OBJECTIVE COMPARISONS

by

G.Rasch.

Lectures given at the UNESCO Seminar,
Voksenåsen,Oslo,1964.

1. Introduction.

The concept of "objectivity" raises fundamental problems in all sciences. For a statement to be scientific "objectivity" is required. However, exactly what "objectivity" means is disputed among philosophers and I am not going to enter into that debate.

We will take a quite specific point of departure. The leading principles originated in test psychology and the present theory will be formulated within a psychological framework with "individual", "stimulus", "situation", "response" and "reaction" as central terms. In the end you may realize that the theory does not deal with psychology in particular, but has much wider scope.

It is convenient to start by considering an intelligence test - a sequence of questions or items which are ordered by increasing difficulty. 1000 recruits in the Danish army answered each of the items right or wrong, + or -. All persons completed all items, no items were slipped (slightly idealized). The information from this investigation may be represented in a (0,1) data matrix with e.g. individuals as rows and items (stimuli) as columns.

In traditional analysis, a raw score x - number of correct answers - is counted for each individual. The individuals are considered to be a "representative sample" from some more or less well-defined "universe" or population, the distribution of which with regard to x is formed and cut into a number of pieces. With an enumeration of these pieces the standardization is completed and it is then claimed that "intelligence" has been "measured". Sometimes the analysis is more elaborate in using for instance inter-item correlations. This kind of analysis is, however, of the same type as the usual standardization in so far as it refers to a population and the information gained depends upon the population.

Skinner has vigerously attacked this kind of statistical analysis, maintaining that the order to be found in animal and human behaviour should be extracted from investigations into individuals and that psychometric methods are inadequate for such purposes since they deal with groups of individuals. Zubin expresses a similar point of view concerning abnormal psychology: "Recourse must be had to individual statistics treating each patient as a separate universe. Unfortunately, present day statistical methods are entirely group-centered so that there is a real need for developing individual centered statistics.

A basic aim of the present work is just to take care of the individual. The first step is to realize a basic uncertainty regarding whether an individual will answer a question + or -. It may happen that a clever person gives a wrong answer to an easy question, he may temporarily feel uneasy or be distracted by outside noise. He may also be irritated by easy questions, and more or less purposively give wrong answers - later become interested and answer correctly. Conversely a stupid person may hit upon a correct answer by chance or even by way of a wrong train of thoughts.

Some psychologists are opposed to probabilistic models because they prefer to see a cause behind every act.

A probabilistic model, however, does not imply that the behaviour in a test situation <u>is</u> haphazard, but only that the data may be <u>represented</u> by a chance model.

A model is not meant to be true. Even in classical physics models are temporary - good enough for some purposes. In the last analysis, however, even deterministic models in physics are in need of probabilistic reformulations, since the observational data themselves do not follow deterministic laws, only the parameters in the models do so. Thus if psycologists insist on deterministic models they really are trying to be "plus royal que le roi".

2. A simple model for measuring.

Anyhow, we shall start by allotting a probability to an individual ν of answering an item i correctly and write

$$(2.1) \qquad p\{+|\nu,i\} = \frac{\lambda_{\nu i}}{1+\lambda_{\nu i}} \, , \quad p\{-|\nu,i\} = \frac{1}{1+\lambda_{\nu i}}, \quad \lambda_{\nu i} \geq 0 \, .$$

If observations could be repeated we might estimate

$$(2,2) \qquad \lambda_{\nu i} = \frac{p\{+|\nu,i\}}{p\{-|\nu,i\}}$$

directly, but in intelligence testing independent replication usually are out of question. For estimation purposes a further specification of $\lambda_{\nu i}$ therefore would be necessary. With a view to mathematical simplicity I shall suggest a partition of $\lambda_{\nu i}$ into two factors, one pertaining to the person, the other one to the item, i.e.

$$(2.3) \qquad \lambda_{\nu} = \xi_{\nu} \varepsilon_{i}$$

- not with standing the apparent boldness of this assumption.

Later on much stronger grounds for attempting this choice will emerge. However, it would not be wise just going ahead and make use of the model. Models should never be believed in, they are never more than tentative. Therefore investigations of a model should be directed such as to disclose its weaknesses pointing to substantial improvements, if possible.

Many statisticians and users of statistics have been somewhat vague and loose on this point. In many problems the use of a certain class of statistical specifications has been recommended, but an effective examination of the adequacy of the chosen specification is rarely seen. As a case in point it may be mentioned that a simple additive model for many years has been in common use in two way analysis of variance with one observation per cell, but not until 195o a paper appeared on how to check the model (J.W.Tukey, Biometrics, Vol.5, p. 233-242).

In our discussion therefore much attention will be paid to ways and means for effectively controlling the applicability (not to be confused with the "validity") of the model.

The model as given by (2.1) and (2.3) implies that

$$p\{+|\nu,i\} = \frac{\xi_\nu \varepsilon_i}{1+\xi_\nu \varepsilon_i}, \quad p\{-|\nu,i\} = \frac{1}{1+\xi_\nu \varepsilon_i} \tag{2.1a}$$

and accordingly

$$\frac{p\{+|\nu,i\}}{p\{-|\nu,i\}} = \xi_\nu \varepsilon_i . \tag{2.3a}$$

As presented this model is multiplicative, but turning it into an additive model is just a matter of transforming logarithmically and is no particularly important question.

The unit of measurement is arbitrary. We might choose the parameter of a standard person for a unit, but that would be impractical as he cannot be preserved. More suitable a particular item, say i = o, may be chosen as having $\varepsilon_o = 1$ and then

$$\xi_\nu = \frac{p\{+|\nu,o\}}{p\{-|\nu,o\}} \tag{2.4}$$

appears to be the "betting odds" - based upon an objective probability - of person ν for a correct answer to the standard item. Next we may "hunt for" a standard person (ν = o) with ξ_o = 1 and then

(2.5) $$\varepsilon_i = \frac{p\{+|o,i\}}{p\{-|o,i\}}$$

becomes the betting odds for the standard person for giving a correct answer to item i. In both cases the parameter is a simple function of a particular probability.

Considering now what happens to $p\{+|\nu,i\}$ in various combinations of items and persons, it is obvious that for a fixed ε_i persons with small parameter values have a small chance for answering correctly, while the probability approaches unity when ξ_ν is very large.

Tentatively we may therefore think of the personal parameters as his "degree of ability" in such problems as were given in the intelligence test in question.

