

4.11 Example: Associations between species

It is instructive to consider some additional examples of the test of association where the variables are not evenly distributed. Specifically, we wish to cater for situations where the variables of interest are occurrences, densities or counts of species' abundances, as commonly encountered in community ecology.

In such cases, we should consider using the *index of association* ($I_{\{A\}}$ or $I_{\{A\}}^*$) as our measure, because, in the majority of cases, there can be no sense in including joint absences (i.e., where both species take a value of zero, corresponding to sites where neither species occurs) in the consideration of whether the individuals of the two species are likely to be found together or not. Sites that contain neither species at all are simply non-informative in this regard.

Abra prismatica & *Goniada maculata*

From the 'Ekofisk macrofauna counts' dataset (see the previous page for how to access this data set), choose **Analyse > Univariate > Association...** and choose the bivalve, *Abra prismatica*, as 'Variable 1' and the polychaete *Goniada maculata* as 'Variable 2' (leaving defaults for the rest), then click **OK**.

Test of Association

Worksheet 1 (active):
Ekofisk_macrofauna_counts

Worksheet 2:
Ekofisk_macrofauna_counts

Measure of association:
Index of Association

Max permutations:
9999

Alternative hypothesis:
☒ Positive association
☐ Negative association
☐ Either a positive or negative association (two-tailed)

Variable 1:
Abra prismatica

Variable 2:
Goniada maculata

☐ Within levels of a factor
Dist

Output:
☐ Results to worksheet
☒ Output scatter plot
 Output values of test-statistic under permutation:
☐ to worksheet
☒ to graph (histogram)

OK Cancel Help

The output file shows a statistically significant positive association between these two variables; the index of association is $I_{\{tiny\{A\}} = 80.3$ and $P < 0.01$.

Test of association3

Test of association
Test the association between two variables

Data worksheet
Name: Ekofisk_macrofauna_counts
Data type: Abundance
Sample selection: All
Variable selection: All

Parameters
Worksheet 1: Ekofisk_macrofauna_counts
Variable 1: Abra prismatica
Worksheet 2: Ekofisk_macrofauna_counts
Variable 2: Goniada maculata
Measure of association: Index of Association (IoA)
Separate test within levels of a secondary factor: False
Max permutations (excluding observed statistic): 9999
Results to worksheet: False
Output scatter plot: True
Output distribution to worksheet: False
Output distribution to graph: True

Null hypothesis: There is no association (or a negative association).
Alternative hypothesis: There is a positive association (one-tailed).

