FW663 -- Laboratory Exercise

Capture-Recapture Analysis with Programs MARK and CAPTURE

Program CAPTURE is a general program for the analysis of capture-recapture experiments of closed populations. It computes the methods described in the Otis et al. (1978) monograph and in the Los Alamos NERP report (White et al. 1982). Note especially the restriction to closed populations. Program CAPTURE runs on PC-compatible microcomputers. The microcomputer version is identical to the version of the program that runs on mainframes from back in the 1980's. Hence limitations in the mainframe version still exist in the personal computer version of the code. In particular, to improve portability of the code among machines, the default is for the program to read input from the file CAPTIN and write output to the file CAPTLP. Eric Rexstad has written a front end program for CAPTURE called 2CAPTURE (Rexstad and Burnham 1991) with the intent of making it easier to create input files than the brute force approach of using an editor to create a command file. Unfortunately, 2CAPTURE refuses to run on the CNR network. Appendix A of White et al. (1982) is a user's manual for Program CAPTURE, so you may want to begin learning the input commands to CAPTURE. Additional estimators added to the program are documented in Rexstad and Burnham (1991).

Program CAPTURE does not allow covariates to model capture and recapture probabilities. Nor does it allow capture and recapture probabilities to be set in common across groups of animals. Therefore, all the likelihood-based models in CAPTURE can now be estimated with Program MARK using the closed captures model. With MARK, constraints on capture and recapture probabilities can be implemented, as well as covariates provided in the design matrix.

The purpose of today's exercise is to gain familiarity with the closed capture models by using Program MARK to construct each of the likelihood models of Program CAPTURE for a simple data set with only one group. These models are M_0, M_t, M_b, M_{bh}, and M_{bt}. We have provided a SAS program to generate data sets for input to Program MARK. MARK will then run these same data in Program CAPTURE, so that you can see if you have constructed the model correctly by comparing the estimates you obtained from MARK with those from CAPTURE.

The class directory for this exercise (J:\CLASSES\FW663\EXERCISE.15) on the network at the CNR microcomputer laboratory contains the file

J:\CLASSES\FW663\EXERCISE.15\CAPTSIM.SAS

that will create an input data set to Program MARK. Copy this file to your local directory and use SAS to run it. The file CAPTURE2.INP created in your user directory by this SAS program contains a set of simulated capture-recapture data for input to Program MARK, with each student's file containing a different simulation. Hence, your answers will not match your neighbors, although all data sets were simulated from the same model. A second objective of today's exercise is to expose you further to the random variation of the capture-recapture process.

Thus, you will generate multiple data sets from the same model, but get different estimators. We are interested in seeing how well the estimators in Program CAPTURE perform, i.e., is the bias of the estimator important, how do different models perform, and how well does the model selection procedure work.

For your information, Tom Stanley recently completed a Ph.D. Dissertation on model selection in the closed capture models. He investigated both the discriminant classification method used in CAPTURE and AIC. His recommendation is to use a model averaging procedure based on AIC.

Questions for Discussion

1. What model was selected for your data set?

2. What were the population estimate and standard error?

3. Did the confidence interval cover the true value (available in class)?

4. Were the estimates from the models selected as good as estimates from the true model (to be identified in class)?

5. What else?

<u>Additional Problem for Motivated Students</u>

Program CAPTURE has TASK SIMULATE to generate data, select a model, compute an estimate, and tabulate the results. You can evaluate the performance of the program for a particular scenario with this procedure. Try generating data with TASK SIMULATE to match that simulated by the SAS program CAPSIM.SAS, and generate 1000 replicates to fully evaluate the bias, precision, and model selection algorithm of CAPTURE. The simulation capabilities of MARK are still being developed, and are barely up to this task, but you may want to explore them.

Literature Cited

Otis, D. L., K. P. Burnham, G. C. White, and D. R. Anderson. 1978. Statistical inference from capture data on closed animal populations. Wildlife Monogr. No. 62:1-135.

Rexstad, E. A., and K. P. Burnham. 1991. User's guide for interactive program CAPTURE. Colorado Cooperative Wildlife Research Unit, Colorado State University, Fort Collins, Co. 29pp.

White, G. C., D. R. Anderson, K. P. Burnham, and D. L. Otis. 1982. Capture-recapture and removal methods for sampling closed populations. LA-8787-NERP, Los Alamos Nat. Lab., Los Alamos, NM. 235pp.

Over-Parameterized Closed Captures Models

When a model is specified in Program MARK with unidentifiable *p* estimates, the estimates of $\hat{N}$ are just M_{t+1}. Here's why.

The estimate of population size is basically

$$\hat{N} = \frac{M_{t+1}}{1 - (1 - p_1)(1 - p_2) \cdots (1 - p_t)},$$

where the numerator is the number of individuals captured and the denominator is the probability that an animal was captured during the study. That is, the probability of not being initially captured on the first occasion is $1 - p_1$, not being captured on the second occasion is $1 - p_2$, and so on to $1 - p_t$. The product of these terms is the probability of never being captured during the study. Thus, 1 minus this product is the probability of being captured at least once during the study. Therefore, the denominator is a "correction" to inflate the numerator.

When a closed captures model is over-parameterized, the last term in the denominator becomes zero, because p_t is estimated as 1. That is, if no constraints are applied to the *p*'s and *c*'s to estimate p_t, its value is estimated as 1. Making the product zero results in $\hat{N} = M_{t+1}$.

An additional problem with closed capture models in Program MARK is that often the number of parameters is not correctly computed because values of $\hat{N}$ close to M_{t+1} are assumed to not be estimated – causing an error in the algorithm to determine the number of parameters actually estimated and messing up the AIC value.