Conversely, if items with small and large values of ε_i are given to the same person he gets small and large chances, respectively, for giving the right answers. Thus ε_i may be thought of as "the degree of easiness" of the item and the reciprocal $\delta_i = 1/\varepsilon_i$ as "the degree of difficulty" - with certain limitations, of course, as for instance that the person in question should possess a cultural background that makes the items comprehensible to him.

To (2.1a) we shall add the assumptions that the answers given by one person are considered to be independent of those of other persons and also of the answers he has given to preceding questions.

Off hand, psychologists often dispute the latter assumption, but on second thought they usually give in. Traditionally, answers are scored 1 or 0 and correlations between items are computed. These correlations quite often amount to + 0,90 or even more. On this background the assumption of independence may be shocking. There is,however, no discrepancy here; on the contrary, the positivity of the correlations may be directly inferred from the assumptions, as soon as it is clearly understood, that the independence is assumed only within each person. In fact, for any two items i and j we have

(2.6) $$p\{+|\nu,i\} = \frac{\xi_\nu \varepsilon_i}{1+\xi_\nu \varepsilon_i}, \quad p\{+|\nu,j\} = \frac{\xi_\nu \varepsilon_j}{1+\xi_\nu \varepsilon_j}$$

from which it is seen that

ξ_ν small makes both p's close to zero

and

ξ_ν large " " " " " unity.

Thus the p's follow each other and this produces a correlation over a group of individuals with differing ξ's. Accordingly stochastic independence within individuals is fully compatible with a parametrically generated interdependence within a population.

As an analogy from the elementary theory of errors we may consider duplicate chemical observations:

$$x_\nu = \xi_\nu + u_\nu, \qquad y_\nu = \xi_\nu + v_\nu$$

where ξ_ν is, say, the concentration of nitrogen in some solution, u_ν and v_ν being the errors of measurement, presumed to be independent. For each solution, then x_ν and y_ν are uncorrelated, but on titration of several solutions with different concentrations the variation of ξ_ν would produce a "correlation", the strength of which would depend on this variation as compared to the variances of the error terms.

Returning to the intelligence test we should perhaps regard a test situation as a "learning" process or perhaps rather as an "adaptation" especially if the testees are unfamiliar with the type of items used. If a test with "new" items, e.g. consisting of letter matrics, is administered to a highly "intelligent" group but starting with later items, they may fail miserably. Some sort of finding out what the items are about seems to take place within the first items. Perhaps we ought to cut off the first few items from the analysis.

+ and - answers are recorded as 1 and o, respectively. This, however, is not an arbitrary scoring, but just an alternative registration from the independence assumption it follows that if we put

$$a_{\nu i} = \begin{cases} 1 \\ o \end{cases}$$

(2.1a) more compactly may be written:

$$p\{a_{\nu i}\} = \frac{(\xi_\nu \varepsilon_i)^{a_{\nu i}}}{1+\xi_\nu \varepsilon_i} \quad (2.7)$$

Consider an answer pattern for a given person:

$$(2.8) \quad p\{a_{\nu 1} \ldots a_{\nu k} | \xi_\nu\} = p\{a_{\nu 1} | \xi_\nu\} \ldots p\{a_{\nu k} | \xi_\nu\}$$

$$= \frac{(\xi_\nu \varepsilon_1)^{a_{\nu 1}} \ldots (\xi_\nu \varepsilon_k)^{a_{\nu k}}}{(1+\xi_\nu \varepsilon_1) \ldots (1+\xi_\nu \varepsilon_k)}$$

$$= \frac{\xi_\nu^{a_{\nu 1} + \ldots a_{\nu k}} \varepsilon_1^{a_{\nu 1}} \ldots \varepsilon_k^{a_{\nu k}}}{(1+\xi_\nu \varepsilon_1) \ldots (1+\xi_\nu \varepsilon_k)} \quad .$$

Here a remarkable thing has happened. Traditionally it has for no obvious reason been customary to count correct answers. But if the model is applicable we <u>must</u> do so because

$\sum_{i=1}^{k} a_{\nu i} = a_{\nu .}$ has so to speak been inflicted upon us -

$a_{\nu .}$ enters explicitly into probability of the set of answers. Thus, counting is a consequence of the model - later we shall discuss whether the model is arbitrary or not.

Accordingly we are going to consider: $p\{a_{\nu .} | \xi_\nu\}$.

A given value $a_{\nu .} = r$ can be achieved in a number of ways.

We have to find the probability for each way and then add these probabilities:

<u>r = 1.</u> In this case the realizations are

$$i: \quad 1 \;\; 2 \; \ldots \; k$$

$$a_{\nu i}: \begin{cases} 1 & o & \ldots & o \\ o & 1 & \ldots & o \\ - & - & - & - \\ o & o & \ldots & 1 \end{cases}$$

with the probabilities

$$\frac{\xi_\nu \varepsilon_1}{1+\xi_\nu \varepsilon_1} \cdot \frac{1}{1+\xi_\nu \varepsilon_2} \cdots \frac{1}{1+\xi_\nu \varepsilon_k}$$

$$\frac{1}{1+\xi_\nu \varepsilon_1} \cdot \frac{\xi_\nu \varepsilon_2}{1+\xi_\nu \varepsilon_2} \cdots \frac{1}{1+\xi_\nu \varepsilon_k}$$

- - - - - - - - - - - - - - -

$$\frac{1}{1+\xi_\nu \varepsilon_1} \cdot \frac{1}{1+\xi_\nu \varepsilon_2} \cdots \frac{\xi_\nu \varepsilon_k}{1+\xi_\nu \varepsilon_k} ,$$

the sum being

$$p\{a_{\nu .} = 1 | \xi_\nu\} = \frac{\xi_\nu (\varepsilon_1 + \ldots + \varepsilon_k)}{(1+\xi_\nu \varepsilon_1) \ldots (1+\xi_\nu \varepsilon_k)}$$

$$= \frac{\xi_\nu \gamma_1 (\varepsilon_1, \ldots \varepsilon_k)}{D(\xi_\nu)} \quad \text{, say.}$$

<u>$r = 2$.</u> Now $\binom{k}{2}$ realizations take place:

$$\begin{array}{ccccccc} i & : & 1 & 2 & 3 & \ldots & k \\ & & 1 & 1 & o & \ldots & o \\ a_{\nu i} & : & 1 & o & 1 & \ldots & o \\ & & o & 1 & 1 & \ldots & o \\ & & - & - & - & - & - \end{array}$$

with a total probability of

$$p\{a_{\nu .} = 2 | \xi_\nu\} = \frac{\xi_\nu^2 (\varepsilon_1 \varepsilon_2 + \varepsilon_1 \varepsilon_3 + \varepsilon_2 \varepsilon_3 + \ldots + \varepsilon_{k-1} \varepsilon_k)}{(1+\xi_\nu \varepsilon_1)(1+\xi_\nu \varepsilon_2) \ldots (1+\xi_\nu \varepsilon_k)}$$

$$= \frac{\xi_\nu^2 \gamma_2 (\varepsilon_1, \ldots \varepsilon_k)}{D(\xi_\nu)} .$$

