Grouping in Sociometry

As a method of representing human interaction patterns in sociometry [8], matrix representation will be able to be used. Now suppose that the human interaction of persons of size n, is represented by sociometry method as below.

e_{ij} means the attitude that i-person has towards j-person, e_{ji} means that attitude that j-person has towards i-person, that is to say, the treatment that i-person receive from j-person.

Person	1	2	$\cdots$	j	$\cdots$	n
1	e_{11}	e_{12}	$\vdots$	e_{1j}	$\vdots$	e_{1n}
2	e_{21}	e_{22}	$\vdots$	e_{2j}	$\vdots$	e_{2n}
3	e_{31}	e_{32}	$\vdots$	e_{3j}	$\vdots$	e_{3n}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
i	e_{i1}	e_{i2}	$\cdots$	e_{ij}	$\cdots$	e_{in}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	e_{n1}	e_{n2}	$\vdots$	e_{nj}	$\vdots$	e_{nn}

So the i-th row means the attitude that i-person has towards the other persons, and the j-th column means the treatmənt that j-person receive from the others.

Suppose that these e_{ij} which represent human relations, are quantified previously by some methods fulfilling the property of validity.

Generally $e_{ij} \neq e_{ji}$, and e_{ii} is of no sequence, so that it need not be measured. Then it is necessary to make clear the group structure in n persons from the matrix representation of human relations and interactions. For this purpose, for example, the existence of sub-groups must be ascertained. Now let e_{ij} be the representation of the degree of intimacy that i-person has towards j-person, e_{ij} being the larger, the degree being higher. Of course, $e_{ij} \gtreqless 0$.

In this case we may say that the sub-groups consists of persons who are intimate with each other. We give a numerical value x_i to i-person from a stand point, in order to represent the group structure (for example, to find the existence or non-existence of sub-groups). The pattern of $x_i (i=1, 2, \cdots\cdots, n)$ may be considered to show the group structure. We quantify each person from the stand point that the persons, the x values of whom are near to each other, will be intimate with each other, and the persons, the x values of whom are distant, will be alienated from each other. Of course, these are only spoken relatively. The vector X is an index of group structure. Now we consider an matric Q

$$Q = -\sum_{i=1}^{n} \sum_{j=1}^{n} e_{ij} (x_i - x_j)_2$$

If e_{ij} is larger, x_i and x_j are desirable to be near from the point of view of quantification, and if e_{ij} is smaller, x_i and x_j are desirable to be distant. This mean that x values must be given to maximize the Q values. For this, we must consider to give the x values to maximize Q value under some conditions. This consideration will be appropriate. It is reasonable that we make the variance of x_i constant as the restriction condition,

$$\frac{1}{n} \sum_{i=1}^{n} x_i^2 - \bar{x}^2 = a^2, \quad \text{where} \quad \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i, \quad a \text{ is constant } (\neq 0).$$

Without loss of generally we take $\bar{x} = 0$.

So the restriction is

$$\frac{1}{n} \sum_{i=1}^{n} x_i^2 = a^2$$

It is our purpose to maximize the Q with respect to $x_i (i=1, 2, \cdots\cdots, n)$, under the condition $\frac{1}{n} \sum_{i=1}^{n} x_i^2 = a^2$ (x_i lies on the surface of a sphere, the raduis of which is a, $a \neq 0$). Q is a differentiable function (of course, continuous) and the domain of x_i is bounded as above mentioned. So the general theorem tells us that Q attains a maximum value in the domain of x_i. Thus we can generally require the x_i values $(i=1, 2, \cdots\cdots, n)$.

The practical method of require x_i is as followings.

$$G=\frac{-\sum_{i=1}^{n}\sum_{j=1}^{n}e_{ij}(x_i-x_j)^2}{\frac{1}{n}\sum_{i=1}^{n}x_i^2}$$

It is our purpose to maximize G. The maximum value of G is independent of a.

