Discussie-rubriek

STOCHASTIC UNFOLDING: TWO NULL MODELS AND THEIR FLAWS.

Ivo W. Molenaar, Statistiek en Meettheorie, FSW, R.U. Groningen.

In the June 1979 issue of MDN, a paper by Van Schuur and Stokman on stochastic unfolding of rank order data was followed by a note of Van der Eijk and Van der Noort, stating that the notion of "expected frequency of errors under random preference ordering" was incorrectly defined in the former paper, and proposing an alternative definition "not marred by these flaws".

The present note, however, detects some flaws in the Van der Eijk & Van der Noort proposal, and leads the author to maintain his preference for the Van Schuur-Stokman definition.

1. A bogus flaw and a real one

Let three stimuli A, B and C form a J-scale in this order. If Π_{AB} denotes the probability that A is preferred over B by a randomly chosen subject, then

$$\Pi_{AB} < \Pi_{AC} < \Pi_{BC} \ . \quad ^{*)}$$

It is desired to calculate the expected frequency of the two inadmissible rank orders ACB and CAB "under the null hypothesis". In analogy to Mokken scaling, this null model is taken as statistical independence of the answers for various items (here: stimulus pairs), taking into account the observed relative popularity of the items. The extra problem encountered here is that this simple null model would lead to $2^3 = 8$ possible patterns for three stimuli. If it was the task of the subject to produce an ordering of the stimuli, however, the two circular patterns A < B < C < A and A > B > C > A have zero frequency. Van Schuur and Stokman brush this problem aside by looking at the conditional distribution obtained by eliminating the intransitive patterns and inflating the remaining six probabilities by division

*) Note that $\Pi_{AB} = 1 - \Pi_k$ in the notation of the previous paper.

through 1 minus the probability of an intransitive response.

At this point their critics (MDN 4(1979) nr. 2, page 32) state that "the calculated 'probability' of intransitive answers is not a probability at all, as it does not refer to any imagineable event (simple or composite) in a sample space". The present author asks the critics:

(a) to imagine an order-obtaining strategy or sample space obtained by three independent decisions on AB, AC and BC with probabilities Π_{AB}, Π_{AC} and Π_{BC} respectively, followed by an independent repetition of the decisions if and only if one of the two intransitive patterns appears;

(b) to consider that Andrichs (1978 , 1979) has published on exactly this model, independently derived, in recent issues of Psychometrika and Biometrics, two journals that might not have admitted the work if it really did not properly define probabilities;

(c) to realize that the model which they reject here is a simple case of structural zeros in two of the 2^3 cells with quasi-independence governing the distribution across the other six, see e.g. Bishop, Fienberg & Holland (1975) Ch.5.

A second remark by Van der Eijk and Van der Noort tells no more than that they "do not believe" that the redistribution of the probability of an intransitive response across the six transitive ones according to this initial size is justifiable or plausible. This is left as a matter of taste on which it is difficult to argue.

A third remark, entirely correct this time, points out that the redistribution leads to a marginal probability Pr(ACB v CAB v CBA) which is not equal to the marginal probability $1-\Pi_{BC}$ from which the independence model had started (with a trivial exception when $\Pi_{AB} = \Pi_{AC} = \Pi_{BC} = .5$). This is a real flaw; it means that the estimation of the original probabilities Π_{AB}, Π_{AC} and Π_{BC} from the data is biased. A correction can be obtained from formula (6) of Van der Eijk and Van der Noort. It means that three unknown probabilities $x = \Pi_{BC}$, $y = \Pi_{AC}$ and $z = \Pi_{AB}$ have to be found such that e.g. the observed relative frequency of the preference for

B above C is not put equal to x, but to the expected probability Pr(ABC v BAC v BCA) as calculated from x, y and z. Let us put

$$t = 1 - zx(1 - y) - y(1 - x)(1 - z) = 1 - y + xy + yz + xz$$

for the probability of a transitive response in the simple independence model with 2^3 categories, and let a, b, and c be the observed relative frequencies of the preferences BC, AC and AB respectively. Then we must solve x, y and z from

$$ta = xyz + x(1 - y)(1 - z)+(1 - z)xy = x - xz + xyz;$$
$$tb = xyz + yz(1 - x) + xy\ (1 - z) = xy + yz - xyz;$$
$$tc = xyz + yz\ (1 - x) + (1 - x)(1 - y)z = z - xz + xyz.$$

For these four equations in the four unknowns t, x, y, z no explicit solution seems to be available, but Charles Lewis wrote a little computer program like leading to an iterative solution. It takes something like 10 iterations to get a solution accurate in six decimal places.

This solution, however, is less satisfactory than I had hoped. It follows from ta - tc = x - z (subtract the two equations) that the difference between x and z is the difference of a and c shrunken by a factor t. But y need not be close to b, and it can be shown that for equidistant observed marginals (a - b = b - c) the solution is always y = ½, even when for example a = .3, b = .2 and c = .1.