<u>r arbitrary</u> ($\leq k$). Generalizing we get:

$$(2.9)\quad p\{a_{\nu .} = r \mid \xi_\nu\} = \frac{\xi_\nu^r}{D_\nu} \gamma_r (\varepsilon_1 \ldots \varepsilon_k)$$

where γ_r consists of $\binom{k}{r}$ terms each of which is a product of r different ε-parameters:

$$(2.10)\quad \gamma_r(\varepsilon_1, \ldots, \varepsilon_k) = \varepsilon_1 \ldots \varepsilon_r + \varepsilon_1 \ldots \varepsilon_{r-1}\varepsilon_{r+1} + \ldots + \varepsilon_{k-r+1} \ldots \varepsilon_k .$$

The next step which is quite decisive I shall, to the benefit of participants that are not too well versed in the relatively advanced theory of probability, express in terms which are more illustrative than exact, but add that the translation into a correct exposition should be easy for those who are familiar with the concept of conditional probability.

Consider a large number (N) of persons with the same ξ_ν. $Np\{a_{\nu .} = r \mid \xi_\nu\}$ then "stands for" the number of them which gave a total of r correct answers and similarly $Np\{a_{\nu 1}, \ldots, a_{\nu k} \mid \xi_\nu\}$ "stands for" the number of them which showed the particular pattern of answers $(a_{\nu 1}, \ldots, a_{\nu k})$. Thus

$$\frac{Np\{a_{\nu 1}, \ldots, a_{\nu k} \mid \xi_\nu\}}{Np\{a_{\nu .} = r \mid \xi_\nu\}}$$

"stands for" the relative number of persons showing that particular pattern among those who had $a_{\nu .} = r$. Reducing by N we get what is called <u>the conditional probability of the pattern $(a_{\nu 1}, \ldots a_{\nu k})$, given that $a_{\nu .} = r$:</u>

$$(2.11)\quad p\{a_{\nu 1}, \ldots, a_{\nu k} \mid a_{\nu .} = r_1 \xi_\nu\} = \frac{p\{a_{\nu 1}, \ldots, a_{\nu k} \mid \xi_\nu\}}{p\{a_{\nu .} = r \mid \xi_\nu\}}$$

Inserting (2.7) and (2.9) into (2.11) we get:

$$p\{a_{\nu 1},\ldots,a_{\nu k}\,|\,a_{\nu .}=r,\xi_\nu\} = \frac{\varepsilon_1^{a_{\nu 1}}\cdots\varepsilon_k^{a_{\nu k}}}{\gamma_r(\varepsilon_1,\ldots,\varepsilon_k)} \tag{2.12}$$

The fundamental importance of (2.12) is that $D(\xi_\nu)$ and ξ_ν cancel which implies that the conditional probability is independent of ξ_ν.

Accordingly we have established a basis for estimating the item parameters from which all sampling problems have been eliminated.The conditional probabilities (2.12) give an information on item parameters that is valid for <u>all</u> persons with $a_{\nu .} = r$, provided of course the model is applicable.

Before proceeding further it may be worth while to consider in some detail a special case of (2.12) that of only two items for person ν. The four possible combinations of answers have probabilities that follow immediately from (2.1a) and the independence:

(2.13)

i \ j	+	−	
+	$\frac{\xi_\nu\varepsilon_i}{1+\xi_\nu\varepsilon_i}\cdot\frac{\xi_\nu\varepsilon_j}{1+\xi_\nu\varepsilon_j}$	$\frac{\xi_\nu\varepsilon_i}{1+\xi_\nu\varepsilon_i}\cdot\frac{1}{1+\xi_\nu\varepsilon_j}$	$\frac{\xi_\nu\varepsilon_i}{1+\xi_\nu\varepsilon_i}$
−	$\frac{1}{1+\xi_\nu\varepsilon_i}\cdot\frac{\xi_\nu\varepsilon_j}{1+\xi_\nu\varepsilon_j}$	$\frac{1}{1+\xi_\nu\varepsilon_i}\cdot\frac{1}{1+\xi_\nu\varepsilon_j}$	$\frac{1}{1+\xi_\nu\varepsilon_i}$
	$\frac{\xi_\nu\varepsilon_j}{1+\xi_\nu\varepsilon_j}$	$\frac{1}{1+\xi_\nu\varepsilon_j}$	1

More compactly this table may be written:

(2.14)

i \ j	+	−
+	$\frac{\xi_\nu^2\varepsilon_i\varepsilon_j}{D}$	$\frac{\xi_\nu\varepsilon_i}{D}$
−	$\frac{\xi_\nu\varepsilon_j}{D}$	$\frac{1}{D}$

Repeating the above argument, if needed, we consider now the conditional probability of + in i, given + in either i or j (but not in both) :

$$(2.15) \qquad p\{+ \text{ in } i \,|\, + \text{ in either } i \text{ or } j, \xi_\nu\} = \frac{\xi_\nu \varepsilon_i / D}{\xi_\nu(\varepsilon_i + \varepsilon_j)/D} = \frac{\varepsilon_i}{\varepsilon_i + \varepsilon_j}$$

for persons with the same parameter ξ_ν.

Again the conditional probability is independent of the person parameter ξ_ν, and this implies that the persons may be chosen in any way we wish. It will easily be recognized that (2.15) is a special case of (2.12).

Suppose we have:

n_{ij} = number of persons with + in either i or j
a_{ij} = number among the n_{ij} with + in i

Remembering that persons are presumed to answer independently of each other we have a pure binomial situation:

$$(2.16) \qquad p\{a_{ij}\} = \binom{n_{ij}}{a_{ij}} \left(\frac{\varepsilon_i}{\varepsilon_i + \varepsilon_j}\right)^{a_{ij}} \left(\frac{\varepsilon_j}{\varepsilon_i + \varepsilon_j}\right)^{b_{ij}}$$

where

$$b_{ij} = n_{ij} - a_{ij} \, .$$

The best estimate of the probability $\frac{\varepsilon_i}{\varepsilon_i + \varepsilon_j}$ is:

$$(2.17) \qquad \frac{a_{ij}}{n_{ij}} \approx \frac{\varepsilon_i}{\varepsilon_i + \varepsilon_j}$$

where $\approx$ means "stochastic equality", i.e. equality, but for chance variations. $\simeq$ will be used for approximate equality in mathematical sense (e.g. $\sqrt{2} \simeq 1.41$).