Results

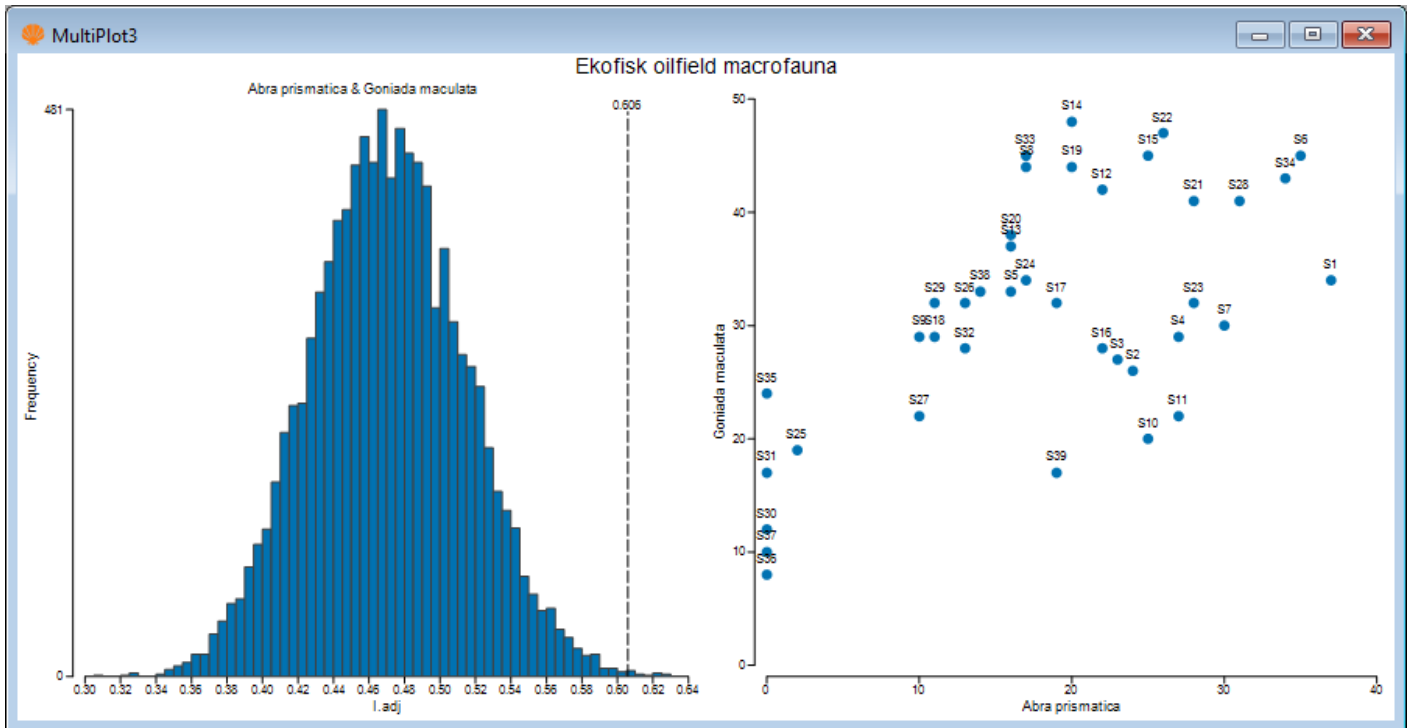
	Statistic (I.adj)	Statistic (IoA)	P value	Possible Permutations Large*	Actual Permutations	Unique Values of I.adj.perm	Num >= observed	Num Tied Data Values	Total
Abra prismatica & Goniada maculata	0.6059	80.2959	0.0014		9999	7089	13	66	7

*greater than 100,000,000

Note:
The test statistic used here is the adjusted index of association (I.adj) which permits either positive or negative relationships to be expressed (scaled from -1 to +1). $I_{adj} = (2 * IoA / 100) - 1$

Outputs
Plot: Graph5
Plot: Graph6

This association is also evident visually in the scatter plot:

