Partitioning ordered variables into discrete states for discriminant analysis of ecological classifications

LOUIS LEGENDRE
Département de biologie, Université Laval, Québec (Qué.), Canada G1K 7P4
AND
PIERRE LEGENDRE
Département de sciences biologiques, Université de Montréal, C.P. 6128, Succursale A, Montréal (Qué.), Canada H3C 3J7

Received October 28, 1982

LEGENDRE, L., and P. LEGENDRE. 1983. Partitioning ordered variables into discrete states for discriminant analysis of ecological classifications. Can. J. Zool. **61**: 1002–1010.

This paper describes a procedure to partition ordered variables into discrete states for the discrimination of an ecological classification. At each step, the best partition is that which maximizes a log likelihood ratio for nonhomogeneity of distributions across the groups of the classification. The partitioning procedure ends when the probability of the log likelihood χ^2 statistic reaches its minimum value. An actual ecological example is given of the discrete discriminant analysis of a benthos classification, with vegetation and oxygen concentration as discriminant variables. The O_2 observations are first partitioned into discrete states, using the partitioning algorithm described before. Following three-dimensional contingency table analysis, it is concluded that the benthos classification is independent of oxygen concentrations.

LEGENDRE, L., et P. LEGENDRE. 1983. Partitioning ordered variables into discrete states for discriminant analysis of ecological classifications. Can. J. Zool. **61**: 1002–1010.

Cet article décrit une méthode qui permet de partitionner des variables ordonnées en classes, en vue de discriminer entre les groupes d'une classification écologique. A chaque étape, la meilleure partition est celle qui maximise un rapport de vraisemblance pour la non homogénéité des distributions parmi les groupes de la classification. La partition se termine lorsque la probabilité du χ^2 atteint sa valeur minimale. A titre d'exemple écologique, une classification de benthos animal est soumise à l'analyse discriminante discrète, en utilisant la végétation et la concentration d'oxygène comme variables discriminantes. Les mesures de l'oxygène sont tout d'abord partitionnées en classes discrètes, au moyen de l'algorithme décrit précédemment. Au terme de l'analyse du tableau de contingence à trois dimensions, on conclut que la classification benthique est indépendante des concentrations d'oxygène.

Introduction

Qualitative (nominal) data are often disregarded by ecologists because they cannot be treated directly together with the more usual quantitative variables, even though many essential variables are qualitative, and many quantitative variables are much more efficiently sampled as rank ordered (ordinal). Among the qualitative data are the dominant species, the biological association, the presence or absence of a species or of some substance (e.g., pollutant), the values exceeding or lower than a physiological threshold, the type of substrate, etc. Cyclical variables such as the direction of the wind or of a current, coded into nonordered states, are other examples of qualitative data encountered in ecology. Quantitative information, on the other hand, may often be sampled efficiently as importance or abundance scores; samples can also be rapidly enumerated in the laboratory as coded abundance scores (Frontier 1969, 1973), the cost of data collection being reduced, or the number of samples increased for a given cost.

This paper describes a procedure whereby quantitative or rank-ordered variables can be partitioned into discrete states, in order to use them together with qualitative discrete variables in problems of ecological discrimination. The importance of this problem has been emphasized by various authors, including Fienberg (1970). The rationale of the proposed method is the following.

Ecological data analysis often requires the assessment of the relationships between two data sets, describing in different ways the same objects of study (these objects being the samples, the quadrats, the stations, and so forth). For example, the ecologist may wish to explain the distribution of species in a territory as a function of observed environmental characteristics, or to compare two groups of taxa within a region, or even to establish the similarity between two sets of environmental variables. In theory, the best approach to such a problem would be that of canonical correlation analysis, in which maximum correlations between linear combinations of the two sets of variables (canonical variates) are computed. Unfortunately this symmetric approach requires the relationships between variables to be strictly linear (Gauch and Wentworth 1976). This is not often the case in ecology.