To maximize G, with respect to x_l, $(l=1, 2, \cdots\cdots, n)$

$$\frac{\partial G}{\partial x_l}=0$$

So
$$-\frac{\partial}{\partial x_l}\Big(\sum_{i=1}^{n}\sum_{j=1}^{n}e_{ij}(x_i-x_j)^2\Big)-G\frac{\partial}{\partial x_l}\Big(\frac{1}{n}\sum_{i=1}^{n}x_i^2\Big)=0$$

$$\Big\{-\sum_{\substack{j=1\\ j\neq l}}^{n}(e_{lj}+e_{jl})\Big\}x_l+\sum_{\substack{j=1\\ j\neq l}}^{n}(e_{lj}+e_{jl})x_j=\lambda x_l, \qquad (l=1, 2, \cdots\cdots, n)$$

where
$$\lambda=\frac{G}{n},$$

Let
$$e_{lj}+e_{jl}=a_{lj}$$

$$\Big\{-\sum_{\substack{j=1\\ j\neq l}}^{n}a_{lj}\Big\}x_l+\sum_{\substack{j=1\\ j\neq l}}^{n}a_{lj}x_j=\lambda x_l \qquad (l=1, 2, \cdots\cdots, n)$$

To solve these equations is our problem. λ's are considered to be latent roots of matrix $B=(b_{ij})$, where

$$b_{ii}=-\sum_{\substack{j=1\\ j\neq i}}^{n}a_{ij}, \qquad b_{ij}=a_{ij} \qquad (i\neq j)$$

The maximum of value of λ is the value of G required, because

$$\lambda=\frac{G}{n} \quad \text{and} \quad \text{Max}\,\lambda=\frac{1}{n}\,\text{Max}\,G.$$

The latent vector $(x_1, x_2, \cdots\cdots, x_n)$ corresponding to the maximum λ is that we require. Matrix B is symetric. So the latent roots are real. The maximum value of λ is calculated from the secular equation $|B-\lambda E|=0$, where E is unit matrix. It is possible to require the maximum by the well known successive approximation method [7, 9]. Using the relation $\sum_{i=1}^{n}b_{ij}=0$, $\sum_{j=1}^{n}b_{ij}=0$, we know that $|B-\lambda E|=0$ has the root of $\lambda=0$. The corresponding vector is constant except zero. In any case, this is not our solution, for it has no discriminative power. Now let λ_1 be maximum, let $x_1, x_2, \cdots\cdots, x_n$ be the corresponding vector.

So the relation $\sum_{i=1}^{n} x_i=0$ holds, because if roots $\lambda=0$, $\lambda=\lambda_1$ of symetric matrix are distinct, then the corresponding latent vectors are orthogonal to each other. In our case $x_2/x_1, x_3/x_1, \cdots\cdots, x_n/x_1$. $(x_1\neq 0)$ are required by the calculation method.

From the condition $\frac{1}{n}\sum_{i=1}^{n} x_i^2=a^2$, $x_1, x_2, \cdots\cdots, x_n$ are required.

The hardness of sub-groups is measured, in some sense, by $\eta^2=\frac{\sigma_b^2}{\sigma^2}$, where σ^2 is total variance and σ_b^2 is variance between sub-groups, when we group n persons by using by the numerical values x's.

[Example]

Matrix of e_{ij}

Person	1	2	3	4
1		2	−2	−1
2	2		−1	−1
3	−2	−1 .		1
4	−2	−2	2	

Matrix of b_{ij}

$$\begin{pmatrix} 3 & 4 & -4 & -3 \\ 4 & 1 & -2 & -3 \\ -4 & -2 & 3 & 3 \\ 3- & -3 & 3 & 3 \end{pmatrix}$$

Then from the first step of successive approximation that gives the largest of the absolute value of λ, we obtain $\lambda=12$. $\lambda>0$, so it is shown by the theory of calculation that this λ is the largest latent root.

Thus if $a^2=1$

$$x_1=\ 1.131, \qquad x_2=\ 0.860,$$
$$x_3=-1.007, \qquad x_4=-0.984.$$

Sub-group (x_1, x_2), (x_3, x_4) are recognized. The between variance of two sub-groups as 0.99, so the correlation ratio is equal to 0.99.