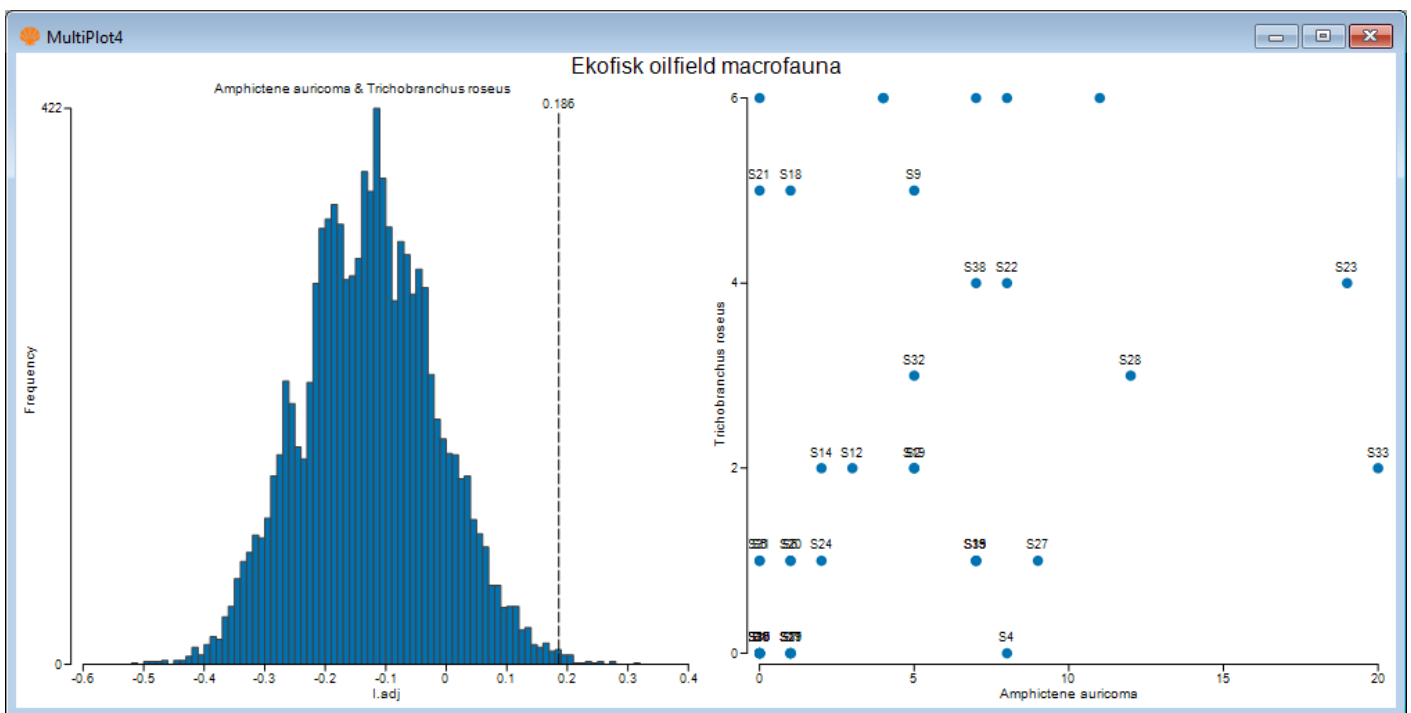
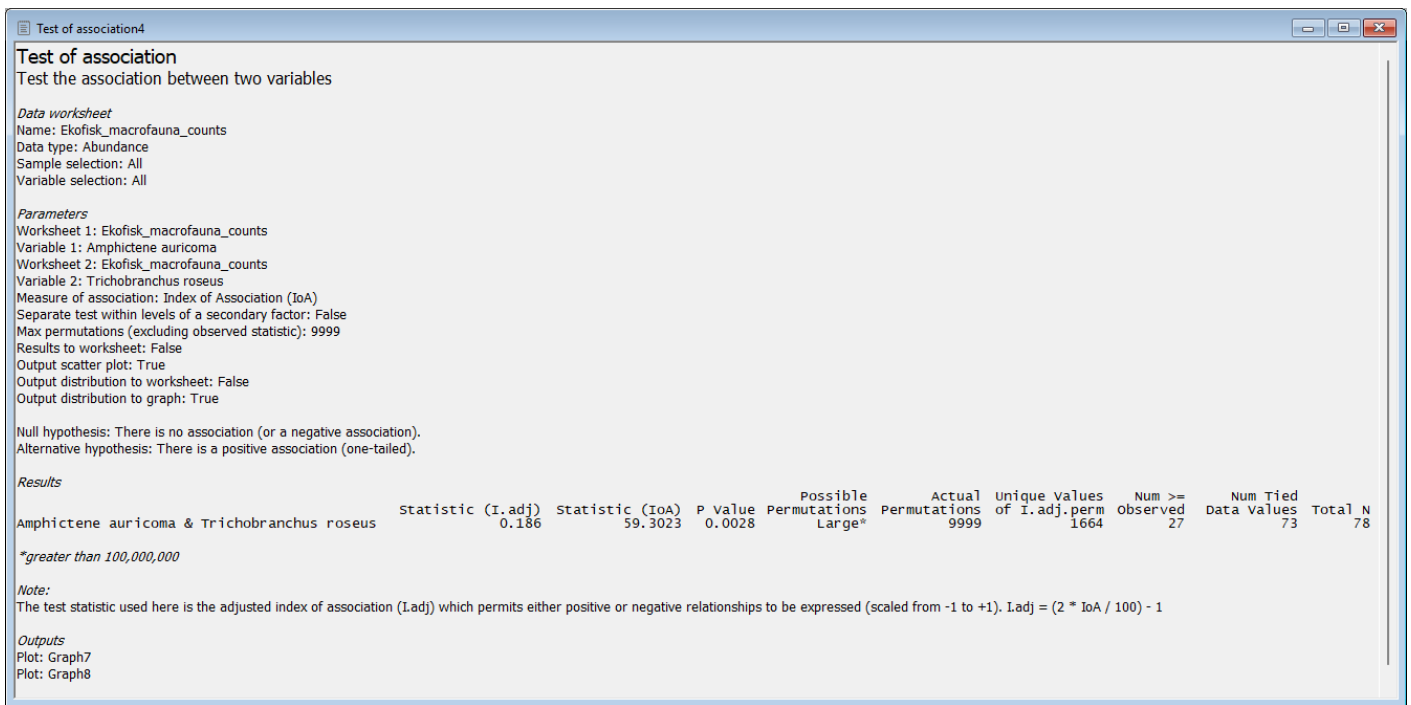