The equation may be rewritten:

$$(2.18) \qquad \frac{a_{ij}}{b_{ij}} \approx \frac{\varepsilon_i}{\varepsilon_j} \, .$$

According to our assumptions this relation should hold regardless of the population from which the n_{ij} persons are taken.

We might for instance have two groups marked by ' and ", differing in, say, previous education and then

$$\left(\frac{\varepsilon_i}{\varepsilon_j}\right)' \text{ should be "equal" to } \left(\frac{\varepsilon_i}{\varepsilon_j}\right)'' .$$

This points the way to regarding the problem of whether tests are "culture-free" or not as an empirical issue.

For the utilization of (2.18) it is convenient to work with logarithms. Thus from

$$(2.19) \qquad l_{ij} = \log \frac{a_{ij}}{b_{ij}} \approx \log \varepsilon_i - \log \varepsilon_j$$

it follows that

$$(2.2o) \qquad \begin{array}{rcl} \log \varepsilon_h - \log \varepsilon_i & \approx & l_{hi} \\ \log \varepsilon_i - \log \varepsilon_j & \approx & l_{ij} \\ \log \varepsilon_j - \log \varepsilon_h & \approx & l_{jh} \\ \hline 0 & \approx & 0 \end{array}$$

which leads to estimating $\varepsilon_h : \varepsilon_i : \varepsilon_j$ and to a control on the model as well.

This method, extended to a larger number of items, has in fact been used in practice, but other methods are more powerful.

An approximative method based directly upon the definition (2.7) is the following:

Consider n_r persons with $a_{\nu .} = r$. As they are likely to have parameters fairly close to each other we shall substitute all the ξ_ν's by a common value $\xi^{(r)}$ - some sort of mean value of them - in (2.7) to obtain

$$(2.21) \qquad p\{+|\xi^{(r)},i\} \simeq \frac{\xi^{(r)}\varepsilon_i}{1+\xi^{(r)}\varepsilon_i} \cdot p\{-|\xi^{(r)},i\} \simeq \frac{1}{1+\xi^{(r)}\varepsilon_i}$$

Applying now the binomial theorem

$$(2.22) \qquad p\{a_i^{(r)}|\xi^{(r)},i\} \simeq \binom{n_r}{a_i^{(r)}} \frac{(\xi^{(r)}\ \varepsilon_i)^{a_i^{(r)}}}{(1+\xi^{(r)}\ \varepsilon_i)^{n_r}}$$

we get the estimate

$$(2.23) \qquad \frac{a_i^{(r)}}{n_r} \approx \frac{\xi^{(r)}\varepsilon_i}{1+\xi^{(r)}\varepsilon_i} \quad , \quad \frac{n_r-a_i^{(r)}}{n_r} \approx \frac{1}{1+\xi^{(r)}\varepsilon_i}$$

from which we derive the equation

$$(2.24) \qquad l_i^{(r)} = \log \frac{a_i^{(r)}}{n_r-a_i^{(r)}} \approx \log \xi^{(r)} + \log \varepsilon_i \ .$$

In practice we first construct a table with the elements $a_i^{(r)}$:

(2.25)

r \ i	1 ...	i ...	k	
1	$a_1^{(1)}$ ...	$a_i^{(1)}$ ...	$a_k^{(1)}$	n_1
⋮				
r	$a_1^{(r)}$ ...	$a_i^{(r)}$ ...	$a_k^{(r)}$	n_r
⋮				
k-1	$a_1^{(k-1)}$...	$a_i^{(k-1)}$...	$a_k^{(k-1)}$	n_{k-1}

and from this a new table with the elements $l_i^{(r)}$ in so far as $a_i^{(r)}$ differs from both 0 and n which we shall to begin with assume is the case everywhere.

For these elements and their row-averages and their column-averages $l_.^{(r)}$ and $l_i^{(\cdot)}$ and the total average $l_.^{(\cdot)}$ we have according to (2.24):

(2.26)

$$
\begin{array}{llll}
l_1^{(1)} \approx \log \xi^{(1)} + \log \varepsilon_1, \ldots, & l_i^{(1)} \approx \log \xi^{(1)} + \log \varepsilon_i, \ldots, & l_k^{(1)} \approx \log \xi^{(1)} + \log \varepsilon_k & l_{.}^{(1)} \approx \log \xi^{(1)} + \log \bar{\varepsilon} \\
\vdots & & & \\
l_1^{(r)} \approx \log \xi^{(r)} + \log \varepsilon_1, \ldots, & l_i^{(r)} \approx \log \xi^{(r)} + \log \varepsilon_i, \ldots, & l_k^{(r)} \approx \log \xi^{(r)} + \log \varepsilon_k & l_{.}^{(r)} \approx \log \xi^{(r)} + \log \bar{\varepsilon} \\
\vdots & & & \\
l_1^{(k-1)} \approx \log \xi^{(k-1)} + \log \varepsilon_1, \ldots, & l_i^{(k-1)} \approx \log \xi^{(k-1)} + \log \varepsilon_i, \ldots, & l_k^{(k-1)} \approx \log \xi^{(k-1)} + \log \varepsilon_k & l_{.}^{(k-1)} \approx \log \xi^{(k-1)} + \log \bar{\varepsilon} \\
\hline
l_1^{(\cdot)} \approx \log \bar{\xi} + \log \varepsilon_1, \ldots, & l_i^{(\cdot)} \approx \log \bar{\xi} + \log \varepsilon_i, \ldots, & l_k^{(\cdot)} \approx \log \bar{\xi} + \log \varepsilon_k & l_{.}^{(\cdot)} \approx \log \bar{\xi} + \log \bar{\varepsilon}
\end{array}
$$

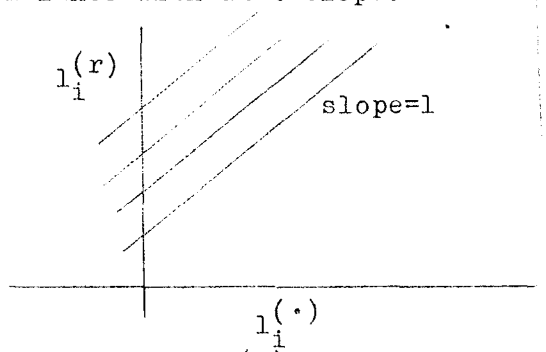
where $\log \bar{\xi}$ and $\log \bar{\varepsilon}$ denote the averages of $\log \xi^{(r)}$ and of $\log \varepsilon_i$ respectively.

