Chapter

5

Multidimensional semiquantitative data

5.0 Nonparametric statistics

Statistical testing often refers to the concepts of *parameter* and *reference population*, as explained in Section 1.2. Section 4.3 showed that the mean, standard deviation and correlation are *parameters* of the multinormal distribution, so that this statistical distribution and others play a key role in testing *quantitative* data. When the data are *semiquantitative*, however, it does not make sense to compute statistics such as the mean or the standard deviation. In that case, hypothesis testing must be conducted with *nonparametric statistics*. This expression cover all statistical methods developed for analysing either *semiquantitative* (rank statistics; Section 5.2) or *qualitative* (Chapter 6) data.

Nonparametric tests are *distribution-free*, i.e. they do not assume that the samples were drawn from a population with a specified distribution (e.g. multinormal). Because of that, nonparametric statistics are useful not only when descriptors are semiquantitative, but also when quantitative descriptors do not conform to the multinormal distribution and researchers do not wish, or succeed, to normalize them. Many of the nonparametric tests for semiquantitative data are called *ranking tests* because they are based on ranked observations instead of quantitative values. Another advantage of nonparametric statistics is computational simplicity. Last but not least, nonparametric tests may be used with small samples, a situation that frequently occurs with ecological data; permutation tests based upon parametric statistics (Section 1.2) share this last advantage of nonparametric tests. For semiquantitative data, the nonparametric statistics corresponding to the *mean* and *variance* (Section 4.1) are the *median* and *range*, respectively.

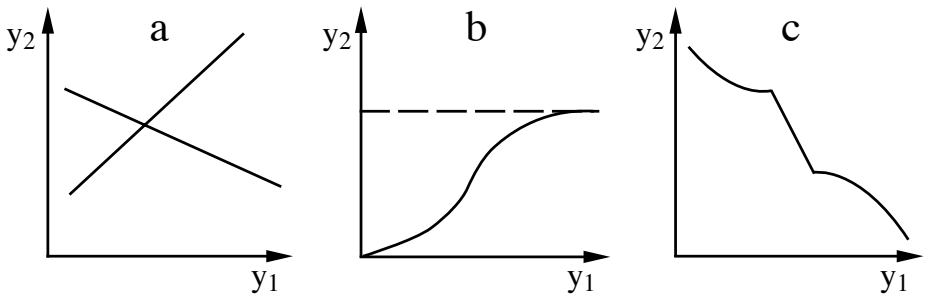

Figure 5.1 Three types of monotonic relationships between two descriptors: (a) linear (increasing and decreasing); (b) logistic (increasing monotonic); (c) atypical (decreasing monotonic).

Ranking tests should be used in the following situations:

- One or several descriptors among those to be compared are semiquantitative.

Monotonic

- The purpose of the study is not to evidence linear, but *monotonic* relationships between quantitative descriptors. In a bivariate monotonic relationship, one of the two descriptors keeps increasing or decreasing as the other increases (Fig. 5.1); the increase (or decrease) is not necessarily *linear* nor *smoothly curvilinear*.

Ranking tests or *permutation tests* (Section 1.2) can be used in the following cases:

- One or several quantitative descriptors are not normally distributed (tests of normality and multinormality are described in Section 4.6) and researchers do not wish to normalize them or do not succeed in doing so. Normalizing transformations are described in Subsection 1.5.6.

- The number of observations is small.

The present chapter first summarizes the methods available in the nonparametric approach, with reference to the corresponding parametric methods (Section 5.1). Tests for differences among groups using quantitative, semiquantitative or qualitative descriptors are compared in Section 5.2. Rank correlation coefficients are presented in Section 5.3. Section 5.4 is devoted to the Kendall coefficient of concordance, which is a generalization of the Spearman correlation coefficient to several descriptors. Most statistical computer packages, including R, offer nonparametric testing procedures.

5.1 Quantitative, semiquantitative, and qualitative multivariates

As discussed in Section 1.4, ecological descriptors may be of different levels of precision (Table 1.2). Ecologists generally observe several descriptors on the same objects, so that multidimensional ecological variates may be either quantitative, semiquantitative, or qualitative, or mixed, i.e. consisting of descriptors with different precision levels. For many years, quantitative ecology has been based almost exclusively on quantitative descriptors and on parametric tests, even though there exist a large number of methods that can efficiently analyse semiquantitative or qualitative multivariates as well as multivariates of mixed precision levels. These methods have become increasingly popular in ecology, not only because non-quantitative descriptors often provide unique information, but also because parametric statistics cannot be tested for significance when quantitative data do not conform to a number of conditions, including normality. This section briefly reviews numerical methods for analysing multivariates with various levels of precision.

Table 5.1 summarizes and compares methods described elsewhere in the present book. In the same row are found corresponding methods, listed under four column headings. The applicability of methods increases from left to right. Methods in the first (left-hand) column are restricted to *quantitative* multivariates, which must also, in most cases, be linearly related or/and multinormally distributed. Methods in the second column have been developed for *semiquantitative* descriptors exhibiting *monotonic* relationships. These methods may also be used (a) with quantitative descriptors, in particular when they do not follow the conditions underlying methods in the first column, and (b) for the combined analysis of quantitative and semiquantitative descriptors. Methods in the third column were developed for the numerical analysis of *qualitative* descriptors. They may also be used for analysing quantitative or semiquantitative descriptors exhibiting nonmonotonic relationships, after dividing these continuous descriptors into classes. Methods for qualitative descriptors thus represent a first type of techniques for multivariates of mixed precision, since they can be used for analysing together quantitative, semiquantitative, and qualitative descriptors, after the former have been divided into classes. An alternative is to recode multiclass qualitative descriptors into dummy variables (Subsection 1.5.7) and use parametric methods (first column of Table 5.1) on the resulting assemblage of quantitative and binary descriptors; this approach is often used in regression and canonical analyses (Chapters 10 and 11).

The methods listed in the right-hand column can be used for analysing data tables containing mixtures of quantitative, semiquantitative and qualitative descriptors. Of special interest are the distance-based methods (dbRDA, PCoA, nMDS, clustering), which can be applied after computing an association coefficient for mixed-level data. These methods are very general, since they may replace equivalent methods in the other three columns; the cost is sometimes greater mathematical and/or computational complexity.

Table 5.1 Methods for analysing *multidimensional* ecological data sets, classified here according to the levels of precision of descriptors (columns). For methods concerning data series, see Table 12.2. The Subject index at the end of the book shows where each method is described.

