of populations living together in an area and interacting with each other is known in ecology as a community. Typical communities consist of hundreds or even thousands of species living together in an area.

▲ Figure 13 A coral reef is a complex community with many interactions between the populations. Most corals have photosynthetic unicellular algae called zooxanthellae living inside their cells

Field work – associations between species

Testing for association between two species using the chi-squared test with data obtained by quadrat sampling.

Quadrats are square sample areas, usually marked out using a quadrat frame. Quadrat sampling involves repeatedly placing a quadrat frame at random positions in a habitat and recording the numbers of organisms present each time.

The usual procedure for randomly positioning quadrats is this:

- A base line is marked out along the edge of the habitat using a measuring tape. It must extend all the way along the edge of the habitat.
- Random numbers are obtained using either a table or a random number generator on a calculator.
- A first random number is used to determine a distance along the measuring tape. All distances along the tape must be equally likely.
- A second random number is used to determine a distance out across the habitat at right angles to the tape. All distances across the habitat must be equally likely.
- The quadrat is placed precisely at the distances determined by the two random numbers.

If this procedure is followed correctly, with a large enough number of replicates, reliable estimates of

▲ Figure 15 Quadrat sampling of seaweed populations on a rocky shore population sizes are obtained. The method is only suitable for plants and other organisms that are not motile. Quadrat sampling is not suitable for populations of most animals, for obvious reasons.

If the presence or absence of more than one species is recorded in every quadrat during sampling of a habitat, it is possible to test for an association between species. Populations are often unevenly distributed because some parts of the habitat are more suitable for a species than others. If two species occur in the same parts of a habitat, they will tend to be found in the same quadrats. This is known as a positive association. There can also be negative associations, or the distribution of two species can be independent.

There are two possible hypotheses:

H_0: two species are distributed independently (the null hypothesis).

H_1: two species are associated (either positively so they tend to occur together or negatively so they tend to occur apart).

We can test these hypotheses using a statistical procedure – the chi-squared test.

The chi-squared test is only valid if all the expected frequencies are 5 or larger and the sample was taken at random from the population.

Method for chi-squared test

1 Draw up a contingency table of observed frequencies, which are the numbers of quadrats containing or not containing the two species.

	Species A present	Species A absent	Row totals
Species B present			
Species B absent			
Column totals			

Calculate the row and column totals. Adding the row totals or the column totals should give the same grand total in the lower right cell.

2 Calculate the expected frequencies, assuming independent distribution, for each of the four species combinations. Each expected frequency is calculated from values on the contingency table using this equation:

$$\text{expected frequency} = \frac{\text{row total} \times \text{column total}}{\text{grand total}}$$

3 Calculate the number of degrees of freedom using this equation.

degrees of freedom = (m – 1)(n – 1)

where m and n are the number of rows and number of columns in the contingency table.

4 Find the critical region for chi-squared from a table of chi-squared values, using the degrees of freedom that you have calculated and a significance level (p) of 0.05 (5%). The critical region is any value of chi-squared larger than the value in the table.

5 Calculate chi-squared using this equation:

$$X^2 = \Sigma\frac{(f_o - f_e)2}{f_e}$$

where f_o is the observed frequency

f_e is the expected frequency and

Σ is the sum of.

6 Compare the calculated value of chi-squared with the critical region.

- If the calculated value is in the critical region, there is evidence at the 5% level for an association between the two species. We can reject the hypothesis H_0.
- If the calculated value is not in the critical region, because it is equal or below the value obtained from the table of chi-squared values, H_0 is not rejected. There is no evidence at the 5% level for an association between the two species.

Data-based questions: Chi-squared testing

Figure 16 shows an area on the summit of Caer Caradoc, a hill in Shropshire, England.

The area is grazed by sheep in summer and hill walkers cross it on grassy paths. There are raised hummocks with heather (*Calluna vulgaris*) growing in them. A visual survey of this site suggested that *Rhytidiadelphus squarrosus*, a species of moss growing in this area, was associated with these heather hummocks. The presence or absence of the heather and the moss was recorded in a sample of 100 quadrats, positioned randomly.

Results

Species	Frequency
Heather only	9
Moss only	7
Both species	57
Neither species	27

Questions

1 Construct a contingency table of observed values. [4]

2 Calculate the expected values, assuming no association between the species. [4]

3 Calculate the number of degrees of freedom. [2]

4 Find the critical region for chi-squared at a significance level of 5%. [2]

5 Calculate chi-squared. [4]

6 State the two alternative hypotheses, H_0 and H_1, and evaluate them using the calculated value for chi-squared. [4]

7 Suggest ecological reasons for an association between the heather and the moss. [4]

8 Explain the methods that should have been used to position quadrats randomly in the area of study. [3]

Figure 16 Caer Caradoc, Shropshire

Statistical significance

Recognizing and interpreting statistical significance.

Biologists often use the phrase "statistically significant" when discussing results of an experiment. This refers to the outcome of a statistical hypothesis test. There are two alternative types of hypothesis:

- H_0 is the null hypothesis and is the belief that there is no relationship, for example that two means are equal or that there is no association or correlation between two variables.
- H_1 is the alternative hypothesis and is the belief that there is a relationship, for example that two means are different or that there is an association between two variables.

The usual procedure is to test the null hypothesis, with the expectation of showing that it is false. A statistic is calculated using the results of the research and is compared with a range of possible values called the critical region. If the calculated statistic exceeds the critical region, the null hypothesis is considered to be false and is therefore rejected, though we cannot say that this has been proved with certainty.

When a biologist states that results were statistically significant it means that if the null hypothesis (H_0) was true, the probability of getting results as extreme as the observed results would be very small. A decision has to be made about how small this probability needs to be. This is known as the significance level. It is the cut-off point for the probability of rejecting the null