Note that *standardising* the variables first (e.g., by their totals), as would be desirable here, is not necessary to do as a separate step, and would actually have no effect, as that operation is done automatically as part of the calculation of the index itself (see [Somerfield & Clarke \(2013\)](#)).

Amphictene auricoma & Trichobranchus roseus

Another example demonstrates how joint absences (double zeros) across the samples can occasionally yield counter-intuitive results when examining scatter plots. Let's run the test of association on the following two species of polychaete worms: *Amphictene auricoma* & *Trichobranchus roseus*.

The output file and graphics (based on the index of association) are shown below:



Despite having detected a significant positive association between these two species ($I_{\{A\}} = 59.3$ and $P < 0.01$), a classic pattern of positive correlation is not particularly obvious in the scatter plot. For these two species, more than a quarter of the data values are equal to zero.[¶] We can, however, trust the outcome of the test, but this example serves to show how a raw scatter plot of all joint values (including the joint absences) may not assist us in ascertaining the nature of species' co-occurrence relationships. Other types of plots may be helpful in clarifying similarity in the patterns of species across multiple samples (e.g., boxplots, means plots, line plots, coherence plots, etc.).

Abra prismatica & Chaetozone setosa

Another example demonstrates how the test of association works in the case of a negative association. Consider the following two species: *Abra prismatica* & *Chaetozone setosa*. *A. prismatica* is a bivalve mollusc that can be negatively affected by contaminants in the field, while *C. setosa* is an opportunistic species that can flourish in polluted areas.

Running the test of association on this pair of species is done under the alternative hypothesis of there being a *negative association* between them, like so:

Test of Association

Worksheet 1 (active):
Ekofisk_macrofauna_counts

Worksheet 2:
Ekofisk_macrofauna_counts

Measure of association:
Index of Association

Max permutations:
9999

Alternative hypothesis:

- ☐ Positive association
- ☒ Negative association
- ☐ Either a positive or negative association (two-tailed)

Variable 1:
Abra prismatica

Variable 2:
Chaetozone setosa

☐ Within levels of a factor
Dist

Output:

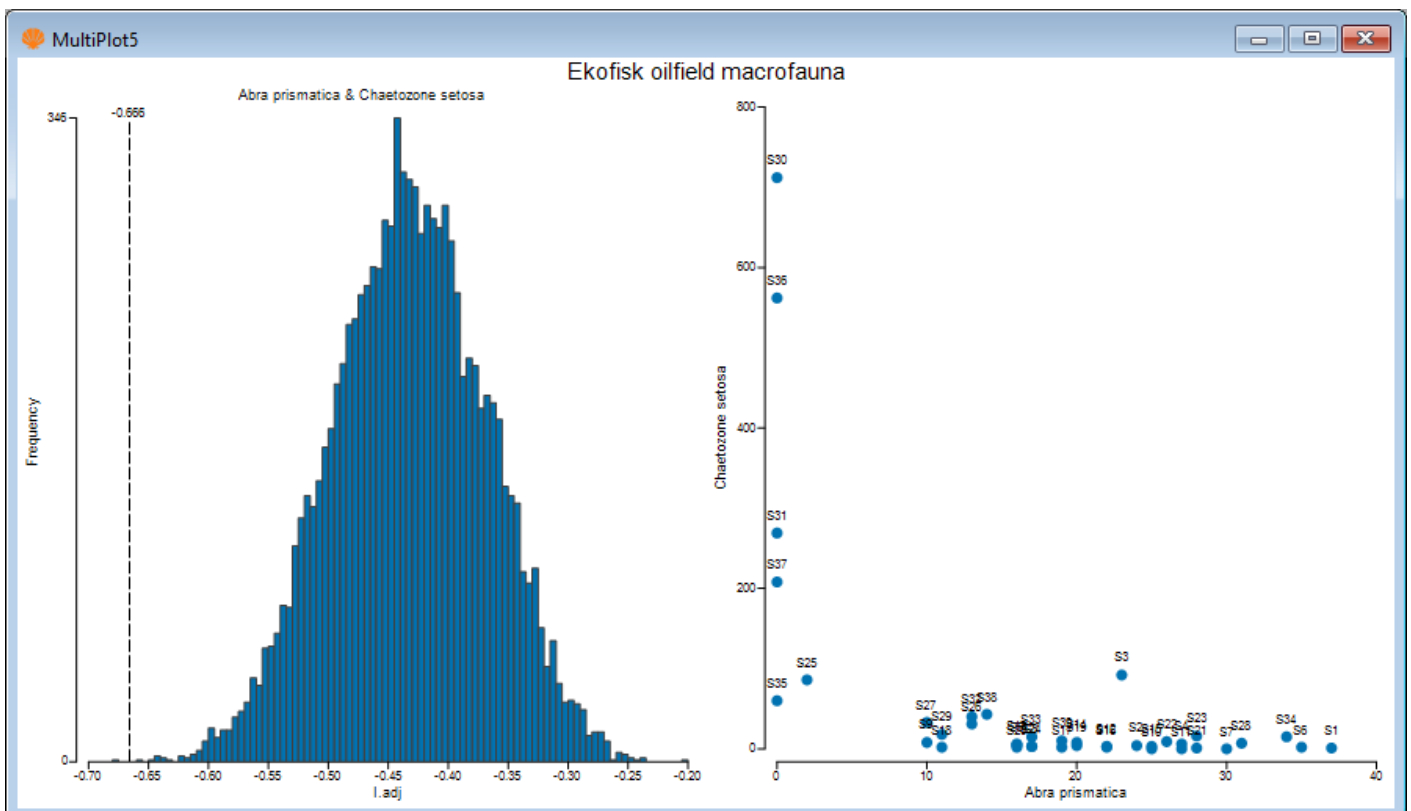
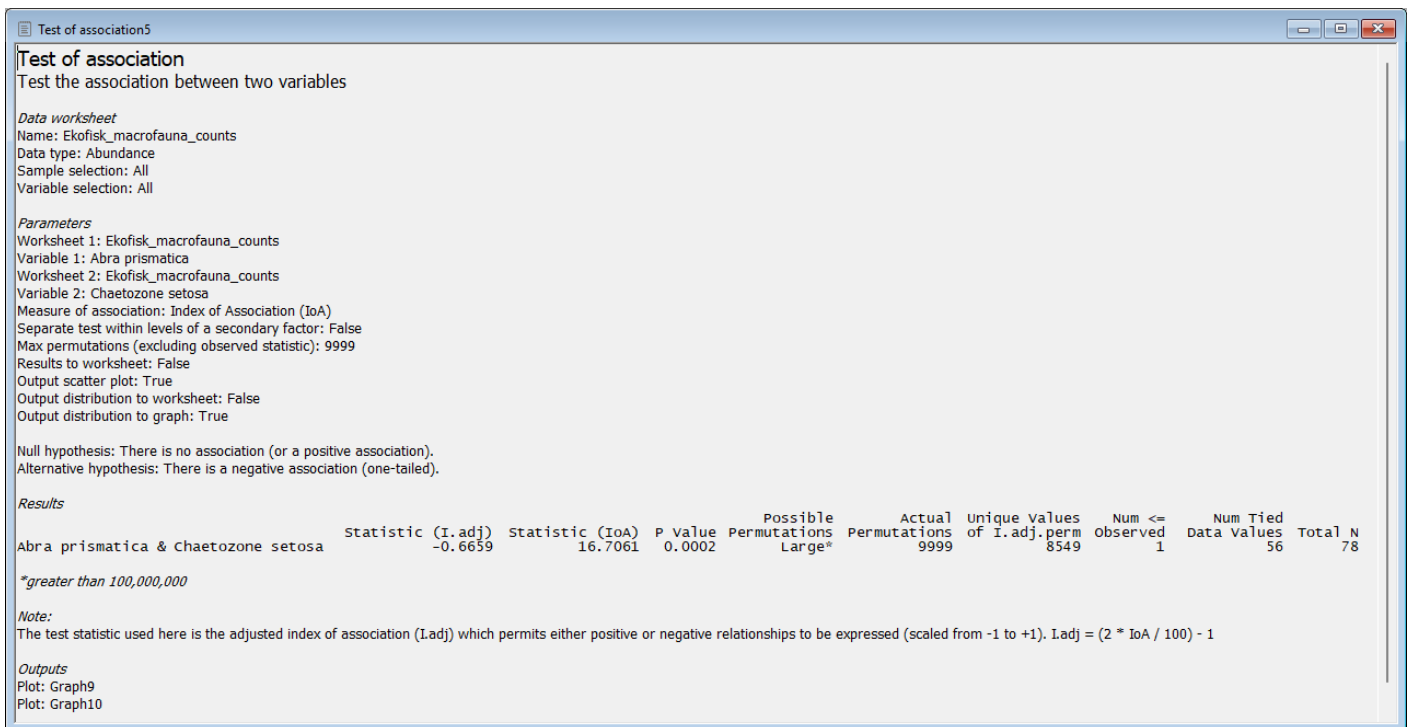
- ☐ Results to worksheet
- ☒ Output scatter plot