Deducting e.g. the column-averages from the first row we get

$$(2.27) \qquad \ell_1^{(1)}-\ell_1^{(\cdot)} \approx \log \frac{\xi^{(1)}}{\bar{\xi}}, \dots, \ell_i^{(1)}-\ell_i^{(\cdot)} \approx \log \frac{\xi^{(1)}}{\bar{\xi}}, \dots.$$

from which it follows that on plotting the first row against the bottom row we should find a sequence of points clustering around a straight line with unit slope. The intersection with the axis estimates $\log \frac{\xi^{(1)}}{\bar{\xi}}$. Similarly for all the other rows and also for the columns, the latter leading to estimates of $\log \frac{\varepsilon_i}{\bar{\varepsilon}}$, $i=1,\dots,k$. One of the ε_i's or $\bar{\varepsilon}$ may be chosen as unity, $\bar{\xi}$ then being estimated from $\ell_{\cdot}^{(\cdot)}$. The main point in the procedure is, however, that it implies a very severe check on the model, namely that provided the model holds we should get two bunches of parallel lines with unit slope.

(2.28)

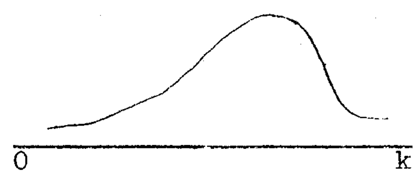


When some of the $a_i^{(r)}$ equal 0 or n_r, making $\ell_i^{(r)} \pm \infty$, the procedure is slightly modified. First we consider a rectangular subset with finite $\ell_i^{(r)}$-values. The averages thus obtained may then serve as a basis for collecting the rows and columns which were left out at the beginning. Items thus fitting the model we shall call <u>conformal.</u> Conformal items in some sense "hang together". In test construction the items should be "closely related" - for ex. not items in arithmetic and analogues in the same test.

The method worked very well for two tests L (letter matrices) and N (numerical sequences), but for two other tests, V(verbal analogies) and F (figure combinations), straight lines were obtained, to be sure, however, with a variety of slopes.

Why this disparity ? In constructing the items equal care was excercised in all subtests in applying uniform principles for all items. However, during the construction generous time limits were given with the result that the raw score distributions for V and F had the following form:

(2.29)



Thus V and F were far too easy with unlimited time in contrast to L and F where the time limits chosen scarcely had undesirable effect.

This was just one month before the tests had to be used, so what to do ? There was no time for trying to add more difficult items. The easy way out was to put on more stringent time limits to V and F. After that the appearance of the raw score distribution was satisfactory, but this - of course - spoiled the test, it was no longer a pure test of capacity, but contaminated with speed in solving the kind of problems.

This suggested reanalyzing the data by grouping people according to number of items completed. For each working speed the results turned out very well and it was possible to estimate item parameters for each working speed. Furthermore we got the same ratio of ε's for all working speeds.

One of the lessons to be learned from this is: It was not the model that failed, but the test construction certainly did ! With sloppy test construction it is not to be expected that the model should work well !

Leaving now the approximate method a side and proceeding with the general development of the theory we consider n_r persons with $a_{\nu .}=r$ and have to find the probability for their set of patterns $(a_{\nu 1},\dots,a_{\nu k})$, $\nu=1$, $\dots n_r$, remembering that the individuals answer independently

We get

$$p\{(a_{\nu 1},\ldots,a_{\nu k})|(a_{\nu .} = r)\}$$

$$= p\{a_{11} \ldots a_{1k}|a_{1.} = r\} \ldots p\{a_{n_r 1}\ldots a_{n_r k}|a_{n_r .}=r\}$$

and on applying (2.12) to each term

$$p\{(a_{\nu 1},\ldots a_{\nu k}) \mid (a_{\nu .} = r)\}$$

$$(2.3o) \qquad = \frac{\varepsilon_1^{a_{11}}\ldots\varepsilon_k^{a_{1k}}}{\gamma_r} \cdot \frac{\varepsilon_1^{a_{21}}\ldots\varepsilon_k^{a_{2k}}}{\gamma_r} \cdot \frac{\varepsilon_1^{a_{n_r 1}}\ldots\varepsilon_k^{a_{n_r k}}}{\gamma_r}$$

$$= \frac{\varepsilon_1^{a_{\cdot 1}}\ldots\varepsilon_k^{a_{\cdot k}}}{\gamma_r^{n_r}} .$$

Whereas (2.12) gives the probability of a specific answer pattern given the row sum,(2.3o) gives the probability for a <u>set</u> of answer patterns with a common row sum.

Next we imagine all matrices with $a_{\nu\cdot} = r$ and a fixed set of item marginals. All such matrices have the same probability since (2.3o) only depends upon item marginals, not on the separate $a_{\nu i}$. Therefore the probability of the marginals $a_{\cdot i}$ becomes

$$(2.31) \qquad p\{a_{\cdot 1}\ldots a_{\cdot k}|(a_{\nu .}=r)\} = \begin{bmatrix} r & \ldots & r \\ a_{\cdot 1} & \ldots & a_{\cdot k}\end{bmatrix} \frac{\varepsilon_1^{a_{\cdot 1}}\ldots\varepsilon_k^{a_{\cdot k}}}{\gamma_r^{n_r}}$$

where $\begin{bmatrix} r \ldots r \\ a_{\cdot 1}\ldots a_{\cdot k}\end{bmatrix}$ denotes the number of such matrices and may be thought of as some kind of generalized binomial coefficient.

In principle it is possible from (2.31) to derive estimates of the ε's. These estimates will depend on r and on the number of persons that happened to have $a_{\nu .}=r$, but they will not depend on the person parameters, giving rise to $a_{\nu .} = r$.

Furthermore, - still under the proviso that the model holds - estimations arising from different values of r - small, medium or large - must give substantially (meaning: apart from chance variations to be accounted for by (2.31)) the same results.

This would seem to be a very satisfactory state of affairs implied by the model suggested: By definition the ε's should characterize the items as such, irrespective of which persons - within certain cultural limits - they be applied to. (2.31) holding for every r demonstrates the possibility of obtaining inferences about the ε's which, to be sure, utilize the results actually acquired from the collection of testees, but which are uninfluenced by which values their parameters might have. In that sense they are independent of the persons used. The inference about the ε's depends on the known observations, not upon the unknown parameters, and statistically speaking the estimates of their ratios are independent of which group of persons were collected for testing.