Quantitative descriptors	Semiquantitative descriptors	Qualitative descriptors	Descriptors of mixed precision
Difference between two samples:			
Hotelling T^2	---	Log-linear models	---
RDA, tbRDA, dbRDA	dbRDA	tbRDA, dbRDA	dbRDA
CCA		CCA	
Difference among several samples:			
MANOVA	---	Log-linear models	---
RDA, tbRDA, dbRDA	db-RDA	tbRDA, db-RDA	db-RDA
CCA		CCA	
Scatter diagram	Rank diagram	Multiway contingency table	Quantitative-rank diagram
Association coefficients R:			
Covariance	---	Information, X^2	---
Pearson *r*	Spearman *r*	Contingency	---
	Kendall τ		
Partial *r*	Partial τ		
Multiple *R*	Kendall *W*		
Species diversity:			
Diversity measures	Diversity measures	Number of species	---
Association coeff. Q	Association coeff. Q	Association coeff. Q	Association coeff. Q
Clustering	Clustering	Clustering	Clustering
Ordination:			
PCA, tbPCA, CA		tbPCA, CA	
PCoA	PCoA	PCoA	PCoA
nMDS	nMDS	nMDS	nMDS
Regression	Regression	Correspondence	Regression
simple linear (I and II)	nonparametric		logistic
multiple linear			dummy
polynomial			
partial linear			
nonlinear, logistic			
smoothing (splines, LOWESS)			
multivariate; see also canonical a.			
Path analysis	---	Log-linear models	
		Logit models	
Canonical analysis:			
RDA, tbRDA, dbRDA	dbRDA	tbRDA, dbRDA	db-RDA
CCA		CCA	
CCorA, CoIA		tbCCorA, tbCoIA	
LDA	---	Discrete discriminant a.	
		Log-linear models	
		Logistic regression	

There are many types of methods for multidimensional analysis (rows of Table 5.1). One interesting aspect of the table is that there is always at least one, and often several methods for descriptors with low precision levels. Thus, ecologists should not hesitate to collect information in semiquantitative or qualitative form since there exist numerical methods for processing descriptors with all levels of precision. However, it is always important to consider, at the stage of the sampling design (Fig. 1.3), how the data will be analysed, so as to avoid problems at the later stage of analysis. These problems may include the lack of human resources to efficiently use advanced numerical methods. Researchers could use the period devoted to sampling to improve their knowledge of methods and become familiar with computer programs and functions.

Coming back to Table 5.1, it is possible to compare groups of objects, described by quantitative multivariate data, using multidimensional analysis of variance (MANOVA). When there are only two groups, another approach is Hotelling's T^2 (Section 7.4). In the case of qualitative multivariate data, the comparison may be done by adjusting log-linear models (Section 6.3) to a multiway contingency table; the relationship between contingency table analysis and analysis of variance is explained in Section 6.0. Multivariate analysis of variance of species presence-absence or abundance tables may be carried out using either transformation-based redundancy analysis (tbRDA) or distance-based redundancy analysis (db-RDA) (Subsections 11.1.5 and 11.1.10), or else canonical correspondence analysis (CCA, Section 11.2).

The simplest approach to investigate the relationships among descriptors considered two at a time (Fig. 5.2) is to plot the data as a scatter diagram, whose semiquantitative and qualitative equivalent are the rank-rank diagram and the contingency table, respectively. Quantitative-rank diagrams may be used to compare a quantitative to a semiquantitative descriptor (Legendre & Legendre, 1982).

Two families of methods follow from these diagrams, for either *measuring* the dependence among descriptors, or *forecasting* one or several descriptors using other ones. The R-mode coefficients of dependence, described in Chapter 4 for quantitative descriptors, in Chapter 5 for semiquantitative descriptors, and in Chapter 6 for qualitative descriptors, measure the dependence between descriptors. These coefficients are summarized in Subsection 7.5.1. It is interesting to note that measures of information and X^2 (chi-square) calculated on contingency tables (Chapter 6) are the equivalent, for qualitative descriptors, of the covariance computed between quantitative descriptors. Methods in the second family belong to regression analysis (Section 10.3), which has nonparametric forms (e.g. the monotone regression method used in Section 9.4), and whose qualitative equivalent is the analysis of correspondence in contingency tables (Section 6.4).

Various measures of species diversity are reviewed in Section 6.5. They are usually computed on quantitative species counts, but Dévaux & Millerioux (1977) have shown that this may be done just as well on semiquantitative counts. When there are no counts, the number of species present may be used to assess diversity; this is indeed

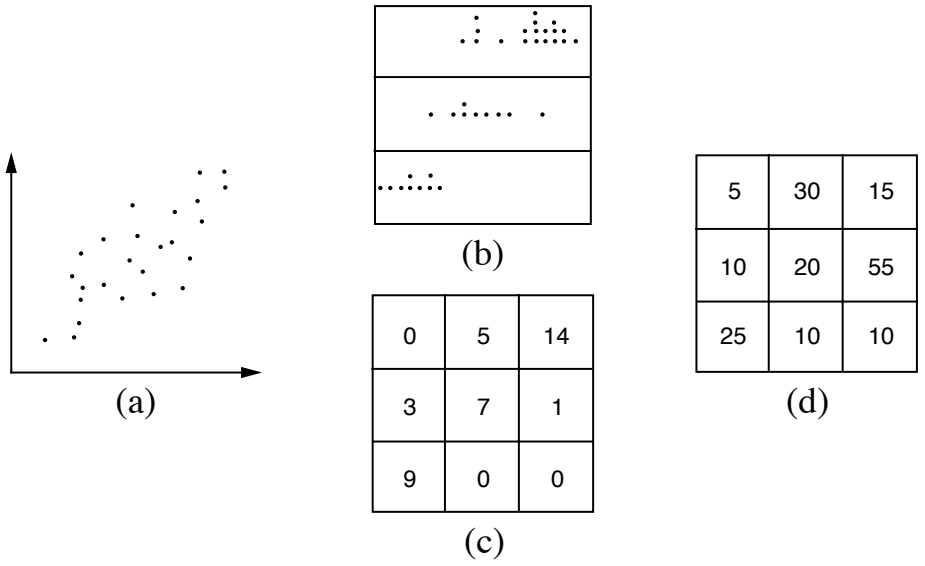

Figure 5.2 Comparison of two descriptors. (a) Scatter diagram (quantitative descriptors on both axes). (b) Quantitative-rank diagram (quantitative descriptor on the abscissa, ranked classes of a semiquantitative descriptor on the ordinate). (c) Rank-rank diagram (ranked classes of semiquantitative descriptors on both axes). (d) Two-way contingency table (nonordered classes of qualitative descriptors on both axes). From Legendre & Legendre (1982).

the first diversity index to have been described in the ecological literature (Patrick, 1949; Subsection 6.5.1).

There are Q-mode association coefficients (Sections 7.3 and 7.4) adapted to descriptors of all levels of precision (see Tables 7.4 and 7.5). Some of the similarity coefficients (Chapter 7: S_{15}, S_{16}) are yet another way of combining quantitative and qualitative descriptors in multivariate data analysis. Concerning clustering algorithms (Chapter 8), most of them are indifferent to the precision of descriptors, since clustering is in general conducted on an association matrix, most often of type Q.

Among the ordination methods in reduced space, principal component analysis (PCA, Section 9.1) is the main method to use with quantitative descriptors, although it can also be applied to semiquantitative data (Subsection 9.1.7). Species abundance or presence-absence data, as well as other types of frequency data, can be analysed by correspondence analysis (CA, Section 9.2) or by transformation-based PCA (tbPCA). Principal coordinate analysis (PCoA, Section 9.3) and nonmetric multidimensional scaling (nMDS, Section 9.4) are indifferent to the precision of descriptors since they are computed on an association matrix (generally Q-type).