The asymmetric approach, more adapted to ecological data (Cassie 1972), consists in first establishing the structure (clustering or ordination) within one of the two data sets, and then analyzing this structure using the data of the second set. The structure of the first set is often known as a classification, that is, a partition of the set of objects (or of variables) into a series of mutually exclusive groups. These groups may be observed directly by the ecologist or they may be predetermined by the system under examination. Alternatively, the structure may be defined through multivariate numerical analysis (cluster analyses, ordination techniques, combination of the two by superimposing a clustering on the scatter diagram of the objects in the ordination space), as discussed for instance by Legendre and Legendre (1983). Given the classification, the ecological approach is then to interpret the structure using potentially explanatory data from a second set. To be of any ecological significance, the analysis has to provide quantitative information on how the variables of the second set discriminate between the groups established on the basis of the variables of the first set. The statistical analysis designed to discriminate between groups using a set of new variables is called discriminant analysis and it is routinely used by a growing number of ecologists. Two types of analysis can be used for this discrimination. These are parametric and discrete discriminant analysis. The parametric analysis is restricted to the case where all the discriminant variables are quantitative, (multi)normally distributed and interact linearly so as to discriminate between the groups making the ecological classification. On the other hand, discrete discriminant analysis (Goldstein and Dillon 1978) may be used when the discriminant variables are qualitative (discrete states). When the discriminant variables are of mixed types (some qualitative, some rank-ordered and (or) some quantitative), discrete discriminant analysis may be used, but only after partitioning the ordered variables (quantitative and rank-ordered) into discrete states in order to use them as the qualitative variables. An alternative method would be the ALSOS program (alternating least squares and optimal scaling) mentioned by Young (1981), for discriminant analysis with mixed-type data.

It must be pointed out that a multidimensional discrete discriminant analysis can never be replaced by the analysis of a set of two-way contingency tables comparing each of the potentially explaining variables in turn to the classification to be explained. Indeed, such an approach forgets all interactions higher than the first degree and prevents testing more complex causal models. The necessity of testing higher-order interaction models will become obvious in the very simple ecological example given in the last section of this paper.

In partitioning ordered variables, two decisions must be made: (*a*) the number of states that are required, and (*b*) where to place divisions between states. There are at least two ways of making these decisions.

First, there may be biological, geological, chemical, or other ecological reasons to place the partitions at known or hypothesized threshold values; ecologically located divisions are to be preferred to mathematically determined ones. A second way consists in partitioning the variable into states of equal width, or alternatively, into states containing the same number of objects. This solves only the *b* and not the *a* problem. Another variant of this solution is the one proposed by Cox (1957), who addressed himself to the problem of finding the optimal partitioning of a normally distributed variable into a predetermined number of states. He solved problem *b* by minimizing a loss function based on variance computations, and he found in this way where the normal distribution would be partitioned to form two to six states.

In this paper, a procedure is presented and criteria are proposed to allow partitioning of an ordered variable into states, such that maximum discrimination of a classification scheme is obtained. The underlying idea is that the final interpretation can only be improved by partitioning variables optimally with respect to maximum discrimination. The procedure answers both problems *a* and *b* above, and can be translated into a computer program.[1]

A real example is also presented, in which an ecological classification is interpreted by a mixture of qualitative and quantitative variables. In much the same way as in this example, Vincent and Bergeron (1983) used our algorithm to partition physical and chemical descriptors of water masses to optimize their power to explain a previous classification of the stations into aquatic vegetation types.

This partitioning algorithm has also been used for purposes quite different from discriminant analysis. André *et al.* (1981) have used it in a comparative study of benthos samplers; they divided species abundance variables into classes optimally related to the "sampling device" descriptor, in order to measure the selectivity of the sampling gears. Hudon *et al.* (1983) have used it to order a mixed-type group of variables as to their power to predict a predetermined classification. The quantitative variables were divided into classes, so that all descriptors could be compared to the reference classification (four substrate types, in that case), using the same coefficient of dependence. On the other hand, J. Ferraris (to be published) used our partitioning method to compare a metric distance matrix, computed between a group of objects, to the ultrametric matrix resulting from

[1]A computer program written in PASCAL is available from the second author, for CDC machines of the series 6000.

the application of a clustering procedure to these same objects. In this case, the classes of the qualitative reference variable correspond to the fusion levels of the clustering dendrogram.

Method

A simple numerical example will be used. A fictitious ordered discriminant variable (D) is to be partitioned into states, in order to maximize discrimination of a 4-group classification (C). The number of observations is $N = 16$. The fictitious data (in contingency table form) are the following:

Groups of C	Successive values of ordered discriminant variable D													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1			1									1	1
2		1				1						1		1
3			1		1				1		1			
4							1	1		2				
Partitions	1	2	3	4	5	6	7	8	9	10	11	12	13	

Variable D is ordered, which means that the successive columns of the table correspond to increasingly higher (or lower) values of D; tied observations are possible (as in group 4) and they prevent any partition to allocate them to different states. In such a table, quantitative variables are automatically reduced to ranks; tied observations are accepted and values of the quantitative variable for which there are no observations are noted as columns of zeros: these have no effect on the following computations and may therefore be eliminated. For instance, each column of the data table could be the observed temperature, rounded to the nearest integer. In the table above, two observations corresponding to the 10th temperature interval fall into the fourth group of the classification. In the 14th temperature interval, one observation falls into the first group of the classification and another into the second.

