ESCI 340  BIOSTATISTICAL ANALYSIS

Project 7: Species Richness in Tidepools

The number of species found in an area is an important concern for most environmental sciences, including ecology, ecological restoration, conservation biology, toxicology, and management of forests, riparian zones, and rangelands. Ecologists have debated for decades why species richness is greater in many tropical areas than otherwise comparable temperate ones, and what determines the potential and observed number of species in a given area. Restoration ecologists have used species richness to measure progress toward restoration goals. Conservation biologists are concerned about the loss of species and they often use species richness to determine conservation priorities. Some toxicologists have used species richness as an index of exposure to contaminants, although others have derived more precise measures from the relative tolerance to pollutants of particular species. Managers have used species richness as a surrogate for environmental quality in the systems they oversee. Many factors influencing species richness have been proposed, and the relative importance of a particular factor at a given site must depend on conditions at the site and in the system in which it exists. Factors that have been shown to affect local species richness include: the number of species found in the larger region; the extent of the area sampled; the number of (micro)habitat types within the study area; proximity to sources of dispersing organisms; disturbance frequency, intensity, and extent; and interactions among these factors. In this project, you will evaluate how well various factors explain patterns in species richness observed in tidepools at Larrabee State Park.

Research Question
What factors affect species richness in tidepools, and what is the relative importance of those factors?

Hypotheses
1. Species in a given tidepool constitute a sample drawn from the regional species assemblage. Local species richness depends primarily on regional species richness.

2. Species-area effect. Species richness increases (nonlinearly) with area available. Species-area relationships have been studied extensively, and several mechanisms have been proposed. The most widely known is the theory of island biogeography by MacArthur and Wilson (1963, 1967), which suggests that species number is determined by a balance between rates of species colonization and extinction.

3. Species richness increases with habitat diversity, because additional (micro)habitats provide conditions that support additional species. Species richness in a given tidepool will be proportional to the number of distinct microhabitats within that tidepool.

4. Supply-side ecology. Species richness is limited by transport mechanisms – water currents in the intertidal zone (Roughgarden, et al. 1988; Doherty and Fowler 1994). Since tidepool isolation increases with height in the intertidal zone, tidepool species richness also will decrease with intertidal height (distance to water).

5. Organism survival and therefore species richness increases with tidepool water depth. On average, deeper tidepools fluctuate less in salinity and provide greater protection from vertebrate predators.

6. Species richness is constrained by both tidepool area and microhabitat diversity, according to mechanisms stated in hypotheses 2 and 3.

7. Species richness is constrained by both microhabitat diversity and height in the intertidal zone, for reasons stated in hypotheses 3 and 4.
Species richness is determined by an interaction between microhabitat diversity and distance from high tide line. As one descends in the intertidal zone, the pool of species that could be found increases. The number actually found depends on the number of microhabitats available. Hence, species richness is proportional to the product of number of habitats and distance from high tide line.

**Models for Each Hypothesis**

1. Species richness \((S)\) is a random deviation from the mean.
   \[
   S_j = \mu + \varepsilon_j \quad K = 2
   \]

2. Species richness is a function of area \((A)\). Although most research on species-area relationships have assumed or found a logarithmic relationship between habitat area and species richness (see material at end for background), we will assume a linear relationship.
   \[
   S_{ij} = \alpha + \beta A_i + \varepsilon_{ij} \quad K = 3
   \]

3. Species richness is proportional to habitat diversity \((H)\).
   \[
   S_{ij} = \alpha + \beta H_i + \varepsilon_{ij} \quad K = 3
   \]

4. Species richness decreases with distance from waterline. Equivalently, \(S\) increases with distance \((D)\) from high tide line.
   \[
   S_{ij} = \alpha + \beta D_i + \varepsilon_{ij} \quad K = 3
   \]

5. Species richness is proportional to tidepool water depth \((Z)\).
   \[
   S_{ij} = \alpha + \beta Z_i + \varepsilon_{ij} \quad K = 3
   \]

6. Species richness is constrained by both tidepool area and microhabitat diversity.
   \[
   S_{ijk} = \alpha + \beta A_i + \gamma H_j + \varepsilon_{ijk} \quad K = 4
   \]

7. Species richness is constrained by both microhabitat diversity and distance from high tide line.
   \[
   S_{ijk} = \alpha + \beta H_i + \gamma D_j + \varepsilon_{ijk} \quad K = 4
   \]

8. Species richness is determined by an interaction between microhabitat diversity and distance from high tide line.
   \[
   S_{ijk} = \alpha + \beta H_i D_j + \varepsilon_{ijk} \quad K = 3
   \]

**Field Methods**

You will work in groups of three or four: one person to record data, another to count species, and the remaining one or two to measure tidepool area, depth, and distance from high tide line. Each group should have the following equipment: (1) measuring tape (30 m), (2) metal meter stick, (3) data form(s), and (4) writing implement. So outfitted, complete the following steps.

