

VARIANCE TESTS FOR SPECIES ASSOCIATION

BU-838-M

by

Revised
November, 1984

Charles E. McCulloch

Biometrics Unit, Cornell University, Ithaca, New York 14853

Abstract: Schluter (1984) discusses species association tests to handle the situation where the data consist of a two-way table (species by samples) of either presence/absence or density data. In this paper, connections between the association tests and standard statistical tests are established and used as a guide to proper interpretation of the association tests. Association tests are sensitive to purely sample-to-sample differences and in some cases will not reflect species interactions at all. However, in cases where the association tests are appropriate, the connections furnish information on the choice of accurate critical values.

Key words: variance test, species association, chi-square, Cochran's Q test, F test.

Key phrases: tests for species association, critical values for species association tests, interpreting species association measures, comparing observed and expected variance.

1. Introduction

Schluter (1984) proposed variance tests for detecting species association with either presence/absence data or density data. These tests were, respectively, a modification and generalization of a test originally suggested in Pielou (1972) and Robson (1972). In Section 2 it is shown that the proposed tests are directly related to standard statistical tests. The association measure, W , used in the association tests is, for presence/absence data, a multiple of Cochran's Q (Cochran, 1950; see also Conover, 1980, pp. 199-205) and, for density data, a simple function of an F -statistic from a two-way analysis of variance of species by samples. In addition, the association measure can be viewed as an average of the pairwise correlation between species.

In Section 3, the connection between the association measure and standard statistical tests is used to provide an interpretation of the measure as sensitive to both sample-to-sample variation and associations between species. This is contrasted with the approach suggested in Pielou (1972) and Robson (1972). Sections 4 and 5 discuss the use of the association measure for testing hypotheses. Section 4 concerns the appropriateness of such tests and Section 5 proposes improved critical values.

2. Equivalence Relations

Following Schluter's (1984) notation, let X_{ij} denote the measurement (presence/absence or density) for the i th species in the j th sample. ($i = 1, 2, \dots, M$, $j = 1, 2, \dots, N$). Also, let T_j denote the number or total density of species for the j th sample $\left(T_j = \sum_{i=1}^M X_{ij}\right)$ and let n_i denote the number or total density for the i th species $\left(n_i = \sum_{j=1}^N X_{ij}\right)$. The association test proposed for presence/absence data is based on the association measure

$$V = S_T^2 / \Sigma \hat{\sigma}_i^2 ,$$

where

$$S_T^2 = \frac{1}{N} \sum_{j=1}^N (T_j - \bar{T}.)^2 ,$$

$$\hat{\sigma}_i^2 = \frac{n_i}{N} \left(1 - \frac{n_i}{N} \right) ,$$

$$\bar{T} . = \frac{1}{N} \sum_{j=1}^N T_j ,$$

or equivalently on

$$W = N \cdot V .$$

The test statistic compares the observed variance, S_T^2 , with an estimate of the variance expected under the hypothesis of no association, $\Sigma \hat{\sigma}_i^2$. This will be sensitive to positive or negative associations. The denominator, $\Sigma \hat{\sigma}_i^2$, will always be an estimate of the total within species variance. However, the numerator is affected by two sources of variability. When there is no association it also estimates the total within species variance. When the species are positively associated, fluctuations in T_j will be accentuated, leading to an inflation of S_T^2 relative to the no association case. Under negative associations, changes in one species will be compensated for by one or more other species and T_j will be relatively stable, leading to small values of S_T^2 .

Cochran (1950) proposed a test for homogeneity of treatments using presence/absence data which has come to be called Cochran's Q test and is included in some introductory texts, such as Conover (1980). If used as a test of homogeneity of samples for species by samples data, the test statistic, Q, is given by

$$Q = \frac{\left(\frac{N-1}{N}\right) \sum_{j=1}^N \left(T_j - \bar{T}_{.}\right)^2}{\sum_{i=1}^M \frac{n_i}{N} \left(1 - \frac{n_i}{N}\right)}$$

$$= \left(\frac{N-1}{N}\right) W \quad . \quad (2.1)$$

Thus, W is a multiple of Cochran's Q statistic. The statistic in the exact form of W has also been proposed; see Leslie (1958) and Carothers (1971).