Output values of test-statistic under permutation:

- ☐ to worksheet
- ☒ to graph (histogram)

OK Cancel Help

The results are shown below:



In the past, we would have looked at the (quite low) value for the index of association ($I_{\text{tiny}\{A\}} = 16.7$), but there would be no obvious way of asserting any particular statistical significance to this, one way or the other. Now, with the new tool for testing associations in PRIMER 8, we can use our adjusted index of association value as the test-statistic ($I_{\text{tiny}\{A\}}^* = -0.67$) and, under permutation, it is clear that this negative association is highly statistically significant ($P = 0.0001$). In other words, for this dataset, where you find one of these species, you do not tend to find the other one, and *vice versa*.

Note that, in all three of the above tests, the distribution of the index of association under permutation is not centred on zero, nor is it necessarily symmetric; yet, in all three cases, it is easy to examine the output provided (consisting of the empirical permutation distribution and the observed value of the test statistic relative to it) in order to ascertain the appropriate alternative hypothesis for any given test.

¶ *We can very quickly calculate the number of zeros (and the number of non-zeros) for any variable(s) in any dataset by running the new 'Tools > Summary Stats...' in PRIMER 8.*

Revision #46

Created 23 March 2025 22:42:22 by Marti

Updated 2 December 2025 03:30:32 by Marti