For the interpretation of ecological structures, regression, which was briefly discussed a few paragraphs above, is the chief technique when the dependent variable is a single quantitative variable. Logistic regression is used when the response is presence-absence data. Various forms of canonical analysis are available to interpret the structure of quantitative data using one or several tables of explanatory variables: redundancy analysis (RDA, Section 11.1), canonical correspondence analysis (CCA, Section 11.2), linear discriminant analysis (LDA, Section 11.3), canonical correlation analysis (CCorA, Section 11.4), and co-inertia analysis (CoIA, Section 11.5). Canonical correspondence analysis, as well as tbRDA, allow the interpretation of the structure of species abundance or presence-absence data using explanatory variables. For non-quantitative data, distance-based RDA (dbRDA) can be used after computing a distance matrix using an appropriate distance function. There are also methods equivalent to discriminant and path analyses for qualitative descriptors.

Table 5.1 shows that ecological data can efficiently be analysed irrespective of their levels of precision. Researchers should use ecological criteria, such as allowable effort in the field and biological meaningfulness of the decimal places to be recorded, to decide about the level of precision of their data. The strictly numerical aspects should play a secondary role in that decision.

5.2 One-dimensional nonparametric statistics

The present book is devoted to numerical methods for analysing sets of *multidimensional* ecological data. Methods for one-dimensional variables are not discussed in depth since they are the subject of many excellent textbooks. Nonparametric tests for one-dimensional descriptors are explained, among others, in the books of Siegel (1956), Hájek (1969), Siegel & Castellan (1988), and Sokal & Rohlf (1995). Because ecologists are often not fully conversant with these tests, the correspondence between approaches for quantitative, semiquantitative, and qualitative descriptors is not always clearly understood. This is why the one-dimensional methods to carry out tests of differences among groups of objects are summarized in Table 5.2.

Independent samples
Related samples

Methods in the table are divided in two main families: those for independent samples, which are the most generally applicable, and those for related samples. Related samples are often called *matched* or *paired* samples (Box 1.1). With such samples of observations, the analysis may focus either on the differences between the matched observation units, or on the differences among the classes of another factor while controlling for the differences between the matched observations. Matching may be achieved, for example, by repeating observations at the same sampling sites at different times, or by making observations at points representing corresponding conditions, e.g. in several lakes with sampling units taken from the same depths in the water columns. Sampling units observed before and after a treatment also form matched pairs. When related samples are analysed using the methods for independent

Table 5.2 Methods to carry out tests of differences among groups of objects (*one-dimensional* data) are classified here according to the levels of precision of the descriptors (columns). Most of these methods are not discussed elsewhere in the present book. Table modified from Siegel (1956) and Legendre & Legendre (1982).

Number of groups (k)	Quantitative descriptors*	Semiquantitative descriptors	Qualitative descriptors
Independent samples:			
$k = 2$	Student t (unpaired)	Mann-Whitney U-test Median test Kolmogorov-Smirnov test etc.	X^2 (2 × no. states) Fisher's exact probability test Logistic regression
$k \geq 2$ (one-way)	One-way ANOVA and F-test	Kruskal-Wallis' H Extension of the median test	X^2 (k × no. states) Discriminant a.
Related samples:			
$k = 2$	Student t (paired)	Sign test Wilcoxon signed-ranks test	McNemar test (binary descriptors)
$k \geq 2$ (two-way)	Two-way ANOVA and F-tests	Friedman's two-way ANOVA by ranks	Cochran Q (binary descriptors)
$k \geq 2$ (multiway)	Multiway ANOVA and F-tests	---	---

* When quantitative data do not meet the distributional assumptions underlying parametric tests, they must be analysed using ranking tests (for semiquantitative descriptors). Another way would be to test the parametric statistics by permutation (Section 1.2).

samples, the matching information is not taken into account and this results in a less powerful statistical test. Within each of the two families, methods in Table 5.2 are classified according to the number of groups (k) that are compared.

Univariate comparison of *two independent samples* ($k = 2$), when the data are *quantitative*, is generally done by using the Student t-statistic to test the hypothesis (H_0) of equality of the group means (i.e. that the two groups of objects were drawn from the same statistical population, or perhaps from populations with equal means, assuming equal standard deviations). When the data are *semiquantitative*, computing means and standard deviations would not make sense, so that the approach must be nonparametric. The Mann-Whitney U-statistic first combines and ranks all objects in a single series and then tests the hypothesis (H_0) that the ranked observations come from the same statistical population or from populations that have the same median. The median test, which is not as powerful as the previous one (except in cases when there are ties), is used for testing the hypothesis (H_0) that the two groups of objects have similar medians. Other nonparametric tests consider not only the positions of the two

groups along the abscissa but also the differences in dispersion and shape (e.g. skewness) of their distributions. The best-known is the Kolmogorov-Smirnov test; this is not the same test as the one described in Section 4.6 for comparing an empirical to a theoretical distribution. The Kolmogorov-Smirnov method discussed here allows one to test the hypothesis (H_0) that the largest difference between the cumulative distributions of the two groups is so small that they may come from the same or identical populations. Finally, when the data are *qualitative*, the significance of differences between two groups of objects may be tested using a X^2-statistic calculated on a two-way contingency table. Section 6.2 describes contingency table analysis for the comparison of two descriptors. In the present case, the contingency table has two rows (i.e. two groups of objects) and as many columns as there are states in the quantitative descriptor. The hypothesis tested (H_0) is that the frequency distributions in the two rows are similar; this is the same as stating the more usual hypothesis of independence between rows and columns of the contingency table (Section 6.0). When the descriptor is *binary* (e.g. presence or absence) and the number of observations in the two groups is small, it is possible to test the hypothesis (H_0) that the two groups exhibit similar proportions for the two states, using Fisher's powerful exact probability test. Logistic regression (Subsection 10.3.7) may also be used in this context; in the regression, the two groups are represented by a binary response variable while the qualitative explanatory descriptors are recoded as a series of dummy variables, coded as shown in Subsection 1.5.7.

The standard parametric technique for testing that the means of *several independent samples* ($k \geq 2$) are equal, when the data are *quantitative*, is one-way analysis of variance (ANOVA). It is a k-group generalization of the Student t-test. In one-way ANOVA, the overall variance is partitioned between two orthogonal (i.e. linearly independent; see Box 1.1) components, the first one reflecting differences among the k groups and the second one accounting for the variability among objects within the groups. The hypothesis (H_0) of equal means is rejected (F-test) when the among-group variability is significantly larger than the within-group component. For *semiquantitative* data, the Kruskal-Wallis' H-test (also called Kruskal-Wallis' one-way ANOVA by ranks) first ranks all objects from the k groups into a single series, and then tests (H_0) that the sums of ranks calculated for the various groups are so similar that the objects are likely to have been drawn from the same or identical populations. When applied to quantitative data that are meeting all the assumptions of parametric ANOVA, Kruskal-Wallis' H is almost as powerful as the F-test. Another possibility is to extend to $k \geq 2$ groups the median test, described in the previous paragraph for $k = 2$. The latter is less powerful than Kruskal-Wallis' H because it uses less of the information in the data. As in the above case where $k = 2$, *qualitative* data can be analysed using a contingency table, but this time with $k \geq 2$ rows.

Multinomial logistic regression

To model multistate qualitative response data, *multinomial logistic regression* is available in R (see Section 5.5) as well as in procedure CATMOD of SAS. Discriminant analysis could be used in the same spirit. See the discussion on discriminant analysis *versus* logistic regression in Section 11.6 (point 2).