(b) The above method uses the relation, $a_{ij}=e_{ij}+e_{ji}$. In some case, this is not appropriate. Then we consider the quantification two-dimensionally. One dimension consists of the attitude that a person has towards other persons. The other consists of the treatment that a person receive from other persons. Suppose that i-person is given numerical value (x_i, y_i), where x_i is the former, y_i is the latter. Every person has a position in two dimension space. The position will allow meaningful interpretations on group structure (or human behaviour) from the psychological and sociological points of view. When (x_i, y_i) are to be given to each person, we consider similarily to (a)

$$Q_1=\frac{-\sum_{i\neq j}^{n}\sum^{n}\left(\sum_{k}\frac{1}{(e_{ik}-e_{jk})^2}\right)(x_i-x_j)^2}{\frac{1}{n}\sum_{i=1}^{n} x_i^2}$$

$$Q_2=\frac{-\sum_{i\neq j}^{n}\sum^{n}\left(\sum_{k}\frac{1}{(e_{ki}-e_{kj})^2}\right)(y_i-y_j)^2}{\frac{1}{n}\sum_{i=1}^{n}y_i^2}$$

where, if $\sum_{k=1}^{n}(e_{ik}-e_{jk})^2=0$ or $\sum_{j=1}^{n}(e_{ki}-e_{kj})^2=0$, then we consider, $x_i=x_j$ or $y_i=y_j$, and omit i or j and operate the above relations; this is reasonable from the point of interpretation. And we determine (x_i, y_i) so as to maximize Q_1, Q_2 or $F(Q_1, Q_2)$, where F is a differentiable function that is to be given a valid meaning, with respect to $x_i, y_i (i=1, 2, \cdots\cdots, n)$. If we put

$$e_{ij}'=\sum_{k}^{n}\frac{1}{(e_{ik}-e_{jk})^2} \quad \text{or} \quad e_{ij}''=\sum_{k}^{n}\frac{1}{(e_{ki}-e_{kj})^2}$$

this problems reduce to (a) mathematically. So the solution is required similarily to (a). But in this case, e_{ii} must be defined in relation to others.

The meaning of e_{ij}' or e_{ij}'' is easy to interpretate and these e_{ij}', e_{ij}'' have the same tendency as e_{ij} in (a). But we must be aware of that Q_1 and Q_2 have different meaning from Q. The method mentioned above is considered to be reasonable though a hypothetical metrical system is used. But the form $Q(Q_1, Q_2)$, which fulfils the property of validity and reproducibility, is to be considered throughly from the mathematico-statistical point of view.

REFERENCE

[1] L. GUTTMAN: An Approach for Quantifying paired Comparison and Rank Order, *The Annals of Mathematical Statistics*, Vol. 17, No. 2, 1946.

[2] C. HAYASHI: On the Quantification of Qualitative Data from the Mathematico-Statistical Point of View, (An approach for applying this Method to the Parole Prediction), *Annals of the Institute of Statistical Mathematics*, Vol. II, No. 1, 1950.

[3] C. HAYASHI: Mathematico-Statistical Methods in Sociometrics, *Mathematics*, No. 3, Vol. 3, 1951.

[4] C. HAYASHI: Quantification Methods from the Mathematico-Statistical Point of View, *Research Memoir of the Institute of Statistical Mathematics*, No. 1, 2, 3, 11, Vol. 6, 1950.

[5] C. HAYASHI, F. MARUYAMA, M. D. ISHIDA, S. TAKAKURA, M. TAGUMA, M. SUZUKI: Sampling Design in Literacy Survey, *The Annals of the Institute of Statistial Mathematics*, Vol. II, No. 1, 1950 (or The Committee of Literacy Survey, Literacy of the Japaneses, 1951, Tokyo University Press)

[6] P. HORST, P. WALLIN, L. GUTTMAN, F. B, WALLIN, J. A. CLAUSEN, R. B. REED, M. W. RICHARDSON, E. ROSENTHAL: The Prediction of Personal Adjustment, *Social Science Research Council*, 1941.

[7] KARÂMAN and BIOT: Mathematical Methods in Engineering.

[8] D. KRECH and R. S. CRUTCHFIELD: Theory and Problems of Social Psychology. 1948, McGraw Hill.

[9] R. v. MISES und H. POLLACZEK-GEIRINGER: Praktische Verfahren der Gleichungsauflösung, *Zeitscchrift für Angewandte Mathematik und Mechanik,* Vol. 9, Heft 1, Heft 2, 1929.

[10] J. v. NEUMAN and O. MORGENSTERN: Theory of Games and Economic Behavior 1946. Princeton University Press.

[11] S. A, STOUFFER, L. GUTTMAN, E. A. SUCHMAN, P. F. LAZARSFELD, S. A. STAR, J. A. CLAUSEN: Mesurement and Prediction, 1950. Princeton University Press.