The variable (D) is then partitioned into two states. In this case, there are 13 such possible partitions, as indicated under the table. For instance, partition 9 (occurring between values 9 and 10 of variable D) would result in the following contingency table:

Groups of C	States of D		Total
	1	2	
1	2	2	4
2	2	2	4
3	3	1	4
4	2	2	4
Total	9	7	16

In each cell of this contingency table is the number of observations, belonging to a given group of the classification, that are allocated to a state of D by the partitioning algorithm. Each number in the first state of D is the sum of those to the left of partition 9 in the data table above, whereas each number in the second state of D is the sum of those to the right of partition 9.

The dependence between the classification (C) and the discriminant variable (D) is measured through Wilks' (1935) likelihood ratio statistic, which is asymptotically distributed as χ^2 when the total number of observations (N) is large:

$$\chi^2 \simeq 2 \sum_{\substack{\text{all the} \\ \text{cells}}} O \ln (O/E) \qquad \text{when } O, E > 0$$

where O is the number of observations in each cell of the contingency table, and E is the corresponding expected value; the statistic is computed using natural logarithms (ln). Under the null hypothesis of independence between C and D, E values are computed as:

$$(\text{total of the row} \times \text{total of the column})/N$$

The same results are achieved if the expected cell frequencies are written as an additive log-linear model:

$$\ln E = [\theta] + [C] + [D]$$

the null hypothesis being then $H_0: [CD] = 0$. Explanations about log-linear models are given in the next section.

When the distribution of D is the same for all the groups of the classification (same relative frequencies of D in each row of the contingency table), and conversely the distribution of C is the same for all the states of D, then C and D are completely independent. In this case, looking at the states of one variable does not give any information about the distribution among the states of the other. The corresponding log likelihood ratio is zero. This statistic thus measures the nonhomogeneity of the distribution of D across the groups of C, the associated probability being that of independence between C and D. For partition 9 above, the log likelihood ratio is 0.796, with a corresponding probability 0.850 which indicates that the hypothesis of independence between C and D cannot be rejected at the 5% probability level.

The best partition into a given number of states is that with the highest log likelihood ratio, since it has the maximum "weight of evidence" (Good 1950) for nonhomogeneity of distributions across the groups, and thus the lowest probability of independence between C and D. The highest computed log likelihood ratio for two states is that resulting from partition 11 (Table 1).

Similarly, all the possible three-states partitions may be assessed, in order to determine the one resulting in the highest log likelihood ratio. The procedure is the same as for the two-states partitioning, except that partition 11 of the fictitious example is now kept fixed. For instance,

TABLE 1. Stepwise partitioning of fictitious discriminant variable D, given classification C (see the text). Position of successive partitions are shown, as well as maximum log likelihood ratio (χ^2 statistic), number of degrees of freedom (ν), and probability (P) of independence between C and D

No. of states	Partitions	χ^2	ν	P
2	11	6.904	3	0.0750
3	11 and 6	17.995	6	0.0063
4	11 and 6 and (1 or 5 or 10)	20.629	9	0.0144

partition 9, keeping partition 11 fixed, would result in the following three-states discriminant variable:

Groups of C	States of D 1	2	3	Total
1	2	0	2	4
2	2	0	2	4
3	3	1	0	4
4	2	2	0	4
Total	9	3	4	16

The log likelihood ratio resulting from this partition is 10.357. The maximum is that of partition 6 (Table 1), so that the best partition into a three-states variable is at 11 and 6.

For a four-states variable D, keeping partitions 11 and 6 fixed, the best partition would be 1 or 5 or 10, which result in the same maximum log likelihood ratio (Table 1).

The partitioning procedure ends when the probability of the log likelihood χ^2 is minimum, since the dependence of the classification on the discriminant descriptor is then maximum. The best partition of discriminant variable D is therefore into three states, as the probability reaches a minimum (Table 1).