1. Locate a set of tidepools with areas ranging from the largest available to the smallest puddle containing visible organisms. Also try to include tidepools spanning a wide range of tidal heights, from the highest pool available to ones adjacent to the waterline. Complete the following steps for each tidepool.

2. Record the number of visible invertebrate species the tidepool. Please sample in a nondestructive manner. As you begin to count, mobile species are likely to try to hide. Look for those species first, and then count sessile species.
3 Record the maximum water depth (cm) of the tidepool, using the metal meter stick.
4 Record the length (cm), width (cm), and approximate shape (circle, ellipse, square, rectangle) of the tidepool.
5 Record the minimum distance (m) to open water.
6 Repeat steps 2–5 with each of your remaining tidepools.

A general guideline for determining minimum sample size in statistical modeling is to retain twenty residual degrees of freedom. Because models for hypotheses 6 and 7 require estimates for three parameters, you should try to sample at least 23 tidepools. Work rapidly and efficiently!

**Data Analysis**

1 Enter your data into some electronic format.
2 Calculate variance in number of species ($S$) per tidepool. This is the residual sum of squares ($RSS$) for model 1.
3 Create scatterplots for $S$ versus each predictor variable. Assess the distributions relative to the assumptions of linear regression.
4 Calculate the area ($A$) of each tidepool, and create new variable containing area values. Use a formula for area that is appropriate to the shape of each tidepool.
5 Perform linear regression on data for $S$ and $A$. Determine $RSS$ for model 2 and estimates for $\alpha$, $\beta$.
6 Fit the remaining models to your species data ($S$), using linear regression. Determine $RSS$ for each model.
7 Calculate an AIC (or $AIC_c$) score for each model, using $RSS$ values and $K$ given above. Note that using $RSS$ to determine the maximized log-likelihood of a model requires assuming that $\varepsilon_i$ are normally distributed.
8 Identify the model with the smallest AIC (or $AIC_c$) score, and subtract it from $AIC_c$ scores for the other models to determine $\Delta_i$ values.
9 Use $\Delta_i$ values to determine Akaike weights, $w_i$.
10 Identify the confidence set for best model.
11 Determine probabilities for each model in confidence set.
12 Interpret results of your analysis to answer the research question.
## Data Sheet

<table>
<thead>
<tr>
<th>Pool #</th>
<th>No. Species</th>
<th>Length(m)</th>
<th>Width(m)</th>
<th>Shape</th>
<th>Depth(m)</th>
<th>No. Habitats</th>
<th>Distance to waterline(m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Shapes: square, rectangle, circle, ellipse, triangle, trapezoid, …

Habitats: (1) bottom, (2) open water, (3) vertical wall, (4) 2-sided corner, (5) 3-sided corner, (6) pocket
Appendix: Species-Area Relationships

In many ecological systems, the number of species found increases nonlinearly with the area sampled. [A linear relationship is more likely in our case due to the brief time during which tidepools are isolated and their frequent (twice daily) connection during high tides. Hence, the model for hypothesis 2 above is linear. If you are interested in the usual form for species-area relationships, this Appendix provides some background.] The relationship between area and number of species has important implications for ecology, conservation biology, and natural resource management. Two main hypotheses have been proposed to explain species-area relationships. The first argues that species number increases with habitat diversity, and that larger areas tend to contain more kinds of habitats. The second hypothesis suggests that the species number is determined by a balance between rates of species colonization and extinction. This hypothesis was developed into the theory of island biogeography by MacArthur and Wilson (1963, 1967). The theory predicts that, in the absence of other variables, more species will inhabit larger islands than smaller ones because the colonization-extinction balance is reached with more species on larger islands. For more information about each hypothesis, see chapter 7 in Gotelli (2008).

The simplest description of the species-area relationship is given by equation A1.

\[ S = bA^z \]  

where  
\( S \) = number of species 
\( A \) = area sampled 
\( b \) = fitted constant 
\( z \) = slope of species area relationship

In equation A1, the constant \( b \) depends on the group of organisms sampled and the relative species richness of that group within the region sampled. Equation A1 can be transformed to linear form by taking the logarithm of both sides, yielding equation A2.

\[ \log S = \log b + z \log A \]  

When this kind of species-area relationship applies to a system, equation A2 shows that plotting log-transformed data on species number and island area will produce a straight line. The slope of this line will be \( z \).

Island biogeography has been applied widely to systems of non-oceanic islands. These applications assume that patches of habitat are isolated from each other to some degree. This assumption implicitly requires that the species sampled cannot live in areas surrounding habitat patches. Other assumptions are listed on page 166 in Gotelli (2008).

References


