DEPARTMENTS

Technological Tools

A New Means of Presenting the Results of Logistic Regression

Introduction

The use of logistic regression analysis in ecological studies has greatly increased in recent years. It is a popular and useful statistical tool for predicting the probability of occurrence of a categorical dependent variable (e.g., presence or absence, male or female) based on predictor variables. The results of logistic regression have been presented in a number of ways in the scientific literature: equations with statistics (e.g., Sydeman et al. 1991, Stewart et al. 1996, Bolger et al. 1997, Gross and Kapuscinski 1997, Morrison 1998, Wiser et al. 1998*a*); probability response curves (e.g., Sydeman et al. 1991, Van Sickle et al. 1996, Wiser et al. 1998*a*); and bar charts of the percentage deviance explained by different models (e.g., Wiser et al. 1998*b*). However, these traditional means of presenting the results have many limitations in the information that they provide. We propose a new method for presenting logistic regression data, describe how it can be achieved with current software, and suggest that it should be routinely incorporated in future updates of statistical packages.

Traditional presentation

Fig. 1a shows one of the commonest current methods of presenting logistic regression output, using hypothetical data that describe the probability of a pool being occupied by an invertebrate in relation to pH. The two main limitations of this type of figure are:

1) The observed data for the dependent variable, i.e., presence and absence in pools, are arranged at zero and one on the *Y* axis, making it impossible to ascertain the number of pools at each interval along the *X* axis. This makes the distribution of the data difficult to visualize. The influence of the observed data on the slope

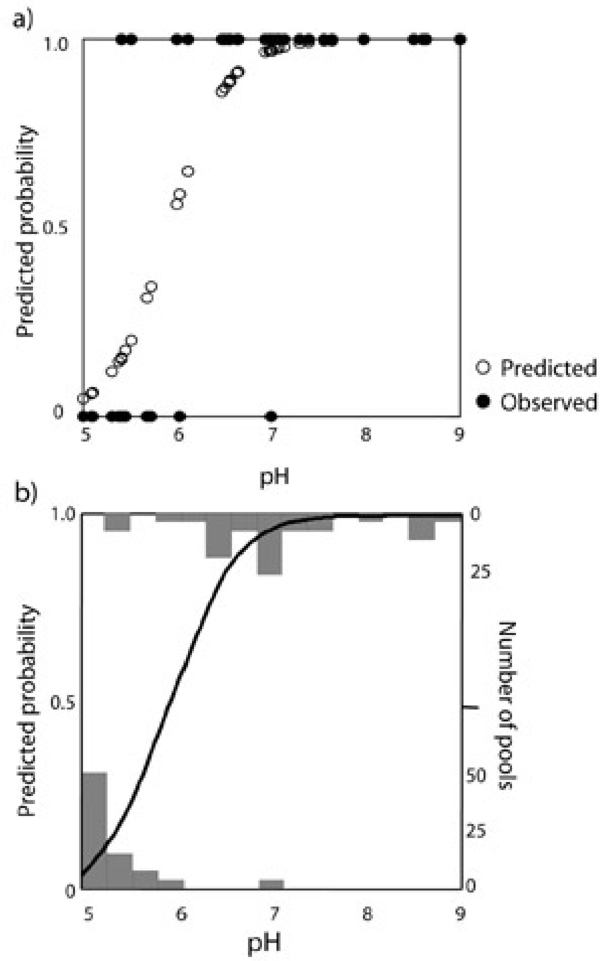

Fig. 1. Fitted logistic regression curves showing that the probability of pool occupation by an invertebrate (presence and absence) is dependent on pH. Both graphs present the same hypothetical data, but (a) is the traditional method of presenting logistic regression graphs produced in SPSS using overlay scatterplots, and (b) is the new method produced using a combination of SPSS and PowerPoint, where the histograms represent the observed data and the line is the predicted probability that a pool will be occupied.

and shape of the logistic regression curve are also unclear.

2) Sample size cannot be ascertained from the figure alone because the number of pools at each interval along the axis is not apparent.

New presentation

Fig. 1b shows the same hypothetical data as shown in Fig. 1a, but in this case the observed data are presented in the form of frequency histograms for each category of the dependent variable, with the associated scale on the right-hand axis. These changes overcome the limitations of the traditional method, outlined above, because the frequency of the observed data at each interval along the axis is now clearly displayed. It is now easy to interpret the summed effect of these points on the logistic regression curve. For example, in Fig. 1a it is impossible to determine how many unoccupied pools have a pH of between 5 and 6. However, we can see from Fig. 1b that there are ~80 pools within this category. It is also now possible to assess the sample size from the figure alone, as the observed data points are displayed against a scale.