This stepwise procedure (keeping the previous $s - 1$ partitions fixed when partitioning into s states) does not always give exactly the same results as the simultaneous partitioning of the variable into s states: sometimes, it is possible to find a simultaneous partitioning with a slightly higher χ^2 than with the stepwise procedure. In the simultaneous procedure, for every successive partitioning into 2, 3, ..., s, ... states, all the possible $(N-1)!/(N-1-s)$ s-states partitions are tried in turn. The latter procedure must be preferred when computer time is not limited. However, in view of the rapid increase in computing time with the simultaneous algorithm, especially when N is large, the stepwise procedure is described here as a practical alternative.

In the fictitious example above, equivalent partitions were encountered when partitioning into four states. If such a situation had been encountered in the first steps of the procedure, a predetermined rule of decision could have been used for deciding between otherwise equivalent solutions. Alternatively, the various equivalent partitions could have been tried in turn, the final choice being based on the overall minimum probability.

With more complex variables than the fictitious example above, a first probability minimum may be reached after partitioning the variable into a few states, and then another and lower minimum be found after partitioning into a larger number of states. It is suggested to stop the partitioning at the first probability minimum encountered, in order to keep the number of empty cells in the multidimensional contingency table as low as possible, and to establish the simplest possible correspondence between the classification and the discriminant variables (see the discussion below). Another decision might be more appropriate, however, in view of specific ecological problems: for instance, retaining the minimum corresponding to the number of states of D closest to the number of groups of C.

A similar approach was described by Pielou (1969) to divide quadrats into groups, using presence–absence data. The subdivision is made by choosing a "critical species" and dividing on that species. Quadrats are first divided into those containing and those not containing a first species. Data in the left-hand column of the contingency table are the numbers of occurrences of every other species, in quadrats containing the first species. Data in the right-hand column are similar numbers of occurrences, but in quadrats not containing the first species. χ^2 is then computed and the procedure is repeated for all the species in turn. As in our method, the first division of quadrats is on the species with the highest χ^2. The first two groups of quadrats are then divided and redivided by the same method, until all evidence of heterogeneity is removed.

Following the partitioning procedure described here, it is therefore possible to use any type of ordered variables, together with nonordered ones, in the discrete discriminant analysis of ecological classifications.

Ecological example

Baie de Penouille is a small cove of baie de Gaspé, in the Gulf of St. Lawrence (Fig. 1). Sampling was conducted at eight stations which were visited six times between July 1976 and June 1977 (Anonymous 1978). Various physical and chemical properties (including O_2 concentration of the water) were measured, and the vegetation growing at each station was recorded. Benthic organisms were quantitatively sampled and 48 taxa (among which annelids, molluscs, crustaceans, arachnids, and insects) were enumerated. A cursory analysis of the Penouille data is used here as an example; it is

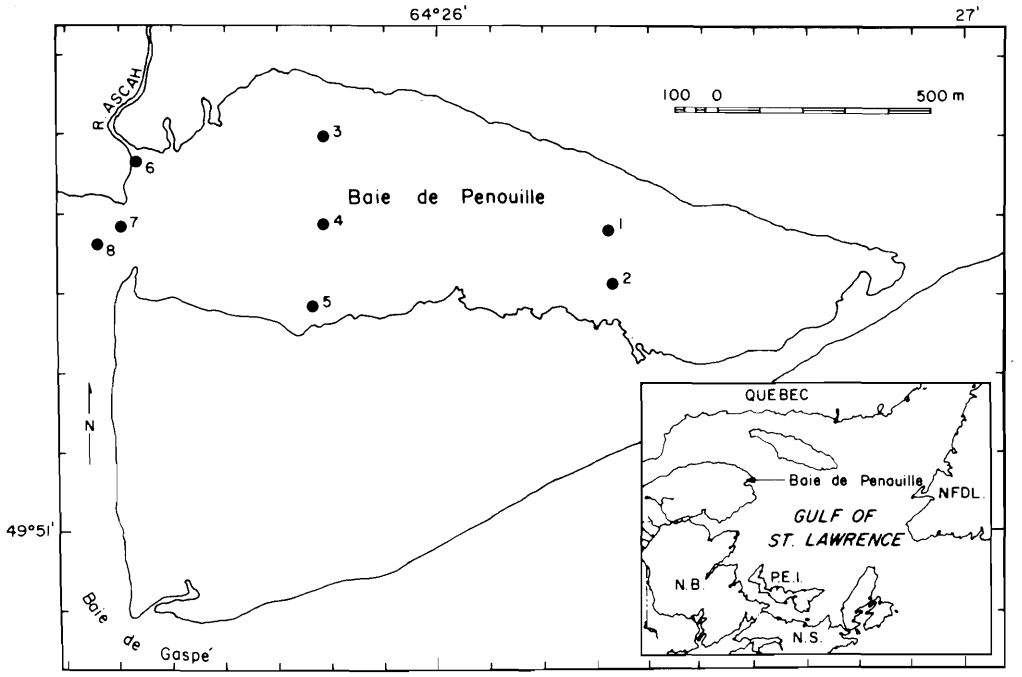

FIG. 1. Location of the sampling stations in baie de Penouille (Gulf of St. Lawrence).

not intended to provide a full ecological interpretation of the data set.