Comparing *two related samples* ($k = 2$) is usually done, for *quantitative* data, by testing (H_0) that the mean of the differences between matched pairs of observations is null (Student *t*-test; the differences are assumed to be normally and independently distributed). When the data are *semiquantitative*, one can use the sign test, which first codes pairs of observations (y_i, y_k) as either (+) when $y_i > y_k$ or (–) when $y_i < y_k$, and then tests the hypothesis (H_0) that the numbers of pairs with each sign are equal; an equivalent formulation is that the proportion of pairs with either sign is equal to 0.5. This test uses information about the direction of the differences between pairs. When the relative magnitude of the differences between pairs is also known, it becomes possible to use the more powerful Wilcoxon matched-pairs signed-ranks test. Differences between pairs are first ranked according to their magnitude (absolute values), after which the sign of the difference is affixed to each rank. The null hypothesis of the test (H_0) is that the sum of the ranks having a (+) sign is equal to that of the ranks with a (–) sign. The McNemar test provides a means of comparing paired samples of *binary* data. For example, using binary observations (e.g. presence or absence) made at the same sites, before and after some event, one could test (H_0) that no overall change has occurred.

When there are *several related samples* ($k \geq 2$) and the data are *quantitative*, the parametric approach for testing (H_0) that the means of the k groups are equal is two-way analysis of variance, with or without replication. One classification criterion of the two-way ANOVA accounts for the variability among the k groups (as in one-way ANOVA above, for $k \geq 2$ independent samples) and the other for that among the related samples. Consider, as example, 16 sites (i.e. k groups) that have been sampled at 5 depths in the water column (or at 5 different times, or using 5 different methods, etc.). The nonparametric equivalent, for *semiquantitative* data, is Friedman's *two-way analysis of variance by ranks without replication*, which is based on a two-way table like Table 5.7. In the two-way table, the k groups (e.g. 16 sites) are in rows and the corresponding samples (e.g. 5 depths) are in columns. Values within each column are ranked separately, and the Friedman X^2-statistic (eq. 5.15) is used to test (H_0) that the rank totals of the various rows (e.g. 16 sites) are equal. For *binary* data, the Cochran Q test is an extension to $k \geq 2$ groups of the McNemar test, described above for $k = 2$.

Finally, when there are *several samples ($k \geq 2$), related across several classification criteria* (e.g. 16 sites all sampled at 8 different times, using each time 5 different methods), multiway ANOVA is the standard parametric method for testing the null hypothesis (H_0) that the means of the k groups are equal (F-test). In that case, there are no obvious equivalent approaches for semiquantitative or qualitative data.

How to analyse multivariate data representing related samples is described in Subsection 11.1.10, point 3.

Table 5.3 Numerical example. Perfect rank correlation between descriptors $\mathbf{y}_1$ and $\mathbf{y}_2$.

Objects (observation units)	Ranks of objects on the two descriptors	
	$\mathbf{y}_1$	$\mathbf{y}_2$
$\mathbf{x}_1$	5	5
$\mathbf{x}_2$	1	1
$\mathbf{x}_3$	4	4
$\mathbf{x}_4$	2	2
$\mathbf{x}_5$	3	3

5.3 Rank correlations

Textbooks of nonparametric statistics propose a few methods only for the analysis of bi- or multivariate semiquantitative data. Section 5.1 has shown that there actually exist many numerical approaches for analysing multidimensional data, corresponding to all levels of precision (Table 5.1). These methods, which include most of those described in this book, belong to *nonparametric statistics* in a general sense, because they do not focus on the parameters of the data distributions. Within the specific realm of *ranking tests*, however, the statistical techniques available for multidimensional semiquantitative data are two *rank correlation coefficients* (Spearman r and Kendall τ), which both quantify the relationship between two descriptors, and the *coefficient of concordance* (Kendall W), which assesses the relationship among several descriptors. The two correlation coefficients are described in the present section and coefficient W in Section 5.4.

1 — Spearman r

Spearman corr. coeff.

The Spearman r coefficient, also called ρ (rho), is based on the idea that two descriptors $\mathbf{y}_1$ and $\mathbf{y}_2$ carry the same information if the object with the largest rank on $\mathbf{y}_1$ also has the highest rank on $\mathbf{y}_2$, and so on for all other objects. Two descriptors are said to be in perfect correlation when the ranks of all objects are the same on both descriptors, as in the numerical example shown in Table 5.3. If, however, object $\mathbf{x}_1$ which has rank 5 on $\mathbf{y}_1$ had rank 2 on $\mathbf{y}_2$, it would be natural to use the difference between these ranks, $d_1 = (y_{11} - y_{12}) = (5 - 2) = 3$, as a measure of the difference between the rankings given to this object by the two descriptors. For the whole set of objects, differences d_i are squared before summing them, to prevent differences with opposite signs from cancelling each other out.

The expression for the Spearman *r* correlation coefficient may be derived from the general formula of correlation coefficients (Kendall, 1948):

$$r_{jk} = \frac{\sum_{i=1}^{n} (y_{ij} - \bar{y}_j)\,(y_{ik} - \bar{y}_k)}{\sqrt{\sum_{i=1}^{n} (y_{ij} - \bar{y}_j)^2 \sum_{i=1}^{n} (y_{ik} - \bar{y}_k)^2}} \qquad \textbf{(5.1)}$$

For quantitative data, this equation is used to compute the Pearson linear correlation coefficient (eq. 4.7).

For ranked data, the average ranks $\bar{y}_j$ and $\bar{y}_k$ are equal, so that $(y_{ij} - \bar{y}_j) - (y_{ik} - \bar{y}_k) = (y_{ij} - y_{ik})$. One can write the difference between the ranks of object *i* on the two descriptors as $d_i = (y_{ij} - y_{ik}) = (y_{ij} - \bar{y}_j) - (y_{ik} - \bar{y}_k)$, which leads to:

$$\sum_{i=1}^{n} d_i^2 = \sum_{i=1}^{n} (y_{ij} - \bar{y}_j)^2 + \sum_{i=1}^{n} (y_{ik} - \bar{y}_k)^2 - 2\sum_{i=1}^{n} (y_{ij} - \bar{y}_j)\,(y_{ik} - \bar{y}_k)$$

Isolating the right-hand sum gives:

$$\sum_{i=1}^{n} (y_{ij} - \bar{y}_j)\,(y_{ik} - \bar{y}_k) = \frac{1}{2}\left[\sum_{i=1}^{n} (y_{ij} - \bar{y}_j)^2 + \sum_{i=1}^{n} (y_{ik} - \bar{y}_k)^2 - \sum_{i=1}^{n} d_i^2\right]$$

Using this result, eq. 5.1 is rewritten as:

$$r_{jk} = \frac{\frac{1}{2}\left[\sum_{i=1}^{n} (y_{ij} - \bar{y}_j)^2 + \sum_{i=1}^{n} (y_{ik} - \bar{y}_k)^2 - \sum_{i=1}^{n} d_i^2\right]}{\sqrt{\sum_{i=1}^{n} (y_{ij} - \bar{y}_j)^2 \sum_{i=1}^{n} (y_{ik} - \bar{y}_k)^2}} \qquad \textbf{(5.2)}$$