Method for the new presentation of logistic regression graphs

At present, combination graphs of this type are not available on any of the standard statistics or graphing packages of which we are aware. Although the presentation of this type of graph is therefore more time consuming, we would argue that, in terms of ease of interpretation, it is worth the extra time and effort. We are sure that there are many different methods and design packages available that could ultimately be used to produce these graphs, but as an example we describe here our step-by-step method, which uses a combination of SPSS v. 11.0 and Microsoft Power Point.

SPSS

1) The data view should have three variables, (a) the dependent variable (e.g., coded 0 and 1), (b) the observed data for the predictor variable, and (c) the predicted probability of group membership saved from the logistic regression analysis.

2) Produce two histograms from the predictor variable by using the dependent variable as the panel variable (Graph-Interactive-Histogram). In the options, it is important to change the maximum scale range to approximately twice the maximum category value.

3) Create a scatterplot with the predicted probability on the *y* axis and the predictor variable on the *x* axis (Graph-Interactive-Scatterplot). Under the *Fit* option select Smoother.

Microsoft PowerPoint

1) Copy and paste the two histograms and the scatterplot onto a PowerPoint slide.

2) Ungroup all three graphs and remove any redundant objects. Regroup the remaining objects in each graph separately.

3) Resize each graph so that the dimensions of all three are equal. Use the flip option to flip vertically the histogram of the higher value on the predicted probability axis (i.e., 1 [presence] in this example). Then overlay the two histograms and the scatterplot, ensuring that the logistic regression curve is brought to the foreground.

4) Add text boxes for the right-hand *y* axis labelling and create a scale for both histograms.

5) Edit the graph for final presentation.

Future directions

The new method for graphical representation of the results of logistic regression analysis presented here greatly increases the information that can be extracted from these figures, and should therefore improve the ease of interpretation of the output. However, the manual production of these figures can be time consuming. If software manufacturers incorporate this type of combination graph in future software updates, we hope that this type of figure will become a common feature of logistic regression analyses.

Literature cited

Bolger, D. T., A. C. Alberts, R. M. Sauvajot, P. Potenza, C. McCalvin, D. Tran, S. Mazzoni, and M. E. Soulé. 1997. Response of rodents to habitat fragmentation in coastal southern California. Ecological Applications **7**:552–563.

Gross, M. L., and A. R. Kapuscinski. 1997. Reproductive success of smallmouth bass estimated and evaluated from family-specific DNA fingerprints. Ecology **78**:1424–1430.

Morrison, L. W. 1998. The spatiotemporal dynamics of insular ant metapopulations. Ecology **79**:1135-1146.

Stewart, A. J., W. R. Hill, K. D. Ham, S. W. Christensen, and J. J. Beauchamp. 1996. Chlorine dynamics and ambient toxicity in receiving streams. Ecological Applications **6**:458–471.

Sydeman, W. J., H. R. Huber, S. D. Emslie, C. A. Ribic, and N. Nur. 1991. Age-specific weaning success of northern elephant seals in relation to previous breeding experience. Ecology **72**:2204–2217.

Van Sickle, J., J. P. Baker, H. A. Simonin, B. P. Baldigo, W. A. Kretser, and W. E. Sharpe. 1996. Episodic acidification of small streams in the northeastern United States: fish mortality in field bioassays. Ecological Applications **6**:408–421.

Wiser, S. K., R. B. Allen, P. W. Clinton, and K. H. Platt. 1998*b*. Community structure and forest invasion by an exotic herb over 23 years. Ecology **79**: 2071–2081.

Wiser, S. K., R. K. Peet, and P. S. White. 1998*a*. Prediction of rare plant occurrence: a southern Appalachian example. Ecological Applications **8**: 909–920.

Jennifer Smart,[1,3] William J. Sutherland,[2] Andrew R. Watkinson,[1,2] and Jennifer A. Gill[2]
Schools of Environmental[1] and Biological Sciences[2]
University of East Anglia
Norwich NR4 7TJ UK
[3]Correspondence: jennifer.smart@uea.ac.uk

WinSSS: Stochastic Spatial Simulator

Introduction

WinSSS is a Windows-based program for simulating stochastic spatial models that are individual-based, have discrete spatial structure, and continuous time. This class of models is commonly referred to as *interacting particle systems* or asynchronously updated *probabilistic cellular automata*. It is ideally suited for developing insight and making predictions in spatial ecology. Ecological examples can be found in Dieckmann et al. (2000) and Durrett and Levin (1994).

WinSSS features an elaborate graphical interface that allows one to choose from various models, specify parameters such as birth/death rates and interaction strengths, initialize with various starting configurations, and view spatial dynamics as well as time series and phase diagrams corresponding to spatial windows of various sizes. One can download the program freely at the URL below. This includes a ready-to-run simulator, with pre-programmed models and an HTML tutorial and help window. For those who would like to code their own models, the C++ code can also be obtained from the authors.