After elimination of 24 very rare taxa, a classification of the 48 samples was performed. Counts were normalized by log transformation. The χ^2 metric (Roux and Reysac 1975) was used to calculate a sample-by-sample distance matrix, from which principal coordinates (Gower 1966) were computed, as well as various agglomerative clustering classifications: single and complete linkage, proportional link linkage (Sneath 1966) with 0.75 connectedness, and flexible clustering (Lance and Williams 1966) with $\beta = -0.25$. After drawing the results of the clusterings onto the projection of the sample points on the two first principal coordinates, it was possible to recognize three clusters of samples, three sample points being left unclassified.

Seeking an interpretative model of the benthos classification, the discrete-states classification (Table 2) was compared with the three-states vegetation variable, and to dissolved oxygen. Since the latter is quantitative and continuous, it had to be divided into states before being used together with the qualitative vegetation variable in a discrete discriminant analysis.

Partitioning the O_2 variable into discrete states, given the benthos classification, was done according to the stepwise procedure described in the previous section. The steps of this partitioning are shown in Table 3. Partitioning the O_2 variable into five states resulted in a probability minimum, where the partitioning procedure ended.

Data from Table 2 are allocated to a three-way contingency table (Table 4). Various procedures of discrete discriminant analysis are then available (Goldstein and Dillon 1978). The analysis is conducted here by fitting log-linear models to the cell frequencies of the contingency table, since discrimination problems can be solved using multiway contingency table analysis (Goldstein and Dillon 1978, p. 25). To do so, logarithms of the expected cell frequencies (E) are written as an additive function of main effects and interactions between the variables. The full second-order model for the three variables (Table 4) is:

$$\ln E = [\theta] + [C] + [V] + [O] + [CV] + [CO] + [VO]$$

Symbols in brackets (see heading of Table 4) are the effects: for instance, $[CO]$ is the effect due to the interaction between the benthos classification (C) and the oxygen concentration (O). $[\theta]$ is the grand mean of the logarithms of the expected counts. The effect due to the interaction between all the variables $[CVO]$ is never included since, in the resulting saturated model, the expected cell frequencies simply turn out to be equal to the observed counts: $E = O$.

If only the effect of the interaction between the classification and the oxygen concentration is considered, for instance, the hierarchical model becomes:

$$\ln E = [\theta] + [C] + [V] + [O] + [CO]$$

the null hypothesis being then H_0: $[CV] = 0$ (no interaction between the benthos classification and the

TABLE 2. Data describing the 48 samples from baie de Penouille. The O_2 variable is presented both as raw data and after coding as in Table 3. Vegetation: 1 = epilithic algae, 2 = absent or *Ruppia* and *Zostera*, 3 = *Zostera* with or without *Fucus*. NC, unclassified sample (see text)

Station	Date	Classification	Vegetation	O_2 (mg·L^{-1})	O_2 (coded)
1	77-06-17	2	2	13.5	5
	76-07-06	3	2	6.5	2
	76-07-20	2	2	6.6	2
	76-08-03	2	2	7.3	2
	76-08-16	NC	2	10.4	—
	76-08-30	2	2	7.2	2
2	77-06-17	2	2	12.7	5
	76-07-06	3	2	6.2	1
	76-07-20	2	2	7.1	2
	76-08-03	2	2	7.0	2
	76-08-16	2	2	10.9	5
	76-08-30	3	2	7.4	3
3	77-06-17	2	2	11.9	5
	76-07-06	3	2	7.2	2
	76-07-20	2	2	6.6	2
	76-08-03	NC	2	7.7	—
	76-08-16	2	2	9.7	5
	76-08-30	2	2	8.9	5
4	77-06-17	2	3	11.9	5
	76-07-06	3	3	8.6	3
	76-07-20	3	3	7.2	2
	76-08-03	3	3	7.9	3
	76-08-16	3	3	10.3	5
	76-08-30	3	3	8.1	3
5	77-06-17	3	3	11.9	5
	76-07-06	3	3	7.9	3
	76-07-20	3	3	7.1	2
	76-08-03	3	3	7.0	2
	76-08-16	3	3	10.3	5
	76-08-30	3	3	7.5	3
6	77-06-17	NC	1	11.9	—
	76-07-06	1	1	6.1	1
	76-07-20	1	1	6.1	1
	76-08-03	1	1	6.4	1
	76-08-16	1	1	10.2	5
	76-08-30	1	1	8.8	4
7	77-06-17	3	3	12.3	5
	76-07-06	3	3	7.4	3
	76-07-20	3	3	6.8	2
	76-08-03	3	3	8.2	3
	76-08-16	3	3	9.4	5
	76-08-30	3	3	8.0	3
8	77-06-17	3	3	12.0	5
	76-07-06	3	3	7.2	2
	76-07-20	3	3	7.5	3
	76-08-03	3	3	8.2	3
	76-08-16	3	3	9.2	5
	76-08-30	3	3	8.4	3