The sum of ranks for each descriptor, which is the sum of the first *n* integers, is equal to $\sum_{i=1}^{n} y_{ij} = n(n+1)/2$ and the sum of their squares is $\sum_{i=1}^{n} y_{ij}^2 = n(n+1)(2n+1)/6$. Since the sum of deviations from the mean rank is

$$\sum_{i=1}^{n} (y_{ij} - \bar{y}_j)^2 = \sum_{i=1}^{n} y_{ij}^2 - \frac{1}{n}\left(\sum_{i=1}^{n} y_{ij}\right)^2$$

one can write:

$$\sum_{i=1}^{n} (y_{ij} - \bar{y}_j)^2 = \frac{n(n+1)(2n+1)}{6} - \frac{1}{n}\left[\frac{n^2(n+1)^2}{4}\right] = \frac{n^3 - n}{12}$$

It follows that, when using ranks, the numerator of eq. 5.2 becomes:

$$\frac{1}{2}\left[\sum_{i=1}^{n}(y_{ij}-\bar{y}_j)^2+\sum_{i=1}^{n}(y_{ik}-\bar{y}_k)^2-\sum_{i=1}^{n}d_i^2\right]=\frac{1}{2}\left[\frac{n^3-n}{12}+\frac{n^3-n}{12}-\sum_{i=1}^{n}d_i^2\right]$$

while its denominator reduces to:

$$\sqrt{\sum_{i=1}^{n}(y_{ij}-\bar{y}_j)^2\sum_{i=1}^{n}(y_{ik}-\bar{y}_k)^2}=\sqrt{\left(\frac{n^3-n}{12}\right)\left(\frac{n^3-n}{12}\right)}=\frac{n^3-n}{12}$$

The final formula is obtained by replacing the above two expressions in eq. 5.2. This development shows that, when using ranks, eq. 5.1 simplifies to the following formula for Spearman *r*:

$$r_{jk}=\frac{\frac{1}{2}\left[\frac{n^3-n}{12}+\frac{n^3-n}{12}-\sum_{i=1}^{n}d_i^2\right]}{\frac{n^3-n}{12}}=1-\frac{6\sum_{i=1}^{n}d_i^2}{n^3-n} \qquad \textbf{(5.3)}$$

Alternatively, the Spearman rank correlation coefficient can be obtained in two steps: (1) replace all observations by ranks (columnwise) and (2) compute the Pearson correlation coefficient (eq. 4.7, formula identical to eq. 5.1) between the ranked variables. The result is the same as obtained from eq. 5.3.

The Spearman *r* coefficient varies between +1 and –1, just like the Pearson *r*. Descriptors that are perfectly matched, in terms of ranks, exhibit values $r = +1$ (direct relationship) or $r = -1$ (inverse relationship), whereas $r = 0$ indicates the absence of a monotonic relationship between the two descriptors. (Relationships that are not monotonic, e.g. Fig. 4.4d, can be quantified using polynomial or nonlinear regression, or else contingency coefficients; see Section 6.2 and Subsection 10.3.4.)

Numerical example. A small example (ranked data, Table 5.4) illustrates the equivalence between eq. 5.1 computed on ranks and eq. 5.3. Using eq. 5.1 gives:

$$r_{12}=\frac{-2}{\sqrt{5\times 5}}=\frac{-2}{5}=-0.4$$

The same result is obtained from eq. 5.3:

$$r_{12}=1-\frac{6\times 14}{4^3-4}=1-\frac{84}{60}=1-1.4=-0.4$$

Two or more objects may have the same rank on a given descriptor. This is often the case with descriptors used in ecology, which may have a small number of states or ordered classes. Such observations are said to be *tied*. Each of them is assigned the

Table 5.4 Numerical example. Ranks of four objects on two descriptors, $\mathbf{y}_1$ and $\mathbf{y}_2$.

Objects (observation units)	Ranks of objects on the two descriptors	
	$\mathbf{y}_1$	$\mathbf{y}_2$
$\mathbf{x}_1$	3	3
$\mathbf{x}_2$	4	1
$\mathbf{x}_3$	2	4
$\mathbf{x}_4$	1	2

average of the ranks that would have been assigned had no ties occurred. If the proportion of tied observations is large, correction factors must be introduced into the sums of squared deviations of eq. 5.2, which become:

$$\sum_{i=1}^{n} (y_{ij} - \bar{y}_j)^2 = \frac{1}{12}\left[(n^3 - n) - \sum_{r=1}^{q} (t_{rj}^3 - t_{rj}) \right]$$

and

$$\sum_{i=1}^{n} (y_{ik} - \bar{y}_k)^2 = \frac{1}{12}\left[(n^3 - n) - \sum_{r=1}^{s} (t_{rk}^3 - t_{rk}) \right]$$

where t_{rj} and t_{rk} are the numbers of observations in descriptors $\mathbf{y}_j$ and $\mathbf{y}_k$ that are tied at ranks r, these values being summed over the q sets of tied observations in descriptor j and the s sets in descriptor k.

Significance of the Spearman coefficient is usually tested against the null hypothesis H_0: $r = 0$. When $n \geq 10$, the test statistic is the same as for Pearson r (eq. 4.13):

$$t = \sqrt{\nu} \frac{r_{ki}}{\sqrt{1 - r_{ki}^2}} \qquad \textbf{(5.4)}$$

H_0 is tested by comparing statistic t to the value found in a table of critical values of t with $\nu = n - 2$ degrees of freedom. H_0 is rejected when the probability corresponding to t is smaller than or equal to a predetermined level of significance (α, for a two-tailed test). The rules for one-tailed and two-tailed tests are the same as for the Pearson r (Section 4.2). When $n < 10$, which is not often the case in ecology, one must refer to a

Table 5.5 Numerical example. The order of the four objects (rows) of Table 5.4 has been rearranged in such a way that the ranks on $\mathbf{y}_1$ are now in increasing order

Objects (observation units)	Ranks of objects on the two descriptors	
	$\mathbf{y}_1$	$\mathbf{y}_2$
$\mathbf{x}_4$	1	2
$\mathbf{x}_3$	2	4
$\mathbf{x}_1$	3	3
$\mathbf{x}_2$	4	1

special table of critical values of the Spearman rank correlation coefficient, found in textbooks of nonparametric statistics.

2 — Kendall τ

Kendall corr. coeff.

Kendall τ (tau) is another rank correlation coefficient, which can be used for the same types of descriptors as Spearman *r*. One major advantage of τ over Spearman *r* is that the former can be generalized to a partial correlation coefficient (below), which is not the case for the latter. While Spearman *r* was based on the differences between the ranks of objects on the two descriptors being compared, Kendall τ refers to a somewhat different concept, which is best explained using an example.

