inter·noise 96
25th Anniversary Congress - Liverpool

INCE: 69.3

SOME WAYS TO PROCESS RAW DATA COMING FROM COMMUNITY NOISE SURVEYS

M Maurin

INRETS-LEN, case 24, 69675 Bron cedex, France

1. INTRODUCTION

Complete noise annoyance surveys are involving annoyance questionnaires and related noise recordings in the immediate vicinity of respondents. Thus they yield two complementary sorts of data, and we need both in order to investigate a possible connection between noise nuisances as stimuli and individual responses.
Of course this expected relationship may not be anything else but collective and statistical ; and in any case we need to take some elementary precautions in relation to recording, encoding and processing data. Last in order to gather more numerous data coming from different surveys some new considerations for homogeneity or direct comparability are raising.

2. SOME METHODOLOGY

We intend to deal with the true basic or raw data as they naturally do are, but which are not frequently published under an elemantary form ; this means the natural or integer frequencies instead of conditional percentages, and the raw contengency tables (2 or multi-way, [1, 2, 4]).

i) Here we speak of some segmentation processings applied to a 2-way table with ordered noise levels in variables, and modalities of some other variable in cases, n_{ij} is the number of people exposed to level L_i and answering modality C_j. Besides the classical test of independence, for every column i we consider the two-columns table with grouped together data $n'_{1j} = \sum_{k=1,i} n_{kj}$, $n'_{2j} = \sum_{k=i+1,I} n_{kj}$, and the classical Pearson' statistic

$$\chi^2_{J-1}(i) = n.. \sum_j \{ \frac{n'^2_{1j}}{n'_{1.} \, n_{.j}} + \frac{n'^2_{2j}}{n'_{2.} \, n_{.j}} \} - n.. \quad \text{with } n'.. = n.. \, , \; n'_{.j} = n_{.j} \, .$$

The level i_0 for which the statistic $\chi_{J-1}^2(i_0)$ is maximal is the best value for which responses above and under are the most different ; this is a dichotomy of the population (segmentation into 2 parts) and it may be considered as a threshold value (in dBA) dividing the population in relation to the modalities C_j , fig. 1, [2].

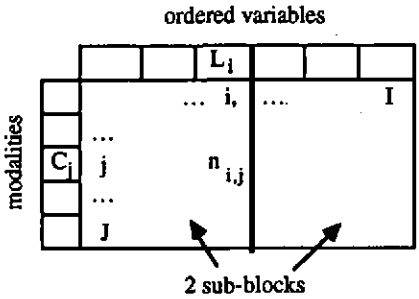

fig. 1 , segmentation of a 2-way table into 2 sub-blocks.

ii) If necessary the same may be done with a tritomy, grouped data $n^*_{1j} = \sum_{k=1,i1} n_{kj}$, $n^*_{2j} = \sum_{k=i1+1,i2} n_{kj}$, $n^*_{3j} = \sum_{k=i2+1,I} n_{kj}$, and the maximum of

$$\chi_{2(J-1)}^2(i_1,i_2) = n.. \sum_j \left\{ \frac{n^{*2}_{1j}}{n^*_{1.}\, n._j} + \frac{n^{*2}_{2j}}{n^*_{2.}\, n._j} + \frac{n^{*2}_{3j}}{n^*_{3.}\, n._j} \right\} - n..$$

iii) More or less analogous multi-way processings enable us to exam some statistical effects due to a third variable (fig. 2), for instance are responses versus C and L dependent of T ? [3].

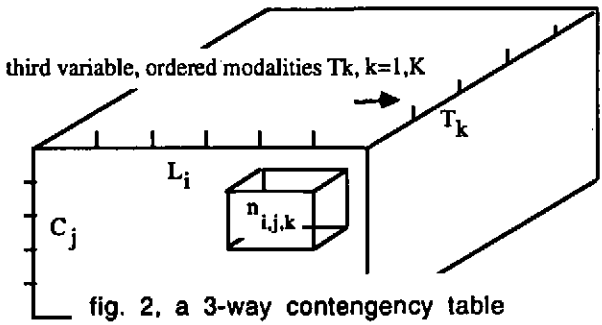

fig. 2, a 3-way contengency table

3. AN EXAMPLE

i) The data are coming from a recent french survey first oriented to night annoyance and extended to day and evening [6]. 1000 people splitted into 18 sites responded, and many noise indices have been measured in a two-steps way, i) first a long time noise measurement (24 or better 36h)

in a central-reference point (Leq for each second), ii) the run of a noise prediction software in each site. The i) step provides every Leq index for every desired time period (true levels), the ii) one the modelised propagation terms between reference point and every desired location, as faces for respondents' dwellings (sometimes 2 or 3), [7]. After some crude refining and qualiity tests we work on about 890 people.