Thus statements about the ε's are available which are <u>uninfluenced by irrelevant parameters</u>, i.e. parameters that have nothing to do with the ε's.

This is our first case of what we shall call <u>a specifically objective estimation of (or inference about) a set of parameters.</u>

Consider now the whole set of answers $((a_{\nu i}))$, $\nu = 1,\ldots n$, $i = 1,\ldots k$, all of which are, according to the model, stochastically independent:

$$p\{((a_{\nu i}))\} = \begin{matrix} p\{a_{11}\} & \cdots & p\{a_{1k}\} \\ - & - & - \\ p\{a_{n1}\} & \cdots & p\{a_{nk}\} \end{matrix}$$

$$= \begin{matrix} \dfrac{\xi_1^{a_{11}}\varepsilon_1^{a_{11}}}{\gamma_{11}} & \cdots & \dfrac{\xi_1^{a_{1k}}\varepsilon_k^{a_{1k}}}{\gamma_{1k}} \\ - & - & - \\ \dfrac{\xi_n^{a_{n1}}\varepsilon_1^{a_{n1}}}{\gamma_{n1}} & \cdots & \dfrac{\xi_n^{a_{nk}}\varepsilon_k^{a_{nk}}}{\gamma_{nk}} \end{matrix}, \tag{2.32}$$

i.e.

$$p\{((a_{\nu i}))\} = \frac{\xi_1^{a_{1.}}\cdots\xi_n^{a_{n.}}\cdot\varepsilon_1^{a_{.1}}\cdots\varepsilon_k^{a_{.k}}}{D} \tag{2.33}$$

where for short

(2.34) $\gamma_{\nu i} = 1 + \xi_\nu \varepsilon_i, \quad D = \prod_{\nu=1}^{n} \prod_{i=1}^{k} \gamma_{\nu i}$.

(2.33) shows that all matrices with two given sets of marginals $(a_{1\cdot}, \ldots a_{n\cdot})$ and $(a_{\cdot 1}, \ldots a_{\cdot k})$ have the same probability. Thus, in order to find the joint probability of the two sets we just have to count the number of (o,1)-matrices with the said marginals. Denoting this number by $\begin{bmatrix} a_{1\cdot} & \ldots & a_{n\cdot} \\ a_{\cdot 1} & \ldots & a_{\cdot k} \end{bmatrix}$ we have:

$$(2.35) \quad p\{a_{1\cdot} \ldots a_{n\cdot} ; a_{\cdot 1} \ldots a_{\cdot k}\}$$

$$= \begin{bmatrix} a_{1\cdot} & \ldots & a_{n\cdot} \\ a_{\cdot 1} & \ldots & a_{\cdot k} \end{bmatrix} \frac{\xi_1^{a_{1\cdot}} \ldots \xi_n^{a_{n\cdot}} \cdot \varepsilon_1^{a_{\cdot 1}} \ldots \varepsilon_k^{a_{\cdot k}}}{D}$$

From this result we may derive the probability of the column marginals by finding all $a_{\cdot i}$ that are compatible with the column marginals and add up the corresponding probabilities:

$$p\{a_{1\cdot} \ldots a_{n\cdot}\}$$

$$(2.36) \quad = \frac{\xi_1^{a_{1\cdot}} \ldots \xi_n^{a_{n\cdot}}}{D} \sum_{(a_{\cdot 1} \ldots a_{\cdot k})} \begin{bmatrix} a_{1\cdot} & \ldots a_{n\cdot} \\ a_{\cdot 1} & \ldots a_{\cdot k} \end{bmatrix} \varepsilon_1^{a_{\cdot 1}} \ldots \varepsilon_k^{a_{\cdot k}}$$

$$= \frac{\xi_1^{a_{1\cdot}} \ldots \xi_n^{a_{n\cdot}}}{D} \gamma(\varepsilon_1 \ldots \varepsilon_k | a_{1\cdot} \ldots a_{n\cdot}), \text{ say.}$$

Now divide (2.36) into (2.35) to obtain the conditional probability of the row marginals given the column marginals:

$$p\{a_{\cdot 1} \ldots a_{\cdot k} | a_{1\cdot} \ldots a_{n\cdot}\}$$

$$(2.37) \quad = \frac{p\{a_{1\cdot} \ldots a_{n\cdot} ; a_{\cdot 1} \ldots a_{\cdot k}\}}{p\{a_{1\cdot} \ldots a_{n\cdot}\}}$$

$$= \begin{bmatrix} a_{1\cdot} & \ldots a_{n\cdot} \\ a_{\cdot 1} & \ldots a_{\cdot k} \end{bmatrix} \frac{\varepsilon_1^{a_{\cdot 1}} \ldots \varepsilon_k^{a_{\cdot k}}}{\gamma(\varepsilon_1 \ldots \varepsilon_k | a_{1\cdot} \ldots a_{n\cdot})} .$$

This result is a generalization of (2.31), showing in fact how the column marginals for different r's may be pooled for estimation purposes. (2.37) is just the probability distribution of their total and may serve as a basis for an estimation of the ε's, utilizing all of the data. (2.37) being in this sense a synthesis of k-1 expresions of the form (2.31)(r = o and r = k are uninformative) preserves of course the specific objectivity of the estimation.

From (2.35) we might have derived the probability of the row marginals:

$$(2.38) \qquad p\{a_{\cdot 1},\ldots,a_{\cdot k}\} = \frac{\varepsilon_1^{a_{\cdot 1}}\cdots\varepsilon_k^{a_{\cdot k}}}{D}\cdot\gamma(\xi_1,\ldots,\xi_n|a_{\cdot 1},\ldots,a_{\cdot k})$$

and obtain a conditional probability symmetrical to (2.37)

$$(2.39) \qquad p\{a_{1\cdot},\ldots,a_{n\cdot}|a_{\cdot 1},\ldots,a_{\cdot k}\} = \begin{bmatrix} a_{1\cdot},\ldots,a_{n\cdot} \\ a_{\cdot 1},\ldots,a_{\cdot k}\end{bmatrix} \frac{\xi_1^{a_{1\cdot}}\cdots\xi_n^{a_{n\cdot}}}{\gamma(\xi_1,\ldots,\xi_n|a_{\cdot 1},\ldots,a_{\cdot k})}$$

which yields an estimation of the persons parameters that is also specifically objective, i.e. unaffected by the unknown parameters of the items.

