

## LAB 06 – MULTINOMIALS AND CMR MODELS FOR CLOSED POPULATIONS

### BEFORE YOU START

Some of the functions in this exercise require the use of the PopTools Excel add-in (<http://www.Poptools.org>). It should already be installed on the computers in the lab. If you are working elsewhere, before beginning this exercise download and install the latest version. When properly installed and functioning, you will see ADD-INS on the menu bar in Excel and PopTools pulldown menu on the menu bar under ADD-INS. There are a couple of things you might try if you can't find it.

If after opening Excel you don't see the ADD-INS menu on the tool bar:

1. Click on FILE in the main menu bar.
2. Select Add-Ins in the menu on the left.
3. Select Excel Add-ins in the Manage field at the bottom of the page and click on Go.
4. Disable the Analysis ToolPak and Analysis-ToolPak – VBA.
5. Enable PopTools and click OK.
6. If PopTools is not in the list of Add-Ins:
  - a. Click on Browse and navigate to "C:\Program Files (x86)\PopTools\PopTools.xla."
  - b. Click OK
7. Restart Excel. ADD-INS should appear in the menu bar. It is OK to re-enable the Analysis ToolPaks if you need them now.

If the Poptools functions aren't working when you open a spreadsheet (e.g. Lab06.xls):

1. Open the file and enable editing
2. ADD-INS/PopTools and select Update PopTools links
3. Click OK everything should update now if the PopTools Add-in is enabled.

### CLOSED POPULATION ESTIMATORS

#### LINCOLN-PETERSON & SEBER

In this exercise you will construct the Lincoln-Peterson and Seber estimators for population size and use Monte-Carlo simulation to estimate the precision and bias of each.

Download the spreadsheet [lab06.xls](#). Go to the top worksheet labeled "Lincoln-Peterson." Column A contains the population of 500 (N) individuals from which you will be sampling labeled 1...N for convenience. Column B is filled with 1s or 0s to indicate whether an individual was marked (1) during the first capture, thus in the unaltered spreadsheet 60 individuals (M) were captured and marked. Also note that the sum of B (in cell H2) is the number marked in the first capture (C). Column D represents the 2nd capture event—50 individuals selected at random without replacement (using the shuffle array function) from the population

in A. Select the cells D2:D51, type in the formula =shuffle(A2:A501), and press CTRL-SHIFT-ENTER. Column E uses the Lookup function to determine the marked status of the individuals captured in the 2nd event from column B. Thus the sum of E (in H9) is the number of marked individuals in the 2nd sample (R).

Using the data in columns B-E calculate the number captured, C; number initially marked, M; number marked recaptured, R; **the Lincoln-Peterson estimator (lecture notes) of population size in the cell H7, and the Seber estimator in H8.**

Note that pressing F9 ‘captures’ a new sample at random from the population.

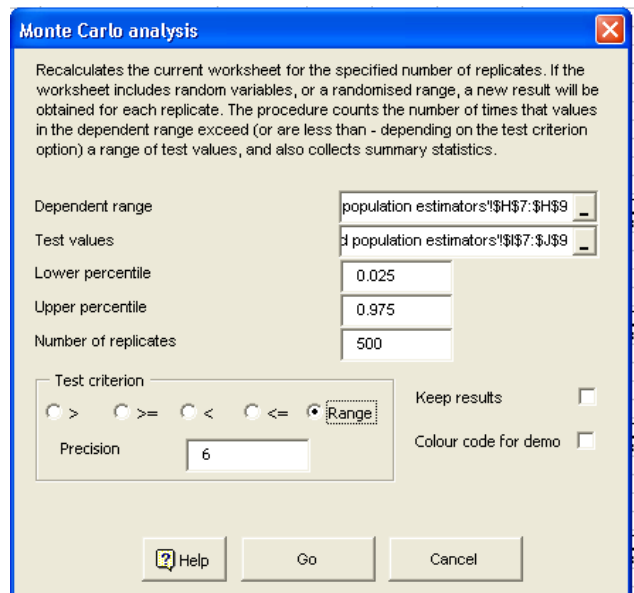
Use the Monte Carlo simulation tool to simulate 500 recapture events of 50 individuals with 50 marked individuals in the population, record the mean population estimates (Lincoln-Peterson and Seber estimators), 95% CLs, and the number of simulations within +10% of the true population size as in the Figure below.

1. Use PopTools/Simulation Tools/Monte Carlo Simulations. Select the range H7 as the dependent cells (This values should change each time the worksheet is resampled).
2. Select the range I7:J9 as the test range (values to which trial will be compared).
3. Leave lower percentile = 0.025 and upper percentile = 0.975 for 95% CLs.
4. Set the number of replicates to 500 and the test criterion to Range. Click Go.
5. Poptools resamples the marked population 500 times and calculates summary statistics
6. Paste the results in cell G13 and label them “M=50”. These results should be similar to the table below.