Numerical example. Kendall τ is calculated on the example of Table 5.4, already used for computing Spearman *r*. In Table 5.5, the order of the objects was rearranged so as to obtain increasing ranks on one of the two descriptors (here $\mathbf{y}_1$). The table is used to determine the degree of dependence between the two descriptors. Since the ranks are now in increasing order on $\mathbf{y}_1$, it is sufficient to determine how many *pairs of ranks* are also in increasing order on $\mathbf{y}_2$ to obtain a measure of the association between the two descriptors. Considering the object in first rank (i.e. $\mathbf{x}_4$), at the top of the right-hand column, the first pair of ranks (2 and 4, belonging to objects $\mathbf{x}_4$ and $\mathbf{x}_3$) is in increasing order; a score of +1 is assigned to it. The same goes for the second pair (2 and 3, belonging to objects $\mathbf{x}_4$ and $\mathbf{x}_1$). The third pair of ranks (2 and 1, belonging to objects $\mathbf{x}_4$ and $\mathbf{x}_2$) is in decreasing order, however, so that it earns a negative score –1. The same operation is repeated for every object in successive ranks along $\mathbf{y}_1$, i.e. for the object in second rank ($\mathbf{x}_3$): first pair of ranks (4 and 3, belonging to objects $\mathbf{x}_3$ and $\mathbf{x}_1$), etc. The sum *S* of scores assigned to each of the $n(n-1)/2$ different pairs of ranks is then computed.

Kendall's rank correlation coefficient is defined as follows:

$$\tau_a = \frac{S}{n(n-1)/2} = \frac{2S}{n(n-1)} \qquad \textbf{(5.5)}$$

Table 5.6 Numerical example. Contingency table giving the distribution of 80 objects among the states of two semiquantitative descriptors, **a** and **b**. Numbers in the table are frequencies (*f*).

	b_1	b_2	b_3	b_4	t_j
a_1	20	10	10	0	40
a_2	0	10	0	10	20
a_3	0	0	10	0	10
a_4	0	0	0	10	10
t_k	20	20	20	20	80

where *S* stands for "sum of scores". Kendall τ_a is thus the sum of scores for pairs in increasing and decreasing order divided by the total number of pairs ($n(n-1)/2$). For the example of Tables 5.4 and 5.5, τ_a is:

$$\tau_a = \frac{2\,(1+1-1-1-1-1)}{4 \times 3} = \frac{2\,(-2)}{12} = -0.33$$

Clearly, in the case of perfect agreement between two descriptors, all pairs receive a positive score, so that $S = n(n-1)/2$ and thus $\tau_a = +1$. When there is complete disagreement, $S = -n(n-1)/2$ and thus $\tau_a = -1$. When the descriptors are totally unrelated, the positive and negative scores cancel out, so that *S* as well as τ_a are 0 or near 0.

Equation 5.5 cannot be used for computing τ when there are tied observations. This is often the case with ecological *semiquantitative* descriptors, which may have a small number of states. The Kendall rank correlation is then computed on a contingency table crossing two semiquantitative descriptors.

Table 5.6 is a contingency (or frequency) table crossing two ordered descriptors. For example, descriptor **a** could represent the relative abundances of arthropods in soil enumerated on a semiquantitative scale (e.g. absent, present, abundant, and very abundant), while descriptor **b** could be the concentration of organic matter in the soil, divided into 4 classes. For simplicity, descriptors are called **a** and **b** here, as in Chapter 6. The states of **a** vary from 1 to *r* (number of rows) while the states of **b** go from 1 to *c* (number of columns).

To compute τ with tied observations, *S* is calculated as the difference between the numbers of positive (*P*) and negative (*Q*) scores, $S = P - Q$. *P* is the sum of all

frequencies f in the contingency table, each one multiplied by the sum of all frequencies located *lower* and on its *right*:

$$P = \sum_{j=1}^{r} \sum_{k=1}^{c} \left[f_{jk} \times \sum_{l=j+1}^{r} \sum_{m=k+1}^{c} f_{lm} \right]$$

Likewise, Q is the sum of all frequencies f in the table, each one multiplied by the sum of all frequencies *lower* and on its *left*:

$$Q = \sum_{j=1}^{r} \sum_{k=1}^{c} \left[f_{jk} \times \sum_{l=j+1}^{r} \sum_{m=1}^{k-1} f_{lm} \right]$$

Numerical example. For the data in Table 5.6:

$$P = (20 \times 40) + (10 \times 30) + (10 \times 20) + (10 \times 20) + (10 \times 10) = 1600$$
$$Q = (10 \times 10) + (10 \times 10) = 200$$
$$S = P - Q = 1600 - 200 = 1400$$

Using this value S, there are two approaches for calculating τ, depending on the numbers of states in the two descriptors. When **a** and **b** have the same numbers of states ($r = c$), τ_b is computed using a formula that includes the total number of pairs $n(n-1)/2$, as in the case of τ_a (eq. 5.5). The difference with eq. 5.5 is that τ_b includes corrections for the number of pairs L_1 tied in **a** and the number of pairs L_2 tied in **b**, where

$$L_1 = \sum_{j=1}^{r} \frac{1}{2} t_j (t_j - 1) \quad \text{in which } t_j \text{ is the marginal total for row } j$$

$$L_2 = \sum_{k=1}^{c} \frac{1}{2} t_k (t_k - 1) \quad \text{in which } t_k \text{ is the marginal total for column } k.$$

The formula for τ_b is:

$$\tau_b = \frac{S}{\sqrt{\frac{1}{2} n(n-1) - L_1}\sqrt{\frac{1}{2} n(n-1) - L_2}} \qquad \textbf{(5.6)}$$

When there are no tied observations, $L_1 = L_2 = 0$ and eq. 5.6 is identical to eq. 5.5.

Numerical example. For the data in Table 5.6:

$$L_1 = \frac{40 \times 39}{2} + \frac{20 \times 19}{2} + \frac{10 \times 9}{2} + \frac{10 \times 9}{2} = 1060$$

$$L_2 = \frac{20 \times 19}{2} + \frac{20 \times 19}{2} + \frac{20 \times 19}{2} + \frac{20 \times 19}{2} = 760$$

$$\tau_b = \frac{1400}{\sqrt{\frac{1}{2}(80 \times 79) - 1060}\sqrt{\frac{1}{2}(80 \times 79) - 760}} = 0.62$$

Without correction for ties, the calculated value (eq. 5.5) would have been

$$\tau_a = (2 \times 1400) / (80 \times 79) = 0.44$$

The second approach for calculating τ with tied observations is used when **a** and **b** do not have the same number of states ($r \neq c$). The formula for τ_c uses the minimum number of states in either **a** or **b**, $min(r, c)$:

$$\tau_c = \frac{S}{\frac{1}{2}n^2\left(\frac{min-1}{min}\right)} \quad \textbf{(5.7)}$$

The significance of Kendall τ is tested against the null hypothesis H_0: $r = 0$ (i.e. independence of the two descriptors). Kendall (1948) has shown that the distribution of τ approximates the normal distribution with mean $\mu_\tau = 0$ and standard deviation $\sqrt{2(2n+5)(9n(n-1))}$. Hence a z-test statistic can be obtained by transforming τ into a standard normal variate z using the formula:

$$z = \left[|\tau|\sqrt{\frac{9n(n-1)}{2(2n+5)}}\right] - \sqrt{\frac{18}{n(n-1)(2n+5)}} \quad \textbf{(5.8)}$$

With this statistic, H_0 can be tested using a table of z (or t_∞). Since z tables are one-tailed, the z-statistic of eq. 5.8 may be used directly for one-tailed tests by comparing it to the value z_α read in the table. For two-tailed tests, the statistic is compared to the value $z_{\alpha/2}$ from the z-table. When n is large, the second term of eq. 5.8 (correction for small n) becomes small: for $n = 30$, its value is 0.0178, and it is 0.0084 for $n = 50$.