The association test for use with density data is a bit more complicated. Using the usual analysis of variance notation, i.e.,

$X_{.j}$ = total density for sample j

$$= \sum_{i=1}^M X_{ij} \quad ,$$

$\bar{X}_{.j}$ = average density for sample j

$$= X_{.j}/M = \sum_{i=1}^M X_{ij}/M \quad ,$$

$\bar{X}_{i.}$ = average density of species i

$$= \sum_{j=1}^N X_{ij}/N \quad ,$$

$X_{..}$ = total density

$$= \sum_{j=1}^N \sum_{i=1}^M X_{ij} \quad ,$$

$\bar{X}_{..}$ = average density

$$= X_{..}/MN$$

W can be written as

$$W = N \cdot S_T^2 / \sum_i \sigma_i^2$$

$$= \sum_{j=1}^N \left(\bar{X}_{.j} - \bar{X}_{..}/N \right)^2 / \sum_{i=1}^M \frac{1}{N} \sum_{j=1}^N \left(X_{ij} - \bar{X}_{i.} \right)^2 .$$

This can be manipulated (see Appendix 1) to show that

$$W = N \cdot \frac{\frac{M}{M-1} F}{\frac{F}{M-1} + 1} , \quad (2.2)$$

where

F = the F-statistic from a two-way analysis of variance (Species by Samples), with (N-1) and (N-1)(M-1) degrees of freedom, used for testing Samples with the Species x Samples mean square as an error term

$$= \frac{\sum_{j=1}^N M \left(\bar{X}_{.j} - \bar{X}_{..} \right)^2 / (N-1)}{\sum_{i=1}^M \sum_{j=1}^N \left(X_{ij} - \bar{X}_{i.} - \bar{X}_{.j} + \bar{X}_{..} \right)^2 / (N-1)(M-1)} . \quad (2.3)$$

The link between the association measure and the F-statistic as exhibited in (2.2), or the equivalent formula for V,

$$V = \frac{\frac{M}{M-1} F}{\frac{F}{M-1} + 1} \quad (2.4)$$

can be used to obtain the following results:

1. The maximum value of V is M, the number of species. (V is approximately equal to M when F is large).
2. V equal to 1 ("no association") is equivalent to F equal to 1.
3. V greater than 1 is equivalent to F greater than 1 and, correspondingly, V less than 1 is equivalent to F less than 1.

In addition to the above equivalences, the association measure is a kind of average of the pairwise correlations between species. One way to calculate an average correlation, \bar{r} , is to calculate an average covariance term and divide that by an average variance term. Defining \bar{r} in

this way yields the equation

$$\bar{r} = \frac{\frac{1}{M(M-1)} \sum_{i=1}^M \sum_{i' \neq i}^M \frac{1}{N-1} \sum_{j=1}^N (x_{ij} - \bar{x}_{i.})(x_{i'j} - \bar{x}_{i'.})}{\frac{1}{M} \sum_{i=1}^M \frac{1}{N-1} \sum_{j=1}^N (x_{ij} - \bar{x}_{i.})^2}.$$

Straightforward algebraic manipulations show that

$$\bar{r} = \frac{V-1}{M-1}$$

or, this can be solved for V as

$$V = (M-1) \bar{r} + 1 \quad (2.5)$$

This establishes directly the connection between techniques which look at pairs of species and the overall association measure, V. The values of \bar{r} and V are related as follows:

1. \bar{r} ranges between $\frac{-1}{M-1}$ (when $V = 0$) and 1 (when $V = M$).
2. V equal to 1 is equivalent to $\bar{r} = 0$.
3. V greater than 1 is equivalent to \bar{r} greater than 0 and, correspondingly, V less than 1 is equivalent to \bar{r} less than 0.

3. Interpretations of the Association Measure

The first two correspondences derived in Section 2 (equations (2.1) and (2.2)) indicate that the association measure is sensitive to systematic sample-to-sample differences. This is true since both Cochran's Q and the F-statistic are designed to test for differences between samples. More precisely, Cochran's Q is used to test the null hypothesis

$$H_0 : p_{i1} = p_{i2} = \dots = p_{iN} \text{ for all } i,$$

where

p_{ij} = probability that species i is present in sample j .

The F-statistic is used to test the null hypothesis

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_N ,$$

where

μ_j = mean density of all species in sample j .

However, the last correspondence (equation (2.5)) indicates that the association statistic also measure average pairwise correlation. While average pairwise correlation seems like a readily interpretable quantity for inferring species association, the sensitivity of the statistic to sample-to-sample variation is worrisome since that will often have nothing to do with species interactions (as pointed out in Table 4 of Schluter (1984)) or even nothing at all to do with ecological processes!