M=50						
MEAN	VARIANCE	LOWER CL	UPPER CL	WITHIN RANGE	VALID ITERATIONS	TIME TAKEN
609.63	146507.03	277.78	1250	99	499	8 SEC
490.03	49104.27	260.1	1300.5	97	500	
5.08	4.12	1	9	123	500	

The first 4 columns are the respective means, variances, and 95% CLs of the Lincoln-Petersen estimator and the Seber estimator. Values in the first two rows of the columns 5 and 6 are the number of estimates that fell within 10% of the true population size and the number of valid iterations (i.e., with  $R>0$ ). The values in the third row of the fifth column is the number of trials with  $R>7$ .

7. Increase the marked population to 100 individuals by changing the values in B2:B101 to 1. Simulate 500 more recaptures of 50 individuals each as above. Paste the results in cell G18 and label them “M=100”.
8. Repeat the simulations with 200 and 350 marked individuals. (Note that even with  $M=350, C+M<N$ ).



Answer the following in a textbox on the spreadsheet:

1. What is the average bias (difference between the mean estimates and truth) for each of the estimators at each of the marking level? On what portion of the valid iterations was the population estimate within 10% of the true population size for each estimator?
2. If each iteration represents an independent estimate of the population, under what conditions would you trust this method to estimate the size of a wild population of animals under your management?

## MARK EXERCISES

I strongly suggest that you read the documentation under closed captures in the help file with program MARK before completing this exercise. For one thing MARK uses entirely different notation and somewhat different model names than Williams et al. (2002). For example,  $p$  connotes initial capture probability (occasions 1... $K$ ),  $c$  connotes recapture probability (occasions 2... $K$ ),  $p_i$  connotes the proportion of mixtures, and  $N$  indicates population estimate. Additionally, several closed population capture models are listed in the data type menu. Read the help file under Data Types, and make sure to read the sections on closed capture with heterogeneity. For this exercise, we will use the full closed captures with heterogeneity for the mixtures. You may also want to view the example file `mcapture.dbf` packaged with MARK for examples of model parameterizations.

1. Download the file `capmod.zip`. This file includes two data sets with identical capture histories. The capture histories are simulated for 5 encounter occasions for a closed population of 200. In this data set, capture probabilities do not vary over time but the individuals come from 3 subpopulations of 120, 50, and 30 animals with capture probabilities of 0.4750 (the norm), 0.2497 (sort of shy), and 0.7109 (can't stay off the bait). For heuristic purposes, in `capmod.inp` the individuals capture histories are segregated among 3 groups. `Capmod4.inp` was generated from a population of 500 with capture probabilities related to a continuous individual covariate (e.g., mass). This data set also has 5 capture occasions and one individual covariate.
2. Start with the file `capmod2.inp` (5 occ.; 1 group; 0 covariates), and create a new MARK results file based on "Closed captures" data type. Note that there are 5  $p$ s, 4  $c$ s, and 1  $N$  estimable. Run the predefined models:  $p.c.N.$ ,  $p(t) c(t) N.$ ,  $p(t) c. N.$ , and  $p. c(t) N.$ , using Design Matrix coding and logit link (DM). Rename the appropriate models  $M(bt)$  and  $M(b)$ .
3. Note that none of these models corresponds to  $M(0)$  or  $M(t)$ . Cast, name, and run the models  $M(0)$  and  $M(t)$  using the Design Matrix. Locate the predefined models you've already run that correspond with the models  $Mb$  and  $Mtb$ .

Be sure to check the parameter counts and delete models that have parameters than truly can't be estimated – as indicated by abnormally large or small SEs and real parameter estimates not near the bounds for the capture and recapture probabilities.

4. Copy the results table and the estimates from the to the CapMod worksheet in the `lab_06` spreadsheet.
5. Reopen the `Capmod2` results file, open the PIM menu, and select Change Data Type, and from the list that appears select Full Closed Captures with Heterogeneity. Using this data type you should be able to model mixtures such as the one in `capmod2.inp`. When confronted with the number of mixtures dialog, enter 3. Now examine the PIM chart again, and note that  $p_i$  (the mixture proportion) appears below the row for  $p$ .

WILD 7960 - Analysis of Wildlife Populations  
[www.forestry.auburn.edu/grand/wildpop](http://www.forestry.auburn.edu/grand/wildpop)

6. Create a new MARK file using capmod4.inp (5 occ. 1 groups, 1 covariate) based on “Huggins Closed captures” data type. Open the PIM Charts. Note that there are 5  $p$ s, 4  $c$ s, and  $N$  does not appear on the list of estimable parameters.  $N$  is a derived parameter from the Horvitz-Thompson estimator and estimable after the model is run. Cast all eight of the closed capture models using the Design matrix. In Huggins’ model heterogeneity is modeled by the covariates so  $M_h$  has only two parameters – an B1 (intercept) and B2(Var1).
7. Copy the Results table including all 8 models to the Excel worksheet named CapMod4.
8. Use model averaging to calculate the unconditional estimate of  $N$  as the weighted average of the estimates from each model for capmod4. Place these results below the Results Table on the CapMod4 worksheet.