Let us finally divide (2.35) into (2.34) to obtain the conditional probability of the whole set of answers, given the total number of correct answers for each person as well as for each item. Clearly all of the parameters cancel and in consequence the said conditional probability

$$(2.4o) \qquad p\{((a_{\nu i}))|(a_{\nu\cdot}),(a_{\cdot i})\} = 1\Big/\begin{bmatrix}(a_{\nu\cdot})\\(a_{\cdot i})\end{bmatrix}$$

becomes independent of all parameters. This result is particularly interesting. The structure of the model is specified by (2.1a) and the stochastic independence and it is the same whichever values the parameters take on. Thus an inference about the adequacy of the model as such showed, in order to deserve the qualification "specifically objective", be independent of all of the parameters. (2.4o) shows that this must in fact be possible even if it is not immediately clear how to do it.

If we had a large number of observed matrices with the same marginals $(a_{\nu\cdot})$ $(a_{\cdot i})$ the matter would be "easy", all realizations of the matrix having the same probability.

The principle of the problem would be the same as testing the equiprobability of the six eyes of a dice from, say, loo throws.

This of course is not feasible. On the face of it we should have just one observed matrix $((a_{\nu i}))$ and this could in no way be exceptional, since according to (2.4o) all (o,1)-matrices that are algebraically compatible with the given marginals are equally probable - however systematic or randomized they may look.

However, the derivation of (2.4o) does not presuppose that all data available were used. Just the same formula would hold for any part of the material.

Thus we may, according to any criterion we like, partition the n persons into a number of groups, $\mu = 1, \ldots m$:

(2.41)

$\mu.\nu$ \ i	1, ...	i, ...	k	total
1.1	$a_{11}^{(1)},\ldots,$	$a_{1i}^{(1)},\ldots,$	$a_{1k}^{(1)}$	$a_{1.}^{(1)}$
$1.n_1$	$a_{n_1 1}^{(1)},\ldots,$	$a_{n_1 i}^{(1)},\ldots,$	$a_{n_1 k}^{(1)}$	$a_{n1.}^{(1)}$
1.	$a_{.1}^{(1)},\ldots,$	$a_{.i}^{(1)},\ldots,$	$a_{.k}^{(1)}$	$a_{..}^{(1)}$
	-	-	-	
$\mu.1$	$a_{11}^{(\mu)},\ldots,$	$a_{1i}^{(\mu)},\ldots,$	$a_{1k}^{(\mu)}$	$a_{1.}^{(\mu)}$
$\mu.n_\mu$	$a_{n_\mu 1}^{(\mu)},\ldots,$	$a_{n_\mu i}^{(\mu)},\ldots,$	$a_{n_\mu k}^{(\mu)}$	$a_{n_\mu .}^{(\mu)}$
$\mu.$	$a_{.1}^{(\mu)},\ldots,$	$a_{.i}^{(\mu)},\ldots,$	$a_{.k}^{(\mu)}$	$a_{..}^{(\mu)}$
	-	-	-	
m.1	$a_{11}^{(m)},\ldots,$	$a_{1i}^{(m)},\ldots,$	$a_{1k}^{(m)}$	$a_{1.}^{(m)}$
$m.n_m$	$a_{n_m 1}^{(m)},\ldots,$	$a_{n_m i}^{(m)},\ldots,$	$a_{n_m k}^{(m)}$	$a_{n_m .}^{(m)}$
m.	$a_{.1}^{(m)},\ldots,$	$a_{.i}^{(m)},\ldots,$	$a_{.k}^{(m)}$	$a_{..}^{(m)}$
total	$a_{.1}^{(\cdot)},\ldots,$	$a_{.i}^{(\cdot)},\ldots,$	$a_{.k}^{(\cdot)}$	$a_{..}^{(\cdot)}$

For each group (2.4o) must hold:

$$(2.42)\quad p\{(a_{.i}^{(\mu)})|(a_{\nu .}^{(\mu)})\} = \begin{bmatrix}(a_{\nu .}^{(\mu)})\\(a_{.i}^{(\mu)})\end{bmatrix} \frac{\prod_{(i)} \varepsilon_i^{a_{.i}^{(\mu)}}}{\gamma(\varepsilon_1,\ldots,\varepsilon_k|(a_{\nu .}^{(\mu)}))}$$

Thus from each group we might estimate the ε's, but provided the model holds for all of the groups these estimates should not deviate significantly from each other. We shall, however, now show that it is possible to decide whether or no they do so, without actually carrying out all these estimations.

Due to the stochastic independence of the blocks the simultaneous distribution of the m sets of item totals, given the m sets of person totals is the product

$$(2.43)\quad p\{((a_{.i}^{(\mu)}))|((a_{\nu .}^{(\mu)}))\}$$

$$= \prod_{\mu=1}^{m} p\{(a_{.i}^{(\mu)})|(a_{\nu .}^{(\mu)})\}$$

$$= \prod_{\mu=1}^{m} \begin{bmatrix}a_{\nu .}^{(\mu)}\\a_{.i}^{(\mu)}\end{bmatrix} \cdot \frac{\varepsilon_1^{a_{.1}^{(.)}} \ldots \varepsilon_k^{a_{.k}^{(.)}}}{\prod_{\mu=1}^{m}\gamma(\varepsilon_1,\ldots \varepsilon_k|(a_{\nu .}^{(\mu)}))} .$$

However, the total column sums of the table (2.41) are the same as if the persons had not been grouped, and therefore we also have according to (2.37)

$$(2.44)\quad p\{(a_{.i}^{(\cdot)})|((a_{\nu .}^{(\mu)}))\} = \begin{bmatrix}((a_{\nu .}^{(\mu)}))\\(a_{.i}^{(\cdot)})\end{bmatrix} \cdot \frac{\varepsilon_1^{a_{.1}^{(\cdot)}} \ldots \varepsilon_k^{a_{.k}^{(\cdot)}}}{\gamma(\varepsilon_1,\ldots \varepsilon_k|((a_{\nu .}^{(\mu)})))} .$$

Obviously, the ε-products in (2.43) and (2.44) are equal, but the denominators are in fact also identical.