The models in WinSSS include mechanisms for invasion of new territory and competition for resources, head-to-head competition, pathogen spread, and various types of successional dynamics. For example, one can easily run simulations of the ``Rock-Scissors-Paper" model that recently appeared in *Nature* (Kerr et al. 2002) describing spatial coexistence of three competing strains of bacteria. The HTML tutorial gives a brief introduction to these spatially extended individual-based models and provides some references for further reading.

Model specification and parameters

To describe the models and simulations, we begin by noting that all the action takes place on a two-dimensional rectangular lattice (or grid) of sites, with a number of options for the lattice size.

Each site in the lattice can be in a number of different states (represented as colors), depending on the specific model. One can think of a site as an individual or a group of individuals, say in a habitat patch. The state of a given site can change to other states at rates that depend in general on the configuration of states at "neighboring" sites. These changes occur in continuous time and very quickly, so when watching the simulation one typically observes sites changing all over the lattice. However, the changes are asynchronous, due to the continuous-time nature of this (Markov) process. The way to think of this is that every site has associated with it an (exponential) alarm clock whose rate depends on the state at that site and the states at neighboring sites. The site whose alarm rings first makes the appropriate change and all neighboring sites recalculate their rates. All the alarm clocks then start over and we wait for the next one to ring. (We remark that the behavior of synchronously updated cellular automata can be similar in some respects but very different in others. For example, updating all the sites at once can lead to very rigid behavior that produces patterns not typically seen in biological populations.)

Rates and interaction neighborhoods

There are two basic types of rates that allow one to build most models of interest. These are "contact" and "spontaneous" rates. Contact rates are for events that depend on the types at neighboring sites. For example, a vacant site might become occupied by an offspring from a given species at a rate that is proportional to the number of individuals of that species currently within some distance of the vacant site. Contact rates can depend linearly or nonlinearly (e.g., a threshold event) on the states at neighboring sites. There are several options for neighborhood size in the simulator. Spontaneous rates are for events that occur independently of nearby sites. For example, an individual might die or change its life stage after some random time through no effect from other individuals.

Window size and time series

The overall lattice size can be selected from a number of options ranging from 100 × 100 up to 500 × 500. The densities of the different species appear, color-coded, in a separate window below the main simulation. These densities are averages over a spatial window that the user chooses. They can be recorded in an accessible file and used to obtain information about spatial length scales, as in Rand and Wilson (1995). The user can also choose to watch the phase plane trajectories corresponding to any two species. All of these observations of densities under various window sizes yield perspective on the effects of randomness, correlations between sites at various distances, and comparisons with the corresponding mass-action ordinary differential equations.

Implementation

The models in WinSSS were developed using Visual C++. The graphical interface employs OpenGL, the premier environment for developing portable, interactive two-dimensional and three-dimensional graphics applications.

To run WinSSS at reasonable speeds with lattice size 250 × 250 and above, a Pentium III 866 with 256M RAM is recommended. WinSSS has been tested on Windows 2000 and Windows XP. Other operating systems in the Windows family (e.g., Windows 98 and Windows NT) should also work, but we have not tested them. We plan to initiate improvements and extensions based in part on user feedback.

Acknowledgments

Although most of the C++ code and the graphical interface for this simulator were written independently, we were inspired by the pioneering efforts of Ted Cox and Rick Durrett, who created the Unix-based simulator S3. Y. Guan and S. M. Krone were supported in part by NSF grant EPS-00-80935 and NIH grant P20 RR016448.

Literature cited

Dieckmann, U., R. Law, and J. A. J. Metz, editors. 2000. The geometry of ecological interactions: simplifying spatial complexity. Cambridge University Press, Cambridge, UK.

Durrett, R., and S. Levin. 1994. Stochastic spatial models: a user's guide to ecological applications. Philosophical Transactions of the Royal Society of London B **343**:329–350.

Kerr, B., M. Riley, M. Feldman, and B. J. M. Bohannan. 2002. Local dispersal promotes biodiversity in a real life game of rock-paper-scissors. Nature **418**: 171–174.

Rand, D. A, and H. B. Wilson. 1995. Using spatio-temporal chaos and intermediate-scale determinism to quantify spatially extended ecosystems. Proceedings of the Royal Society of London B **259**:111–117.

Yongtao Guan and Stephen M. Krone
Department of Mathematics, University of Idaho
Moscow, ID 83844-1103
E-mail: krone@uidaho.edu
Availability: ‹http://www.webpages.uidaho.edu/~krone›

Homebrew Camera Traps

Subscribers to ECOLOG-L post queries a few times a year about camera traps (shutter-trip systems that automatically photograph passing wildlife), asking for recommendations about particular models, or how to find the least expensive options. During one recent exchange a reader suggested the web site below, which has complete instructions about how to build your own camera trap. In this case, the builder was able to make one for about $80. If you're competent with a soldering iron and know a little about electronics, it looks like a relatively easy project.

‹http://www.jesseshuntingpage.com/homebrew-cams.html›

David Inouye
University of Maryland