Thus, samples can indicate species associations for three reasons:

1. Direct interactions between species (positive or negative).
2. Heterogeneous sampling techniques: If techniques vary in effectiveness similarly for two species, a positive association will be induced. If effectiveness varies differentially for two species, a negative association will be induced.
3. Selection of heterogeneous habitats: Similarities in selected habitats will give positive associations, differences will give negative associations.

In cases 2. and 3., interpretation of V as giving information about species interactions or any sort of ecological process is dubious. Unfortunately, such situations are common; some examples are listed below:

1. Different samples are taken using different techniques. For example, in Macan (1976) waterbugs were trapped using three different methods: bottom samples, samples slung by rope in midwater, and surface samples.
2. Differing sampling or censusing efforts are used. For example, Nilsson (1977) estimated his census effort to be 3 hours per hectare in mainland areas and 5 hours per hectare in island areas.
3. With density data, purely sample effects can be introduced if the "correct" calculation of density is not made. With different sized sites, what is the proper divisor of abundance to obtain density? For example, in Nilsson (1977) in a study on small islands, density was measured in territories per hectare. Perhaps territories per kilometer of shoreline or something even more complicated is a better measure.
4. In Wood (1974) study sites are selected along an altitudinal gradient. Soil and other variables change with altitude. He studied earthworms which are sensitive and therefore the large negative association must be interpreted as at least partially due to the intentional selection of sites with differing PH levels rather than species interaction. A random selection of sites would likely give a quite different value.
5. Petr (1972) in a study of benthic fauna, chose study sites to represent diverse habitats and gathered data over time as the area became flooded to form a man-made lake. As association measure calculated from such data would show a large positive association with little useful interpretation. Schulter (1984), in using part of Petr's data, wisely did not use the data for all years and sites.

The preceding is not intended to suggest that the studies cited were poorly conducted. It is to indicate that care must be taken to properly interpret the variance measure and that some studies are not set up to (nor are designed to) measure species interaction through a variance measure.

How can situations be identified in which the association measure can be interpreted as reflecting species association? Basically, a division can be made as to whether or not the samples can be regarded as randomly sampled from or representing the totality of measurements (all taken similarly) of a single community. Any differences between samples introduced by nonrandom sampling (selecting sites to be as diverse as possible or sampling different communities) or differing sampling effort or technique will distort the association measure, making it impossible to separate the effect due to ecological processes alone. The distortion can be in a positive or negative direction.

Contrast the above with the suggestions of Pielou (1972). She suggests using the measure within a single community, when the "samples" are actually sampled from the community. Furthermore, she suggests looking at the *change* in the association measure when a group of species are dropped from consideration. This avoids the potential distortion due to non-ecological processes and the resultant problems in interpretation brought about when it is necessary to compare the association measure to an absolute number (i.e., positive association if greater than 1). The distortion is eliminated since both times the association measure is calculated (before and after removal of a group of species) the purely sample-to-sample variation is the same.

4. Using the Association Measure as a Test Statistic

Beyond the question of proper interpretation of the association

measure is the question of when it is appropriate to *test* for association. Any statistical test of a hypothesis involves drawing conclusions about some set of circumstances broader than the ones actually studied (called the target population). Pielou (1972, p. 338) clearly intended to measure the interdependence within a community by using quadrats (samples) which were assumed to be "... independent of one another," e.g., randomly sampled from the community. In this situation, a statistical test would be appropriate if the quadrats could be justified as a random sample of quadrats in some target population (actual or conceptual). An ideal situation would be one in which sites are actually randomly sampled, but the assumption of random samples can be justified for other situations. The requirement of random samples is dubious in some of the situations for which Schluter (1984) has applied the association tests. In these cases hypothesis or significance tests may be inappropriate.

For example, Nilsson (1977) studied breeding birds on small islands in a single lake in Sweden and two mainland study plots. Since the islands were censused during the study period, they do not form the target population. Otherwise, no test would be necessary, since all circumstances forming the target population would have been measured. It is hard to envision even a conceptual population from which the islands can be regarded a random sample. They are restricted in geographical region, restricted in time and are unique in other ways. It does not make sense to test for the presence of association in a target population which we cannot even begin to define.

Another example is the situation of James and Boeklen (in press). A single study area was used and was censused each of seven consecutive years. Again, since a census was performed a hypothesis test is not necessary if the target population is that particular study site over the years studied.

Are the years themselves a random sample of a target population of years? Seven consecutive years are unlikely to be a good sample of years and James and Boecklen refrained from hypothesis or significance tests.

Note, however, that the James and Boecklen situation is a case where the association measure makes perfect sense as a descriptive statistic. The samples are from the same community and there is no obvious variation that is purely sample-to-sample. It just is not a good situation in which to perform statistical tests.