Spearman r provides a better approximation of Pearson r when the data are almost quantitative and there are but a few tied observations, whereas Kendall τ does better when there are many ties. Computing both Spearman r and Kendall τ_a on the same numerical example, above, produced different numerical values (i.e. $r = -0.40$ versus $\tau_a = -0.33$). This is because the two coefficients have different underlying scales, so that their numerical values cannot be directly compared. However, given their different sampling distributions, they both reject H_0 at the same level of significance. If applied to quantitative data that are meeting all the requirements of Pearson r, both Spearman r and Kendall τ have power nearly as high (about 91%; Hotelling & Pabst, 1936) as their parametric equivalent. In all other cases, they are more powerful than Pearson r. This

Power

refers to the notion of *power* of statistical tests: a test is more powerful than another if it is more likely to detect small deviations from H_0 (i.e. smaller type II error), for constant type I error.

The chief advantage of Kendall τ over Spearman r, as already mentioned, is that it can be generalized to a partial correlation coefficient, which cannot be done with Spearman (Siegel, 1956: 214). The formula for a partial τ is:

$$\tau_{12.3} = \frac{\tau_{12} - \tau_{13}\tau_{23}}{\sqrt{1-\tau_{13}^2}\sqrt{1-\tau_{23}^2}} \quad \textbf{(5.9)}$$

This formula is algebraically the same as that of first-order partial Pearson r (eq. 4.36) although, according to Kendall (1948: 103), this would be merely coincidental because the two formulae are derived using entirely different approaches. The three τ coefficients on the right-hand side of eq. 5.9 may themselves be partial τ's, thus allowing one to control for more than one descriptor (i.e. high order partial correlation coefficients). Siegel & Castellan (1988) give tables for testing the significance of the Kendall partial correlation coefficient.

Rank correlation coefficients should not be computed in the Q mode, i.e. for comparing objects instead of descriptors; see Box 7.1, Chapter 7.

5.4 Coefficient of concordance

Kendall coeff. of concordance

The rank correlation coefficients described in the previous section measure the correlation between two descriptors for n objects. Kendall's coefficient of concordance W (Kendall & Babington Smith, 1939) measures the agreement among several (p) quantitative or semiquantitative variables over a set of n objects. In community ecology, the p variables may be species whose abundances are used to assess habitat quality at n study sites. In taxonomy, they may be p characters measured over n different species, biological populations, or individuals. In the social sciences, the variables are often p "judges" assessing n different subjects or situations.

Friedman's two-way ANOVA

There is a close relationship between Friedman's two-way analysis of variance without replication by ranks (Section 5.2) and Kendall's coefficient of concordance. Indeed, they both address hypotheses concerning the same data table and use the same statistic for testing. They only differ in the formulation of their respective null hypothesis. Consider Table 5.7, which contains illustrative data. In Friedman's test, the null hypothesis is that there is no real difference among the $n = 6$ objects because they pertain to the same statistical population. Under H_0, they should have received random ranks along the $p = 3$ variables, so that their sums of ranks should be approximately equal. Kendall's test focuses on the relationships among the $p = 3$ variables. If the null hypothesis of Friedman's test is true, this means that the variables have produced rankings of the objects that are independent of one another. This is the null hypothesis of Kendall's test of W.

Table 5.7 Numerical example. Ranks of six objects on three descriptors, $\mathbf{y}_1$, $\mathbf{y}_2$, and $\mathbf{y}_3$.

Objects (observation units)	Ranks of objects on the three descriptors			Row sums
	$\mathbf{y}_1$	$\mathbf{y}_2$	$\mathbf{y}_3$	R_i
$\mathbf{x}_1$	1	1	6	8
$\mathbf{x}_2$	6	5	3	14
$\mathbf{x}_3$	3	6	2	11
$\mathbf{x}_4$	2	4	5	11
$\mathbf{x}_5$	5	2	4	11
$\mathbf{x}_6$	4	3	1	8

1 — Computing Kendall W

The Kendall W coefficient is an estimate of the variance of the row sums of ranks R_i divided by the maximum possible value the variance can take; this occurs when all variables are in total agreement. Hence $0 \leq W \leq 1$, the value 1 representing perfect concordance. There are two ways of computing the Kendall W coefficient (i.e. either form of eq. 5.11); they lead to the same result. The computation proceeds in two steps.

Firstly, S or S' is computed from the row-marginal sums of ranks R_i received by the objects:

$$S = \sum_{i=1}^{n} (R_i - \bar{R})^2 \qquad \text{or} \qquad S' = SSR = \sum_{i=1}^{n} R_i^2 \qquad \textbf{(5.10)}$$

where S is a sum of squared deviations statistic over the row sums of ranks R_i and $\bar{R}$ is the mean of the R_i values. SSR designates the Sum of Squared R_i values.

Secondly, the Kendall W coefficient is obtained using either of the following formulas:

$$W = \frac{12S}{p^2 (n^3 - n) - pT} \qquad \text{or} \qquad W = \frac{12S' - 3p^2 n (n+1)^2}{p^2 (n^3 - n) - pT} \qquad \textbf{(5.11)}$$

where n is the number of objects and p the number of variables. To derive these formulas, one has to know that the sum of all ranks in the data table is $pn(n + 1)/2$ and

that, when all judgement vectors $\mathbf{y}_j$ are identical, the sum of R_i^2 is $p^2n(n+1)(2n+1)/6$. T is a correction factor for tied ranks (Siegel, 1956; Siegel & Castellan, 1988; Zar, 1999):

$$T = \sum_{k=1}^{g} (t_k^3 - t_k) \qquad \textbf{(5.12)}$$

in which t_k is the number of tied ranks in each (k) of g groups of ties. The sum is computed over all groups of ties found in all p variables of the data table. $T = 0$ when there are no tied values.

There is a close relationship between the Spearman r_S correlation coefficient and the Kendall W coefficient: W can be directly calculated from the mean ($\bar{r}_S$) of the pairwise Spearman correlations r_S using the following relationship (Siegel and Castellan, 1988; Zar, 1999):

$$W = \frac{(p-1)\,\bar{r}_S + 1}{p} \qquad \textbf{(5.13)}$$

where p is the number of variables among which the pairwise Spearman correlations are computed. Equation 5.13 is strictly true for untied observations only; for tied observations, ties are handled in a bivariate way in each Spearman r_S coefficient whereas in Kendall W the correction for ties is computed over all variables (eq. 5.12). For two variables only, W is simply a linear transformation of r_S: $W = (r_S + 1)/2$. In that case, a permutation test of W for two variables is the exact equivalent of a permutation test of r_S for the same variables.

