

Kendall coefficient of concordance: global test and *a posteriori* tests of individual judges – Program Kendall_W User's Guide

Pierre Legendre
Département de sciences biologiques
Université de Montréal
C.P. 6128, succursale Centre-ville
Montréal, Québec H3C 3J7, Canada

November 2004

E-mail: Pierre.Legendre@umontreal.ca

What does program KENDALL_W do?

This program computes a test of the concordance among several judges (which may be species in ecology) using Kendall's W statistic. The search for species associations is one of the classical problems of community ecology. The program also tests the contributions of individual judges (species) to the overall W statistic, as described in Legendre (2005).

The global and *a posteriori* tests are conducted by permutation. Simulations reported in the paper show that in the classical χ^2 test, type I error is too low when the number of judges is smaller than 20, leading to tests that are overly conservative and, thus, have reduced power. The permutation test used in the program has a correct level of type I error for all numbers of judges and observations.

Principle of the test:

H_0 : the p judges (e.g., species) produced independent rankings of the objects.

H_1 : the judges are not all independent of one another. At least one of the judges is concordant with one, or some of the other judges.

1. Read in the scores given by the judges (species in ecology). Write them in the rows of a work table.
2. Transform the scores from each judge into ranks, taking tied ranks into account.
3. Compute Kendall's coefficient of concordance, W .
4. Permute each judge at random, independently of one another, and compute a value W^* .
5. Repeat step 4 a large number of times.
6. Add the reference value of W to the distribution of W^* values. Calculate the one-tailed permutational probability associated to W .

Within a given permutation test, the three statistics W , χ^2 , and sum of squared ranks (SSR), are monotonic to one another. Since they produce the same permutational probability, they are equivalent for permutation testing. W is also linearly related to the mean of the Spearman r statistics among all pairs of judges (Siegel 1956: 232).

The global test is followed by tests of significance in which each judge, in turn, is permuted (H_0 : independence of a given judge with respect to all other judges).

Procedure for the identification of associated species

When there are two or more groups of judges, the global test has high power but the *a posteriori* tests do not allow distinguishing the members of the groups. When there judges are negatively correlated among the groups, power of the *a posteriori* tests decreases. The recommended procedure for the identification of groups of judges (e.g., species associations) is the following:

1. First, conduct an overall test of concordance using all judges (species).
2. If that test is significant, look for groups of correlated judges (species). Several analytical strategies, described in the paper, are possible.
3. Submit each group of judges (species) to a separate analysis of concordance. The *a posteriori* tests will identify the judges (species) that are significantly associated.
4. Compute probabilities adjusted over all judges (species). Write all the unadjusted probabilities obtained for the different group of judges to a single data file. Use that file as input for the program ADJUSTED P-VALUES available on our Web page. See that program's user's guide (Legendre 2002).

Input file

There is a single input data file. It is an ASCII (i.e., text) file without row or column identifiers. The judges may correspond to the rows or columns of the input data table.

Important note: the input file must have ends-of-lines corresponding to the operating system. This is important for Macintosh users who use MacOS 9 and MacOS X concurrently: an input file used with the MacOS 9 version of the program must have Macintosh-type ends-of-lines whereas the same input file used with the MacOS X version of the program must have Unix-type ends-of-lines.

Output file

See example below.

Disclaimer

This program is provided without any explicit or implicit warranty of correct functioning. It has been developed as part of a university-based research program. If, however, you should encounter problems with this program, the author will be happy to help solve them. Researchers may use this program for scientific purposes, but the source code remains the property of Pierre Legendre. Users of the program may refer to the present user's manual as follows:

Legendre, P. 2004. Kendall coefficient of concordance: global test and *a posteriori* tests of individual judges – Program Kendall_W User's Guide. Département de sciences biologiques, Université de Montréal. 8 pages.

Technical notes

The program is available in a variety of forms:

- FORTRAN source code for Macintosh (file Kendall_W.f), which can be compiled using a FORTRAN compiler. The user may modify the Parameter statement at the beginning of the program, which fixes the size of the largest data matrix that can be analysed (nmax = maximum number of observations, jugemax = maximum number of judges).
- FORTRAN source code for Mac OS X and Windows (file KENDALL_W_G77.FOR), which can be compiled using the g77 FORTRAN compiler. The user may modify the Parameter statement at the beginning of the program, which fixes the size of the largest data matrix that can be analysed (nmax = maximum number of observations, jugemax = maximum number of judges).
- Compiled version for PowerPC processors for Macintosh OS X (file KENDALL_W). The maximum size of the data matrix is 2000 objects and 1000 judges.
- Compiled version for PowerPC processors for Macintosh Classic (file KENDALL_W/PPC). The maximum size of the data matrix is 1000 objects and 500 judges. The program requires 12 Mb RAM for running.
- Compiled version for IBM compatible PC (file KENDALL_W.exe). The maximum size of the data matrix is 2000 objects and 1000 judges. The executable file is a Win32 “console” executable, not a DOS executable. Therefore it cannot run under plain DOS, nor in a DOS window under Windows 3.x. It can only run in Windows 95/98, Windows NT, or Windows XP consoles.

References

Legendre, P. 2002. Adjusted P-values user's guide. Département de sciences biologiques, Université de Montréal. 4 pages. Available at <<http://www.bio.umontreal.ca/legendre/>>

Legendre, P. 2005. Species associations: the Kendall coefficient of concordance revisited. *Journal of Agricultural, Biological and Environmental Statistics* **10**: 226-245.

A preprint is available on the WWW site <<http://www.bio.umontreal.ca/legendre/>>. Select “Tirés-à-part / Reprints” in the left-hand column of the introduction page. You will find the title of the paper in the “Tirés-à-part / Reprints” page. Click on the title and you will receive a copy of the preprint in Adobe Acrobat PDF format.