5. Critical Values for the Tests of Association

For situations in which a hypothesis test is appropriate, the connections made in Section 2 suggest improvements to the critical values. Also, previous work as to the accuracy of critical values can be called upon. Theoretical derivations of critical values are advantageous since the derivation of accurate critical values by simulation requires many replications.

Presence/absence data

For presence/absence data the connection with Cochran's Q test (equation (2.1)) suggests using the critical values $\frac{N-1}{N} \chi^2_{N-1, 1-\alpha/2}$ instead of $\chi^2_{N, 1-\alpha/2}$. In practice, there is not much difference between the two, however, Cochran (1950) shows that $\frac{N-1}{N} \chi^2_{N-1, 1-\alpha/2}$ tends to slightly underestimate the upper tail significance probabilities. Using $\chi^2_{N, 1-\alpha/2}$ would make the underestimation even worse. Thus, $\frac{N-1}{N} \chi^2_{N-1, 1-\alpha/2}$ should be an improvement. Tate and Brown (1970, p. 159) recommend the following rule of thumb for deciding if the $\frac{N-1}{N} \chi^2_{N-1, 1-\alpha/2}$ critical value is an adequate approximation:

Delete all species that appear in all the samples or none of the samples. If the product of the remaining number of species times the number of samples is 24 or more, the approximation is generally satisfactory, as long as the remaining number of species is at least 4.

Another interesting consequence of these connections embodied in this rule of thumb is that deletion of species which appear in all or none of the samples does not affect the value of Cochran's Q statistic (or the association measure W).

Density data

For density data, equation (2.2) suggests using F-tables to find improved critical values for the variance test. Before doing this we need to check that the null model for the F-test is the same as that for the species association test. Unfortunately this is not quite true. The F-test assumes that the data are normally distributed, independent and homoscedastic. This is, in fact, the null model that Schluter (1984) used in his simulation. However, the null model in the species association test has no requirement for independence or homoscedasticity. The null model for the association test is that the overall association is zero (negative and positive cancel out) and this is not the same as all associations being zero. Also, the requirement that the variances be equal across species is not reasonable.

If the samples can be justified as a random sample, then independence and homoscedasticity from sample to sample will be guaranteed. We thus need to investigate the performance of the F-test under violation of the assumption of equal variances from species to species. Box (1954) has shown the influence of unequal variances to be slight. Its effect is to slightly reduce the average size of the F-statistic for testing samples. Thus, somewhat smaller associations are to be expected if variances are very unequal from species to species.

Thus, for testing the null hypothesis of no association, equation (2.2) suggests using critical values of

$$N \cdot \frac{\frac{M}{M-1} F_{N-1, (N-1)(M-1), 1-\alpha/2}}{\frac{1}{M-1} F_{N-1, (N-1)(M-1), 1-\alpha/2} + 1} \quad (5.1)$$

instead of $\chi^2_{N, 1-\alpha/2}$. Again, the difference is not numerically large, but significance probabilities can differ by as much as 60% for reasonable choices of M and N (assuming equal variances). As an example, for N = 3, M = 26 and $\alpha = .10$ (same values as the Simberloff data in Table 2 of Schluter (1984)) the χ^2 critical values are $\chi^2_{3, .05} = .35$ and $\chi^2_{3, .95} = 7.82$, while the critical values from equation (5.1) are

$$3 \cdot \frac{\frac{26}{25} (.051)}{\frac{1}{25} (.051) + 1} = .16$$

and

$$3 \cdot \frac{\frac{26}{25} (3.18)}{\frac{1}{25} (3.18) + 1} = 8.80$$

Acknowledgements

Thanks are due Frances James, Douglas Robson, and Philip Dixon for helpful discussions. Comments by the editor and the referees improved the manuscript.

Literature Cited

Box, G.E.P. 1954. Some theorems on quadratic forms applied in the study of analysis of variance problems, II. Effects of inequality of variance and of correlation between errors in the two-way classification. *Annals of Mathematical Statistics* 25:484-498.