Thus we got a two-way table with roughly 1000 individuals as cases and more than 200 variables (questions and noise levels). Here we deal with three Leq noise indices, a night index on [0 -5 am] roughly corresponding the most silent part of night (unfortunately eluded in [22 pm - 6 am] index), a day index [6 am - 22 pm] and an evening one [18 pm - 22 pm] , and only few questions, night, day and evening annoyance on a four semantic points scale, the presence of some noise protections as collective barriers or individual reinforced windows, the nature of space for each face {hight street, quiet or no street, mixed}.

ii) For instance the table crossing night annoyance level (on the 4 points scale) and L_{Aeq} index on night time [0 - 5], as summarised in table 1 but truly built dB by dB for the processing. The dichotomeous technic yields the "best" threshold here equal to 55 dBA ; if we do need a tritomy the thresholds are 54 and 63 dBA.

annoy.	≤ 43 dBA	]43 46]	]46 49]	]49 52]	]52 55]	]55 58]	]58 61]	]61 64]	> 64
1	28	48	54	61	60	59	45	20	13
2	20	31	29	23	46	45	37	13	21
3	12	15	16	7	16	30	23	14	17
4	2	7	9	7	12	16	14	9	13

Table 1 : $L_{Aeq[0-5]}$ and night annoyance 1 : not to 4 : very.

* Split without specific individual noise protections we get

annoy.	≤ 43 dBA	]43 46]	]46 49]	]49 52]	]52 55]	]55 58]	]58 61]	]61 64]	> 64
1	14	26	21	24	28	21	21	13	7
2	8	15	9	9	15	10	18	4	13
3	7	7	11	4	5	10	11	9	5
4	1	3	2	2	3	10	5	7	4

* and with reinforced windows :

annoy.	≤ 43 dBA	]43 46]	]46 49]	]49 52]	]52 55]	]55 58]	]58 61]	]61 64]	> 64
1	14	22	33	37	32	38	24	7	6
2	12	16	20	14	31	35	19	9	8
3	5	8	5	3	11	20	12	5	12
4	1	4	7	5	9	6	9	2	9

Tables 2 : table 1 split versus individual protections

The difference or statistical effect between the two sub-tables is significative, without reinforced windows the threshold is equal to 55 dBA but with : 63 dBA.

iii) There are some analogous results for day levels of noise (Leq) and annoyance, the tritomy with 58 and 64 dBA is the most striking, (and a dichotomy at 60 for evening levels), as so as split versus the collective noise barriers, and the global exposure of dwelling' s faces. For instance :

night [0 - 5] (threshold values in dBA)

windows	normal	reinforced	*comment*
threshold :	55	63	*OK, suitable*
barriers	without	with	
threshold :	56	62	*some extra psychol. effect ?* [5]
exposure	mixed	all windows on hight streets	
threshold :	54	56	*some passive resignation ?* [5]

day [6 - 22]

windows	normal	reinforced	*comment*
threshold :	58	61	*OK, suitable*
barriers	without	with	
threshold :	58	61	*some extra psychol. effect ?* [5]
exposure	mixed	all windows on hight streets	
threshold :	58	59	*some passive resignation ?* [5]

iv) Of course many other results are available with different time periods and variables, here with segmentations technics (dicho- and possibly tritomy) ; but also with other processing frameworks as annoyance measurement in psycho-physics once we get raw contengency tables, [1].
What may import is the possibility to gather rather important lot of data because algorithms and results are sensitive to null or weak frequencies ; thus this may be an argument in order to render more comparable and "re-unifiable" sets of records coming from different noise surveys.

References

[1] Maurin M., The measurement of noise annoyance responses, Inter-Noise, Munich 85, (1985).
[2] Maurin M., Le traitement des échelles de catégories dans les études sur les nuisances, INRETS n° 37, (1987).
[3] Mood A.M., Graybill F.A., An introduction the theory of statistics, (Mc. Graw Hill, 1963).
[4] Plackett R.L., The analysis of categorical data, (Ch. Griffin, 1974).
[5] Vallet M., first informal discussions and Euronoise 96, (1996).
[6] Vallet M. and col., Recherche d'un indice acoustique de gêne nocturne liée au trafic routier, INRETS, (1995).
[7] Vernet I., Maurin M., Koch JR., Vallet J., Prévision de niveaux sonores nocturnes, études de cas, Euro-Noise'95, 179-184, Lyon, (1995).