Siegel, S. 1956. Nonparametric statistics for the behavioral sciences. McGraw-Hill Series in Psychology, McGraw-Hill, New York.

Appendix: Test run

This example is used as illustration of the method in section 9 of the paper (Legendre 2005). The **input file** “Mites(10x4)”, without row or column identifiers, is the following (also shown in Table 1):

```
-----
0.25087  0.40538  0.24380  0.08362
0.40324  0.25503  0.39303  0.00000
0.26577  0.47620  0.27267  0.06097
0.32350  0.63337  0.47003  0.00000
0.26312  0.29089  0.39223  0.08771
0.33675  0.44836  0.53727  0.10153
0.07956  0.19487  0.19487  0.11251
0.00000  0.18570  0.26261  0.11744
0.00000  0.15430  0.15430  0.00000
0.12769  0.62987  0.27584  0.34578
```

Sites 4, 9, 14, 22, 31, 34, 45, 53, 61, 69, species 13, 14, 15, 23

The judges are the columns of the data file in this particular case. Only the first 10 data rows will be read by the program because the user declares the presence of 10 objects. User's notes can be added at the bottom of the file, as in this example. The program dialogue is the following (user's answers in **bold**):

```
Français: tapez (1)
English:  type  (2)
2

Kendall's test of concordance

Pierre Legendre
Département de sciences biologiques
Université de Montréal.
    © Pierre Legendre, 2000, 2004

How many judges and objects?
4 10

(1) Judges are the rows of the input data file
(2) Judges are the columns of the input data file
2

Name of the input data file?
Input data file: Mites(10x4)

Test (1) with or (2) without correction for ties?
1

How many permutations? (e.g., 999, 9999, ...)
9999

Computation time:      1.05 sec.

Results are found in file 'Concord.out'
End of the program.
```

Output file

The output file ("Concord.out") contains the following information:

Kendall's test of concordance

Pierre Legendre
Département de sciences biologiques
Université de Montréal.
© Pierre Legendre, 2000, 2004

Input data file: Mite subset (10x4)

Number of permutations: 9999

Friedman's Chi-square = 15.89771 nu = 9
Kendall's W = 0.44160 with correction for ties

Global test. H0: Independence of all judges.
All judges are independently permuted.

Prob(chi-square, 9999 perm.) = 0.04480

A posteriori test for each judge.
Results are presented in order of decreasing values of Rbar.
H0: Independence of this judge. This judge only is permuted.

Judge	Rbar(j)	W(j)	Adjusted probability		
			Prob	Bonferroni	adj.Holm
3	0.45704	0.59278	0.00510	0.02040	0.02040
2	0.39655	0.54741	0.02400	0.09600	0.07200
1	0.32657	0.49493	0.07660	0.30640	0.15320
4	-0.16813	0.12391	0.70700	2.82800	0.70700

In the output table, the judges are represented by their input order numbers (1 to 4 in this example), and presented in order of decreasing values of the Rbar(j) statistic (\bar{r}_j in Table 2, where the judges are not presented in order of decreasing values of Rbar(j), but in their input order). The judge input order numbers can be replaced by hand, in the table, by the real judge names.

Table 1. Illustrative example. Columns 2-5: Hellinger-transformed abundances of four mite species at 10 sites selected along the long axis of Figure 5 of the Legendre (2005) paper. The Hellinger transformation was computed for the full data set (70 sites). Columns 6-9: the same data transformed into ranks (with ties). Column 10: sum of the ranks for each site.

	Hellinger-transformed abundances				Ranks (species-wise)				Sum of ranks
	Species 13	Species 14	Species 15	Species 23	Species 13	Species 14	Species 15	Species 23	R_i
Site 4	0.25087	0.40538	0.24380	0.08362	5	6	3	5	19.0
Site 9	0.40324	0.25503	0.39303	0.00000	10	4	8	2	24.0
Site 14	0.26577	0.47620	0.27267	0.06097	7	8	5	4	24.0
Site 22	0.32350	0.63337	0.47003	0.00000	8	10	9	2	29.0
Site 31	0.26312	0.29089	0.39223	0.08771	6	5	7	6	24.0
Site 34	0.33675	0.44836	0.53727	0.10153	9	7	10	7	33.0
Site 45	0.07956	0.19487	0.19487	0.11251	3	3	2	8	16.0
Site 53	0.00000	0.18570	0.26261	0.11744	1.5	2	4	9	16.5
Site 61	0.00000	0.15430	0.15430	0.00000	1.5	1	1	2	5.5
Site 69	0.12769	0.62987	0.27584	0.34578	4	9	6	10	29.0

Appendix – Unix and Windows user’s notes

The Unix (MacOS X) and DOS versions of this program were built with g77, the GNU FORTRAN compiler.

Click on the program’s icon to start it, then answer the questions that appear on the screen. You will be asked to provide the name of the file containing the data. In Windows, make sure the file name is DOS-compatible (maximum 8 characters, no space, etc.). If the input file has an extension, like “txt”, you have to type it when you provide the name of the input file, even if the extension does not appear in the file list.

In Unix and Windows, the file “Concord.out” produced by the program cannot be deleted by the FORTRAN program during a following run. If, after launching, the program ends abruptly and a message is displayed, such as:

In Mac OS X: open: 'new' file exists
 apparent state: unit 7 named Concord.out
 last format: list io
 lately reading direct formatted external IO
 Abort trap
 logout

In Windows: Abnormal program termination following which the console window shuts off,
 this means that a file called “Concord.out” already exists in the current directory. Rename or remove that file before running the program again.