- Carothers, A.D. 1971. An examination and extension of Leslie's test of equal catchability. *Biometrics* 27:615-630.
- Cochran, W.G. 1950. The comparison of percentages in matched samples. *Biometrika* 37:256-266.
- Conover, W.J. 1980. *Practical Nonparametric Statistics*, 2nd edition. Wiley, New York, New York, U.S.A.
- James, F.C. and Boecklen, W. In press. Interspecific morphological relationships and the densities of birds. In *Ecological Communities: Conceptual Issues and the Evidence*, D.R. Strong, D.S. Simberloff, and L.G. Abele, editors. Princeton University Press, Princeton, New Jersey, U.S.A.
- Leslie, P.H. 1958. Statistical appendix. *Journal of Animal Ecology* 27:84-86.
- Macan, T.T. 1976. A twenty-one year study of the waterbugs in a moorland fishpond. *Journal of Animal Ecology* 45:913-922.
- Nilsson, S.G. 1977. Density compensation and competition among birds breeding on small islands in a South Swedish lake. *Oikos* 28:170-176.
- Petr, T. 1972. Benthic fauna of a tropical man-made lake (Volta Lake, Ghana 1965-1968). *Archiv Für Hydrobiologie* 70:484-533.
- Pielou, E.C., 1972. 2^k contingency tables in ecology. *Journal of Theoretical Biology* 34:337-352.
- Robson, D.S. 1972. Statistical tests of significance. Appendix to a paper by E.C. Pielou (1972). *Journal of Theoretical Biology* 34:350-352.
- Schulter, D. 1984. A variance test for detecting species associations, with some applications. *Ecology* 65:998-1005.
- Tate, M.W. and Brown, S.M. 1970. Note on the Cochran Q test. *Journal of the American Statistical Association* 65:155-160.

Wood, T.G. 1974. The distribution of earthworms (Megascolidae) in relation to soils, vegetation, and altitude on the slopes of Mt. Kosciusko, Australia. *Journal of Animal Ecology* 43:87-106.

Appendix

Equivalence between V, W, and the F-statistic.

Using the same notation as in Section 2, a two-way analysis of variance (species by samples) would have the following analysis of variance table:

ANOVA			
<u>Source</u>	<u>Degrees of freedom</u>	<u>Sum of squares</u>	<u>Mean square</u>
Species	M-1	$N\sum(\bar{X}_{i.} - \bar{X}_{..})^2$ = SS(Species)	SS(Species)/(M-1) = MS(Species)
Samples	N-1	$M\sum(\bar{X}_{.j} - \bar{X}_{..})^2$ = SS(Samples)	SS(Samples)/(N-1) = MS(Samples)
Species × Samples Interaction	(N-1)(M-1)	$\sum(X_{ij} - \bar{X}_{i.} - \bar{X}_{.j} + \bar{X}_{..})^2$ = SS(Int)	SS(Int)/(N-1)(M-1) = MS(Int)
Total	MN-1	$\sum\sum(X_{ij} - \bar{X}_{..})^2$	

A test for the significance of samples usually involves calculating the ratio $F = \frac{MS(\text{Samples})}{MS(\text{Int})}$, the usual F-statistic.

This can be shown to be related to W as follows:

$$W = \frac{\sum_{j=1}^N \left(\bar{X}_{.j} - \bar{X}_{..} \right)^2}{\sum_{i=1}^M \frac{1}{N} \sum_{j=1}^N \left(X_{ij} - \bar{X}_{i.} \right)^2}$$

$$\begin{aligned}
 &= \frac{NM^2 \sum_{j=1}^N (\bar{X}_{.j} - \bar{X}_{..})^2}{\sum_{i=1}^M \sum_{j=1}^N (X_{ij} - \bar{X}_{i.})^2} \\
 &= \frac{NM^2 \sum_{j=1}^N (\bar{X}_{.j} - \bar{X}_{..})^2}{\sum_{i=1}^M \sum_{j=1}^N (X_{ij} - \bar{X}_{i.} - \bar{X}_{.j} + \bar{X}_{..} + \bar{X}_{.j} - \bar{X}_{..})^2} \\
 &= \frac{NM (SS(\text{Samples}))}{SS(\text{Samples}) + SS(\text{Int})} \\
 &= \frac{NM \left(\frac{SS(\text{Samples})}{SS(\text{Int})} \right)}{\frac{SS(\text{Samples})}{SS(\text{Int})} + 1} \\
 &= \frac{\frac{N}{M-1} \frac{MS(\text{Samples})}{MS(\text{Int})}}{\frac{MS(\text{Samples})}{(M-1) \cdot MS(\text{Int})} + 1} \\
 &= N \cdot \frac{\frac{M}{M-1} F}{\frac{F}{M-1} + 1} .
 \end{aligned}$$

Also,

$$V = \frac{1}{N} W = \frac{\frac{M}{M-1} \cdot F}{\frac{F}{M-1} + 1} .$$