When the corrected values for t, x, y, z are applied to the triplets listed on page 24 of MDN 4 (2), the expected number of errors under the null model always decreases and the H coefficient becomes even worse. This can be viewed as a confirmation that such triplets of political parties do not form a J-scale for the respondents. It could also mean, as more or less suggested by Van der Eijk and Van der Noort, that our null model is inappropriate. Putting the blame on the null model, however, lightly dismisses the fact that many respondents chose one of the two triplets ACB and CAB which are incompatible with the J-scale ABC.

2. The alternative model

The alternative model proposed by Van der Eijk and Van der Noort tries to obtain the expected frequencies of the k! possible orderings of k stimuli from the side conditions that the marginal frequencies of each of the $\binom{k}{2}$ stimulus pairs and the total frequency are fitted. As $\binom{k}{2} + 1$ is less than k! for all $k \geq 3$, there will be more than one solution. As a next step the authors postulate

(a) only integer values for the expected frequencies are allowed;

(b) each solution with such integer values is equally probable.

The present author is not at all convinced that such postulates are of any use in an attempt to asses how many errors (non-scale orderings) would be obtained under the null hypothesis against which the proposed J-scale is compared. It suffices to think of the weird consequences of such postulates for the chisquare test for independence in the 2 x 2 table. For the observed table at the left of Table 1, the expected frequencies would no longer be calculated as in the second part; the new postulates would tell us that each of the right-hand tables for $j = 0, 1, 2, \ldots, 10$ would be equally probable, thus the average expected frequency, corresponding to $j = 5$, would describe the null hypothesis.

Table 1. An observed 2 x 2 table, its expected frequencies under independence, and the expected frequencies obeying postulates (a) and (b)

1	14	15
9	16	25
10	30	40

3.75	11.25	15
7.25	18.75	25
10	30	40

j	15-j	15
10-j	15+j	25
10	30	40

$j=0,1,2,\ldots,10$

5	10	15
5	20	25
10	30	40

$j=5$

There is no sound reason for expected frequencies to be integer, and even less reason why all possible tables obeying the marginal restrictions would be equally probable. This equiprobability clearly violates the independence assumption which was designed as a base line at the very start of the investigation: if the observed frequencies in the error cells could very well have been produced in an independence model, they are too high for a serious corroboration of our J-scale model. Beating some average of all possible integer-valued models does not provide the corroboration that we seek. Indeed some of these integer-valued models like the $j = 0$ case in Table 1, or like the sixth column of f'_{ij} on page 37 of MDN 4 nr.2, may be much more in favor of our research hypothesis than in favor of independence.

We note in passing that the numerical example containing ten answer patterns, of which one violates the J-scale, indeed leads to $H_{ABC} = -.43$ using the Van Schuur-Stokman null model. Disappointing as this may be, it should be noted that the marginal frequencies of 2, 5 and 7 used in the example imply that the total number of violations must be 0, 1, or 2. Observing 1 then, in this case turns out to be even less than would be expected in the null model: with such a small number of cases the occurrence of one violation is certainly not enough to reject the possibility of independent judgements on item pairs.

3. Discussion

The process of independent comparisons, followed by proportional redistribution of inadmissible patterns across all admissible ones, has been forwarded as an attempt to reconcile the independence and transitivity requirements. Such a process, studied in a similar context in recent papers by Andrich, is just one of the possible ways of solving the conflicting demands. Our attempt to use better estimates of the marginals for the original null model with 8 cells has not been completely convincing. Better proposals would be welcomed, but Van der Eijk and Van der Noort have not convinced us that dropping independence and assuming equiprobability between all integer-valued models leads to a null model which makes sense.

Acknowledgement

Frans Stokman and Charles Lewis were kind enough to provide helpful comments.

References

Andrich, D., A rating formulation for ordered response categories, Psychometrika, Vol. 43, no.4, December 1978, 561-573.

Andrich, D., A model for contingency tables having an ordered response classification, Biometrics, Vol. 35, no. 2, June 1979, 403-415.

Bishop, Y.M.M., Fienberg, S.E., and Holland, P.W., Discrete multivariate analysis, The MIT Press Cambridge, Massachusetts, and London, England, 1975.

Fienberg, S.E. & Holland, P.W., Methods for eliminating zero counts in contingency tables, Random Counts in Scientific Work, vol. 3, p.233-260, G.P. Patil (editor), Pennsylvania State University Press.

Van der Eijk, C. & Van der Noort, W., Some notes concerning stochastic unfolding and the expected frequency of rankorder-patterns, Methoden en Data Nieuwsbrief van de sociaal wetenschappelijke sectie van de VVS, jaargang 4, nr. 2, juni 1979, p.30-38.

Van Schuur, W.H. & Stokman, F.N., A one-dimensional stochastic unfolding model with an application to party preferences in The Netherlands, Methoden en Data Nieuwsbrief van de sociaal wetenschappelijke sectie van de VVS, jaargnag 4, nr. 2, juni 1979, p.3-29.