Basic Capture-recapture Lab

The purpose of this lecture and lab is to introduce you to the absolute basic concepts behind capture-recapture sampling. We will focus on the "Capture" group of "closed" population models for estimating population size and "Jolly-Seber" type models for estimating survival. In this lab we will use program MARK, a free software program developed by Gary White at Colorado State University that is available on the web at

http://www.cnr.colostate.edu/~gwhite/mark/mark.htm

or

http://www.phidot.org/software/mark/

The second page link is maintained by Evan Cooch at Cornell and has links to the primary users guide for MARK.

Goals of the lab:

(1) Provide very basic introduction to MARK
(2) Estimate population size for a closed population using CAPTURE models within MARK using a data set I provide (UF_closed.inp), and a data set you create.
(3) Estimate survival rates in an open population using a Jolly-Seber model and a data set I provide (UF_open.inp).

To Turn In:

Answer each question in 1-3 complete sentences and create the two graphs as described.

Part I Estimating population size with a closed model in CAPTURE

This data set (UF_Closed) is a simulated data set based on sampling for a game fish species in a small lake. Six samples took place over two weeks. During the first two samples, the field crew didn't know what they were doing so they were not very efficient at collecting fish. Their skills improved in later samples.

Although program CAPTURE has been around for 20 years, it is still a very good program for estimating population size for closed populations. There has been a lot of research over the last 5-10 years on different ways to model heterogeneity in capture probability of behavior effects by people like Shirley Pledger, Richard Barker, and Ken Pollock. These models are complicated and some even blend in aspects of open population models. Generally these are types of "mixture models" and we will not explore them at this time. CAPTURE runs the classic closed models and is easily implemented in program MARK. To do this, we will use MARK to interface with CAPTURE. Don't worry, it's easy.

Start a new MARK interface and load the UF_closed data set. This data set represents 6 sampling events over a 2 week interval. Change the number of encounter occasions to six. Click the closed captures module on the left side of the MARK main page, then in the pop up window click the top option "Closed Captures", then press OK. The PIM window should pop up for capture probability. One white box for each sampling interval with different numbers in each representing "different" capture probabilities in each sample interval, this is analogous to the "t" model. If all of the numbers were the same, it would be analogous to the "." model. Click on the Run tab, then click on predefined models, and click select models button. This screen now has tabs for each model parameter capture probability (p) recapture probability (c) and population size (N). Highlight a "." for each parameter to tell MARK to run a model with that parameter constant. Now click ok and you will return to the previous screen. Name this model and click run. The reason you are doing this is to basically run a simple model in MARK to load the data. We will now access program CAPTURE. Why don't we do everything in MARK? Because (I think unfortunately) the really simple, but very good, models in CAPTURE aren't readily available in MARK. MARK does have models that maybe have better statistical properties in some instances, but in many fisheries applications, the CAPTURE models are still the best ones to use now. After you run this simple dot model, go to the "Tests" tab and click "Program Capture" now you will see the option to chose one of several capture models (generally, Mo, Mt, Mb, Mh, Mth etc.) and a few formulations of specific ones (Jackknife Mh and Chao Mh for example, similar models with slightly different formulations). Choose the models (of the basic CATURE models, i.e. you can use a Jacnkknife or Chao Mh if you want) that you think are appropriate.

Don't worry about the different formulations of the model, just look for the notation we used earlier in lecture such as M_t for a time dependent model. The "appropriate" option here will use simple comparison test to find the "best" model, but this test is known to not work that well. Again, first chose biologically reasonable models, then later re-run the CAPTURE module and chose the "appropriate" option and see which one it chooses. When you press OK to run on this screen, a new text pad window will open with the CAPTURE output. Print this out and read through it to see the name of each model that ran and the output.

Questions:

(1) What model did you chose and why?

(2) What is the population size and SE?

(3) This is a simulated data set with a known population of 500 individuals. What is the deviance between the known population and your estimate?

(4) I used the following capture probabilities for each interval: p1 = 0.1; p2 = 0.3; p3 = 0.4; p4 = 0.3; p5 = 0.1; p6 = 0.1. Graph you estimated capture probabilities vs. these known capture probabilities. How well did the model perform? Any trends in model performance (generally high or low)?

(5) Create your own fake data set with time dependent capture probability by completing the Cap_His_Simulation spreadsheet. Provide your own "known" population size and capture probabilities (using different values than 4). Graph you estimated capture probabilities vs. these known capture probabilities. How well did the model perform? Note once you create the fake data in Excel you will need to create a new .inp file to read this into MARK. You can copy the capture histories and expected values over to a new spreadsheet; retype the values in something like notepad, or whatever you want to do. Don't forget the semicolon at the end of each row and to make sure your file has the inp extension on the end.

PART II, Estimating survival using a Jolly-Seber model

Jolly-Seber type models are the primary open population model for estimating survival and abundance. We will use a JS model to estimate survival in this lab.

You can use the handout "Mark Notes" to take you step by step through this lab. What you need to do is develop what you think are the "biologically reasonable" models to fit the following scenario describing this data set.

Data notes:
Load the UF_open.inp data set. This generalized data set is based on a portion of a tagging data set for Gulf sturgeon collected over the last 10 years or so. The fish are tagged as they migrate into the rivers during the spring spawning run. The fish "recruit" to the gear at about 1000-mm TL and as an endangered species there is no harvest on these fish. **Gulf sturgeon are a long-lived species** and **sampling effort and conditions have varied over this ten year period** (hint, the parts in bold give you a bit of information on the most "biologically reasonable models"). Detailed instructions are provided on the "Mark Notes" handout.

(1) What is the most biologically appropriate model and why?

(2) Does this model have AIC support?

(3) What is the survival rate and SE for the "best" model??

(4) What other biologically reasonable models should be considered?

(5) Imagine that this was a short-lived species that has high natural mortality, what types of models would you consider to estimate survival for this species?