TABLE 3. Results of the stepwise partitioning of the O_2 ($mg \cdot L^{-1}$) variable of Table 2, given the benthos classification scheme

Number of states	Partition after the value ($mg \cdot L^{-1}$)	χ^2	ν	Probability
2	6.4	11.712	2	0.00286
3	8.6	18.324	4	0.00107
4	7.3	27.933	6	0.00010
5	8.8	32.642	8	0.00007
6	12.3	36.046	10	0.00008
7	11.9	38.833	12	0.00011

vegetation types) and $[VO] = 0$ (no interaction between the two discriminant variables V and O). These models are hierarchical in the sense that a higher order effect cannot be present unless all lower order effects whose variables are subsets of the higher order effect are also included in the model: for instance, if $[VO]$ is present, then $[V]$ and $[O]$ must be included. Details on computations of these models are given in Fienberg (1970), Bishop *et al.* (1975), Fienberg (1980) and Upton (1978). Program BMDP4F (Dixon 1981) was used to compute the expected cell frequencies, for the hierarchical models adjusted to the multiway contingency table.

The goodness-of-fit of a model is tested with null hypothesis (H_0) that the effects not included in the model are zero. The test is performed using Wilks' likelihood ratio statistic, as in the previous section. When the probability associated with χ^2 is smaller than or equal to a preselected level α, the hypothesis H_0 is rejected. The alternative hypothesis that all the effects included in the model are nonzero cannot be accepted, however, since the only conclusion of the test is that at least some of these effects are nonzero. When a model is found with a probability larger than α, this model can be accepted as fitting the data. However, since the number of observations is small relative to the number of cells in the table, the resulting tests are overly liberal; that is, H_0 is rejected too often relative to α. Accordingly, α is taken as 0.01 rather than 0.05.

When several models are acceptable, the ecologist may choose the most parsimonious, that is the model with the largest possible number of null effects in its H_0 hypothesis. Dixon (1981) proposes to search for the most parsimonious, nonsignificant model, in two steps. First, a screening of all the separate effects, using a test of partial association of the factors, and then fitting the models thought to be most appropriate. Sokal and Rohlf (1981; p. 762) add to this that if any of the marginal totals have been fixed by the design of the experiment, then the corresponding term must be present in all models tested. When there are only three variables, as here, all the possible models can easily be fitted in turn (Table 5).

TABLE 4. Contingency table of the benthos classification (C) as a function of vegetation (V) and oxygen concentration (O). In the last column: configuration cells for interaction (VO). Total number of observations $N = 45$. Data from Table 2

Vegetation (V) (see Table 2)	Oxygen (O) ($mg \cdot L^{-1}$)	Benthos classification (C) 1	2	3	Interaction (VO)
1	0.0– 6.4	3	0	0	3
	6.5– 7.3	0	0	0	0
	7.4– 8.6	0	0	0	0
	8.7– 8.8	1	0	0	1
	8.9–13.5	1	0	0	1
2	0.0– 6.4	0	0	1	1
	6.5– 7.3	0	6	2	8
	7.4– 8.6	0	0	1	1
	8.7– 8.8	0	0	0	0
	8.9–13.5	0	6	0	6
3	0.0– 6.4	0	0	0	0
	6.5– 7.3	0	0	5	5
	7.4– 8.6	0	0	11	11
	8.7– 8.8	0	0	0	0
	8.9–13.5	0	1	7	8

A major problem in ecology is that very often some observed cell frequencies are zero, due to the small number of data relative to the number of cells in the contingency table; in Table 4, for instance, the $N = 45$ data are distributed among $3 \times 5 \times 3 = 45$ cells. In computing interactions, these empty cells may become so arranged that some of the configuration cells are empty; in Table 4, it is shown that five configuration cells for interaction (VO) are empty. These will result in a certain number of zero expected frequencies ($E = 0$), for which adjustment is necessary in the computation of the degrees of freedom. Program BMDP4F automatically provides such an adjustment, the bases of which can be found, for instance in Bishop *et al.* (1975, p. 116 and following pages) or in Dixon (1981, p. 666).