This is a consequence of a general formula, now to be proved, which expresses any γ-function in terms of the elementary functions $\gamma_r(\varepsilon_1, \ldots \varepsilon_k)$. In deriving (2.31) we applied (2.12) to each of n_r persons with the same $a_{\nu.}$, but the argument remains the same if they differ. Thus in analogy to (2.3o)

$$(2.45) \quad p\{(a_{\nu 1}, \ldots a_{\nu k}) | (a_{\nu .})\} = \frac{\varepsilon_1^{a_{11}} \ldots \varepsilon_k^{a_{1k}}}{\gamma_{a_{1.}}} \cdots \frac{\varepsilon_1^{a_{n1}} \ldots \varepsilon_k^{a_{nk}}}{\gamma_{a_{n.}}} = \frac{\varepsilon_1^{a_{.1}} \ldots \varepsilon_k^{a_{.k}}}{\prod_{(\nu)} \gamma_{a_{\nu .}}}$$

from which we get by summation over all matrices $((a_{\nu i}))$ with the same two sets of marginals

$$(2.46) \quad p\{a_{.1}, \ldots a_{.k} | (a_{\nu .})\} = \begin{bmatrix} (a_{\nu .}) \\ (a_{.i}) \end{bmatrix} \frac{\varepsilon_1^{a_{.1}} \ldots \varepsilon_k^{a_{.k}}}{\prod_{(\nu)} \gamma_{a_{\nu .}}},$$

and on comparing with (2.37):

$$(2.47) \quad \gamma(\varepsilon_1, \ldots \varepsilon_k | (a_{\nu .})) = \prod_{\nu=1}^{n} \gamma_{a_{\nu .}} (\varepsilon_1, \ldots \varepsilon_k).$$

From this formula it follows that each term in the denominator of (2.43) may be expressed as the product of the elementary γ_r-functions corresponding to r-values of $a_{\nu .}^{(\mu)}$, $\nu = 1, \ldots n_\mu$. Compiling them for $\mu = 1, \ldots m$ we get just the terms making up the product representation of the denominator of (2.44). Therefore, as stated above,

$$(2.48) \quad \gamma(\varepsilon_1, \ldots \varepsilon_k | ((a_{\nu .}^{(\mu)}))) = \prod_{\mu=1}^{m} \gamma(\varepsilon_1, \ldots \varepsilon_k | (a_{\nu .}^{(\mu)})) .$$

We may now draw the conclusion aimed at, that the probability of m sets of item marginals, given the m sets of personal marginals conditional upon the set of total marginals for the items, as obtained by dividing (2.44) into (2.43) is independent of the ε's:

$$(2.49)\qquad p\{((a_{.i}^{(\mu)}))\,|\,((a_{\nu.}^{(\mu)})),(a_{.i}^{(\cdot)})\} = \frac{\prod_{\mu=1}^{m}\begin{bmatrix}(a_{\nu.}^{(\mu)})\\(a_{.i}^{(\mu)})\end{bmatrix}}{\begin{bmatrix}((a_{\nu.}^{(\mu)}))\\(a_{.i}^{(\cdot)})\end{bmatrix}}\ .$$

This formula is a generalization of (2.4o) to which it **reduces** when each group consists of only one person.

From another point of view (2.49) may be considered as an elaboration of (2.4o)[x], producing consequences that are useful for testing purposes.

---

x) It may in fact be derived directly by repeated applications of (2.4o) without returning to the ε's:

$$p\{((a_{.i}^{(\mu)}))\,|\,((a_{\nu.}^{(\mu)})),(a_{.i}^{(\cdot)})\} = \frac{p\{((a_{.i}^{(\mu)})),(a_{.i}^{(\cdot)})\,|\,((a_{\nu.}^{(\mu)}))\}}{p\{(a_{.i}^{(\cdot)})\,|\,((a_{\nu.}^{(\mu)}))\}};$$

since $(a_{.i}^{(\cdot)})$ is determined algebraically from $((a_{.i}^{(\mu)}))$

$$p\{((a_{.i}^{(\mu)})),(a_{.i}^{(\cdot)})\,|\,((a_{\nu.}^{(\mu)}))\} = p\{((a_{.i}^{(\mu)}))\,|\,((a_{\nu.}^{(\mu)}))\}.$$

On multiplying numerator and denominator by $p\{(((a_{\nu i}^{(\mu)})))\,|\,((a_{\nu.}^{(\mu)}))\}$ we then get

$$p\{((a_{.i}^{(\mu)}))\,|\,((a_{\nu.}^{(\mu)})),(a_{.i}^{(\cdot)})\}$$

$$= \frac{p\{(((a_{\nu i}^{(\mu)})))\,|\,((a_{\nu.}^{(\mu)}))\}\ :\ p\{(a_{.i}^{(\cdot)})\,|\,((a_{\nu.}^{(\mu)}))\}}{p\{(((a_{\nu i}^{(\mu)})))\,|\,((a_{\nu.}^{(\mu)}))\}\ :\ p\{((a_{.i}^{(\mu)}))\,|\,((a_{\nu.}^{(\mu)}))\}}$$

$$= \frac{p\{(((a_{\nu i}^{(\mu)})))\,|\,((a_{\nu.}^{(\mu)})),(a_{.i}^{(\cdot)})\}}{\prod_{(\mu)} p\{((a_{\nu i}^{(\mu)}))\,|\,(a_{\nu.}^{(\mu)}),(a_{.i}^{(\mu)})\}}$$

which on applying (2.4o) to each of the groups as well as to the total material yields (2.49).

Another elaboration of (2.4o) obtains when the items are partitioned into a number of groups, $h = 1,\ldots,l$. Formally the result is perfectly symmetrical to (2.49):

$$(2.5o) \quad p\{((a_{\nu.}^{(h)}))|((a_{.i}^{(h)})),(a_{\nu.}^{(\cdot)})\} = \frac{\prod_{h=1}^{\ell}\begin{bmatrix}((a_{\nu.}^{(h)}))\\((a_{.i}^{(h)}))\end{bmatrix}}{\begin{bmatrix}((a_{\nu.}^{(\cdot)}))\\(a_{\nu.}^{(\cdot)})\end{bmatrix}}.$$

In practice there is the difference that usually persons are numerous and items relatively few.

Having chosen the groupings the control on the model has now been reduced to a matter of ordinary statistical testing technique based upon a study of the bracket symbols. At which point we shall leave it here.

But how to choose the grouping . Our formulae allow for any grouping desired. This state of affairs leaves a great deal of freedom to the statistician with the risk of the model-testing being at the mercy of his personal preferences.

This is a matter that deserves more attention than we are able to give it here. But I may tell something about my own leading star.

My point of departure I take in the statement that models are never true and they are not meant to be so. This point may be illustrated by the case of the pendulum.

The simplest model in this case is the "mathematical pendulum": a heavy point fixed to a weightless string and swinging frictionless in vacuum.