(*The Institute of Statistical Mathematics.*)

CORRECTIONS TO

" ON THE PREDICTION OF PHENOMENA FROM QUALITATIVE DATA AND THE QUANTIFICATION OF QUALITATIVE DATA FROM THE MATHEMATICO-STATISTICAL POINT OF VIEW "

C. HAYASHI

In the above titled article (Ann. Inst. Statist. Math., Vol. 3, No. 2 (1952), 69–98), the following corrections should be made.

i) Page 69, line 2 from the bottom.
Read " point of view consists in " instead of " point of consists in view ".

ii) Page 84, line 14.
Read $\frac{\partial\sigma_{p^2}}{\partial x_{uv}}$ instead of $\frac{\partial\sigma^2}{\partial\sigma_{uv}}$.

iii) Page 84, line 17.
Read $\sum_{j}^{R}\sum_{k}^{2} h_{uv}(jk)$ instead of $\sum_{j=1}^{R}\sum_{k=1}^{2} h_{uv}(jk)$.

iv) Page 84, line 23.
Read " require the largest maximum " instead of " require the maximum ".

v) Page 87, line 17.
Insert the following sentense: " We have only to decide the dividing points in reasonable sense, taking the above equations into consideration with the existence of their solution. If p_i is not applicable in the valid sense, dividing points are decided by max-min method (see [2]) ".

vi) Page 90, line 7.
Read $(1, 3, 4, 5, 6)>(2)$ instead of $(1, 3, 4, 5, 6)<(2)$.

vii) Page 90, line 5 from the bottom.
Read P_j instead of p_i.

viii) Page 91, line 7.
Read $\sigma_{E_j}{}^2=nR_j(1-R_j)$, $R_j=P_jQ_j+(1-P_j)(1-Q_j)$ instead of $\sigma_{E_j}{}^2=nQ_j(1-Q_j)$.

Errata

C. Hayashi, "On the Quantification of Qualitative Data from the Mathematico-Statistical Point of View," *this Annals*, Vol. II, No. 1, 1950.

Page 41, line 36, insert after "increases" the following: "under some conditions with respect to weight vector".

Page 43, line 5, insert the sentense: "If the condition $\sum_i^n l_i^2 \to \infty (n \to \infty) (l_i$ finite) is satisfied and if we take $a_i = \frac{l_i}{\sum l_i} n$, then $\frac{\max a_i^2}{\sum a_i^2} \to 0 \, (n \to \infty)$ holds. This satisfies the condition mentioned above with respect to weight vector. We adopt the weights mentioned above in the valid sense if the success rate of prediction becomes 1 when we take infinitely many apprapriate predicting factors".

C. Hayashi, "On the Prediction of Phenomena from Qualitative Data and the Quantification of Qualitative Data from the Mathematico-Statistical Point of view," *this Annals*, Vol. III, No. 2, 1952.

Page 69, line 2 from the bottom, read "point of view consists in" instead of "point of consists in view"

Page 84, line 14, read $\frac{\partial \sigma_p^3}{\partial x_{uv}}$ instead of $\frac{\partial \sigma^2}{\partial \sigma_{uv}}$.

Page 84, line 17, read $\sum_j^R \sum_k^2 h_{uv}(jk)$ instead of $\sum_{j=1}^R \sum_{k=1}^2 h_{uv}(jk)$.

Page 84, line 23, read "require the largest maximum" instead of "require the maximum"

Page 87, line 17, insert the following sentense: "We have only to decide the dividing points in resonable sense, taking the above equations into consideration with the existence of their solution. If p_i is not applicable in the valid sense, dividing points are decided by max-min method (see [2])".

Page 90, line 7, read $(1, 3, 4, 5, 6) > (2)$ instead of $(1, 3, 4, 5, 6) < (2)$.

Page 90, line 5 from the bottom, read P_j instead of p_i.

Page 91, line 7, read $\sigma_{Ej}^2 = nR_j(1-R_j)$, $R_j = P_jQ_j + (1-P_j)(1-Q_j)$ instead of $\sigma_{Ej}^2 = nQ_j(1-Q_j)$.

H. S. Konijn, "Remark on the Characterization of Minimax Procedures" *this Annals* Vol. IV, No. 2, 1953.

Line 2 of the "Introduction", read "Bayes solutions" instead of "Bayes solution".

Page 103, line 2 of the "Statement of result", read "Byes solutions" instead of "Bayes solution".