The relationship described by eq. 5.13 clearly shows that W will consider p variables to be concordant only if their Spearman correlations are positive. Two variables that give perfectly opposite ranks to a set of objects have a Spearman correlation of –1, hence $W = 0$ for these two variables (eq. 5.13); this is the lower bound of the coefficient of concordance. For two variables only, $r_S = 0$ gives $W = 0.5$; for a group of p uncorrelated variables, $W = 1/p$. So coefficient W applies well to rankings given by a panel of "judges" called in to assess overall performance in sports, quality of wines, or food in restaurants, to rankings obtained from criteria used in quality tests of appliances or services by consumer organizations, or to the study of species associations in multi-species communities. It does not apply to variables used in multivariate analysis where negative as well as positive relationships are informative. Zar (1999), for example, used wing length, tail length and bill length of birds to illustrate the use of the coefficient of concordance. These data are appropriate for W because they are all related to the same common property, the size of the birds.

Numerical example. The calculation of Kendall's coefficient of concordance is illustrated using the numerical example of Table 5.7. The data could be semiquantitative rank scores, or quantitative descriptors coded into ranks. It is important to note that the $n = 6$ objects are ranked on each descriptor (column) separately. The last column gives, for each object i, the sum R_i of its

ranks on the $p = 3$ descriptors. The sum of squared deviations from the mean, $\sum (R_i - \bar{R})^2$ (eq. 5.10 left), is equal to 25.5 for this example. The W-statistic is calculated with eq. 5.11 (left):

$$W = \frac{12 \times 25.5}{9\,(216 - 6)} = 0.1619$$

There are no tied ranks in this example. The F and X^2 (chi-square) statistics are computed as follows (eqs. 5.14 and 5.15, next subsection):

$$F = \frac{(3 - 1) \times 0.1619}{(1 - 0.1619)} = 0.386$$

$$X^2 = 3 \times (6 - 1) \times 0.1619$$

The p-value associated with the F-statistic, found using the F-distribution, is 0.825. The permutational p-value after 999 random permutations within the variables is 0.835. The hypothesis (H_0) that the row sums R_i of Table 5.7 are equal cannot be rejected. The conclusion is that the 3 descriptors differ in the way they rank the 6 objects.

2 — Testing the significance of W

The recommended method for testing the significance of W is to compute the following F-statistic:

$$F = \frac{(p - 1)\,W}{(1 - W)} \qquad \textbf{(5.14)}$$

which is asymptotically distributed like F with $\nu_1 = n - 1 - (2/p)$ and $\nu_2 = \nu_1 (p - 1)$ degrees of freedom (Kendall & Babington Smith, 1939). Numerical simulations showed that this F-statistic had correct levels of type I error for any value of n and p (Legendre, 2010). It is unfortunate that this statistic has been overlooked by authors of recent textbooks on nonparametric statistics who recommend testing the significance of W with Friedman's (1937) X^2-statistic, which is obtained from W as follows:

$$X^2 = p(n - 1)\,W \qquad \textbf{(5.15)}$$

This X^2 (chi-square) statistic is asymptotically distributed like χ^2 with $\nu = (n - 1)$ degrees of freedom. Kendall & Babington Smith (1939) considered this test of W to be satisfactory for moderately large values of p and n only, not for small p. This was confirmed by simulations reported by Legendre (2005), who recommended not to use the theoretical χ^2-distribution to test X^2 when $p < 20$. The X^2-statistic can, however, be tested by permutation.

Permutation test

Permutation tests can be used with all combinations of values of p and n (Legendre, 2005). For the global test of significance, the rank values in all variables are permuted at random, independently over each variable, because the null hypothesis is the independence of the rankings produced by the p variables. The alternative hypothesis (H_1) is that at least one of the variables is concordant with one or more of

the other variables; so when H_0 is rejected, one cannot conclude that all variables are concordant with one another, but only that at least one variable is concordant with one or more of the others. Actually, for permutation testing, the four statistics *SSR* (eq. 5.10), *W* (eq. 5.11), *F* (eq. 5.14), and X^2 (eq. 5.15) are monotonic to one another since *n*, *p* and *T* are constant within a given permutation test; they are thus equivalent statistics for testing since they produce the same permutational probabilities. The test is one-tailed because it only recognizes positive associations between the ranked variables.

Many of the problems subjected to Kendall's concordance analysis involve fewer than 20 variables: the parametric χ^2-test should be avoided in these cases. The *F*-test (eq. 5.14) and the permutation test can be safely used with all values of *p* and *n*.

3 — Contributions of individual variables to Kendall's concordance

The contribution of individual variables (e.g. the *p* species) to the *W*-statistic can be assessed by a permutation test proposed by Legendre (2005). The null hypothesis is the monotonic independence of the variable subjected to the test with respect to all other variables in the group under study. The alternative hypothesis is that this variable is positively correlated with one or several other variables in the set under study (one-tailed test). The statistic *W* can be used directly in *a posteriori* permutation tests; alternatively, one can use two other statistics described in Legendre (2005) that are equivalent to *W* for *a posteriori* tests. Contrary to the global test, only the variable under test (e.g. one of the *p* species) is permuted here. If that variable has values that are monotonically independent of the other variables, permuting its values at random should have little influence on the *W*-statistic. If on the contrary it is concordant with one or several other variables, permuting its values at random should break the concordance and induce a noticeable decrease of *W*.

Concordance analysis is applied to the identification of species associations in Subsection 8.9.2, where an ecological application (mite data) is presented. Another example (fish associations) is found in Section 4.10.2 of Borcard *et al*. (2011).

Concordance analysis is also useful in phylogenetic analysis: prior to phylogenetic reconstruction, the degree of congruence among distance matrices (CADM) corresponding to different types of data or different genes can be tested using a test of significance proposed by Legendre & Lapointe (2004). The distance matrices under comparison are strung out like the descriptors in Table 5.7. The coefficient of concordance (*W*, eq. 5.11) is computed, then tested using the same permutation procedure as in the Mantel test (Subsection 10.5.1). The CADM test is actually a generalization of the Mantel test of correspondence between two distance matrices to any number of distance matrices. It can be used to compare distance matrices computed from evolutionary data (genetic congruence), the topologies of phylogenetic trees derived from these data (topological congruence), or the full phylogenetic trees

including topologies and branch lengths (phylogenetic congruence) (Campbell *et al*., 2011). Applications of this method are found in Campbell *et al*. (2009, 2011).

5.5 Software

All major commercial statistical packages allow the calculation of rank correlation coefficients, as well a choice of the methods listed in Table 5.2. In the R language,

1. Methods listed in Table 5.2 are available in the following functions of the STATS package: ***t.test()*** (*t*-test for independent and related samples), ***aov()*** (different forms of ANOVA), ***wilcox.test()*** (Mann-Whitney and Wilcoxon tests), ***kruskal.test()*** (Kruskal-Wallis test), ***friedman.test()*** (Friedman test), ***chisq.test()*** (chi-square test), ***fisher.test()*** (Fisher exact probability test), and ***mcnemar.test()*** (McNemar test). ***chisq.test()*** and ***fisher.test()*** offer permutation tests among their options. Rank correlation coefficients are available as options in function ***cor()*** of the STATS package, which can also be used to compute correlation matrices among several descriptors.

2. Logistic regression can be computed using the ***glm()*** function of the STATS package. Multinomial logistic regression is computed by function ***mlogit()*** of the MLOGIT package.

3. The global coefficient of concordance and *a posteriori* tests are available in functions ***kendall.global()*** and ***kendall.post()*** of VEGAN. Congruence among distance matrices is available in functions ***CADM.global()*** and ***CADM.post()*** of APE.