Model 15 has the best fit among those with interaction $[VO]$ fixed by the design. According to this model, the benthos classification is independent of the oxygen concentration.

The sole consideration of the various two-way contingency tables between the three variables (classification × vegetation, classification × oxygen and vegetation × oxygen) could not have led to the same conclusions, since the hypothesis of independence (see previous section) is rejected for the three tables: $\chi^2[CV] = 55.53$ ($\nu = 4$, $p < 0.001$), $\chi^2[CO] = 32.65$ ($\nu = 8$, $p < 0.001$) and $\chi^2[VO] = 30.07$ ($\nu = 8$, $p < 0.001$). The

TABLE 5. Models fitted to the contingency table (Table 4). Total number of cells: $3 \times 5 \times 3 = 45$. All models are hierarchical (see text): for instance in model 8, $[\theta]$, $[CV]$ stands for $[\theta]$, $[C]$, $[V]$, $[CV]$

Model	H_0: effects = 0	ν	χ^2	Probability
(1) $[\theta]$, $[C]$	$[V]$, $[O]$, $[CV]$, $[CO]$, $[VO]$, $[CVO]$	42	131.42	0.0000
(2) $[\theta]$, $[V]$	$[C]$, $[O]$, $[CV]$, $[CO]$, $[VO]$, $[CVO]$	42	134.82	0.0000
(3) $[\theta]$, $[O]$	$[V]$, $[O]$, $[CV]$, $[CO]$, $[VO]$, $[CVO]$	40	127.55	0.0000
(4) $[\theta]$, $[C]$, $[V]$	$[O]$, $[CV]$, $[CO]$, $[VO]$, $[CVO]$	40	117.78	0.0000
(5) $[\theta]$, $[C]$, $[O]$	$[V]$, $[CV]$, $[CO]$, $[VO]$, $[CVO]$	38	110.52	0.0000
(6) $[\theta]$, $[V]$, $[O]$	$[C]$, $[CV]$, $[CO]$, $[VO]$, $[CVO]$	38	113.91	0.0000
(7) $[\theta]$, $[C]$, $[V]$, $[O]$	$[CV]$, $[CO]$, $[VO]$, $[CVO]$	36	96.88	0.0000
(8) $[\theta]$, $[CV]$	$[O]$, $[CO]$, $[VO]$, $[CVO]$	20	62.25	0.0000
(9) $[\theta]$, $[CO]$	$[V]$, $[CV]$, $[VO]$, $[CVO]$	18	77.87	0.0000
(10) $[\theta]$, $[VO]$	$[C]$, $[CV]$, $[CO]$, $[CVO]$	20	83.85	0.0000
(11) $[\theta]$, $[CV]$, $[O]$	$[CO]$, $[VO]$, $[CVO]$	16	41.34	0.0005
(12) $[\theta]$, $[CO]$, $[V]$	$[CV]$, $[VO]$, $[CVO]$	16	64.23	0.0000
(13) $[\theta]$, $[VO]$, $[C]$	$[CV]$, $[CO]$, $[CVO]$	18	66.82	0.0000
(14) $[\theta]$, $[CV]$, $[CO]$	$[VO]$, $[CVO]$	4	8.70	0.0690
(15) $[\theta]$, $[CV]$, $[VO]$	$[CO]$, $[CVO]$	5	11.28	0.0461
(16) $[\theta]$, $[CO]$, $[VO]$	$[CV]$, $[CVO]$	6	34.17	0.0000
(17) $[\theta]$, $[CV]$, $[CO]$, $[VO]$	$[CVO]$	0	0.22	—

bivariate approach would therefore have been inconclusive as to the discrimination problem stated. As is always the case, multivariate ecological problems must be resolved using the proper available multivariate analyses.

Acknowledgments

The partitioning algorithm was programmed by Mr. Alain Vaudor, then computer analyst in the Centre de recherche en sciences de l'environnement, Université du Québec à Montréal, and the computer time was provided by the Service de l'informatique of this university. The data for the ecological example were made available to us by Dr. Bernadette Pinel-Alloul, Université de Montréal, and the classification of the benthic samples was performed by Ms. Martine Allard, graduate student at the Unviersité du Québec à Montréal. Dr. Louis-Paul Rivest, Département de mathématiques, Université Laval, provided enlightening comments on multiway contingency tables. Our thanks are also due to Dr. John Downing for his critical review of the manuscript. Individual research grants from the Natural Sciences and Engineering Research Council of Canada to both authors were instrumental in the completion of this work.

ANDRÉ, P., P. LEGENDRE, and P. P. HARPER. 1981. La sélectivité de trois engins d'échantillonnage du benthos lacustre. Ann. Limnol. **17**: 25–40.

ANONYMOUS. 1978. Étude d'impact d'utilisation au parc national Forillon. Tome I: Caractérisation du secteur de Penouille. Rapport final présenté par le Centre de Recherches écologiques de Montréal à Parcs Canada.

BISHOP, Y. M. M., S. E. FIENBERG, and P. W. HOLLAND. 1975. Discrete multivariate analysis: theory and practice. The MIT Press, Cambridge.

CASSIE, R. M. 1972. A computer programme for multivariate statistical analysis of ecological data. J. Exp. Mar. Biol. Ecol. **10**: 207–241.

COX, D. R. 1957. Note on grouping. J. Am. Stat. Assoc. **52**: 543–547.

DIXON, W. J. (*Editor*). 1981. BMDP statistical software 1981. University of California Press, Los Angeles.

FIENBERG, S. E. 1970. The analysis of multidimensional contingency tables. Ecology, **51**: 419–433.

———— 1980. The analysis of cross-classified categorical data. 2nd ed. The MIT Press, Cambridge.

FRONTIER, S. 1969. Sur une méthode d'analyse faunistique rapide du zooplancton. J. Exp. Mar. Biol. Ecol. **3**: 18–26.

———— 1973. Évaluation de la quantité totale d'une catégorie d'organismes planctoniques dans un secteur néritique. J. Exp. Mar. Biol. Ecol. **12**: 299–304.

GAUCH, H. G., JR., and T. R. WENTWORTH. 1976. Canonical correlation analysis as an ordination technique. Vegetatio, **33**: 17–22.

GOLDSTEIN, M., and W. R. DILLON. 1978. Discrete discriminant analysis. John Wiley & Sons, New York.

GOOD, I. J. 1950. Probability and the weighing of evidence. Charles Griffin, London.

GOWER, J. C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika, **53**: 325–338.

HUDON, C., E. BOURGET, and P. LEGENDRE. 1983. An integrated study of the factors influencing the choice of the settling site of *Balanus crenatus* cyprid larvae. Can. J. Fish. Aquat. Sci. **40**. In press.

LANCE, G. M., and W. T. WILLIAMS. 1966. A generalized sorting strategy for computer classifications. Nature (London), **212**: 218.

LEGENDRE, L., and P. LEGENDRE. 1983. Numerical ecology. Elsevier Scientific Publishing Co., Amsterdam.

PIELOU, E. C. 1969. Association tests versus homogeneity tests: their use in subdividing quadrats into groups. Vegetatio, **18**: 4–18.

ROUX, M., and J. REYSSAC. 1975. Essai d'application au phytoplancton marin de méthodes statistiques utilisées en phytosociologie terrestre. Ann. Inst. Océanogr. Paris, **51**: 89–97.

SNEATH, P. H. A. 1966. A comparison of different clustering methods as applied to randomly-spaced points. Classification Soc. Bull. **1**: 2–18.

SOKAL, R. R., and F. J. ROHLF. 1981. Biometry—The principles and practice of statistics in biological research. 2nd ed. W. H. Freeman, San Francisco.

UPTON, G. J. G. 1978. The analysis of cross-tabulated data. John Wiley & Sons, New York.

VINCENT, G., and Y. BERGERON. 1983. La caractérisation d'herbiers aquatiques du lac des Deux-Montagnes (Québec) à partir de paramètres physiques de l'eau. Can. J. Bot. **61**: 400–411.

WILKS, S. S. 1935. The likelihood test of independence in contingency tables. Ann. Math. Stat. **6**: 190–196.

YOUNG, F. W. 1981. Quantitative analysis of qualitative data. Psychometrika, **46**: 357–388.