Lab 10 – Estimating abundance with closed-population capture-mark-recapture data
Due before your next lab

The purpose of this lab is to learn how to estimate abundance using mark-recapture data from surveys of closed populations. A closed population does not experience recruitment, mortality, immigration or emigration. The closure assumption is usually only valid over very short time periods. We will learn how to work with data from open populations later.

Work through the problems below, put your answers in a Word file, and then upload it to ELC. Name the file something like "Chandler-lab10.docx".

Part I: Lincoln-Peterson estimation

Suppose you capture, mark, and release 100 largmouth bass (*Micropterus salmoides*) at Lake Herrick. The next day, you return and capture 50 individuals, 25 of which were marked on the first occasion. What is the Lincoln-Peterson estimate of abundance (N)? You do not need to use MARK or R for this – you can do it by hand. Be sure to show your work.

Part II: Closed-population models in MARK and R

You will use either program MARK or R to fit the four mark-recapture models described in Table 1. Undergrads with access to Windows can use program MARK. Everyone else can use the R package 'mra'. Software instructions are below. The assignment is at the end.

Table 1: A description of the four models to be fitted to the stinkpot data.

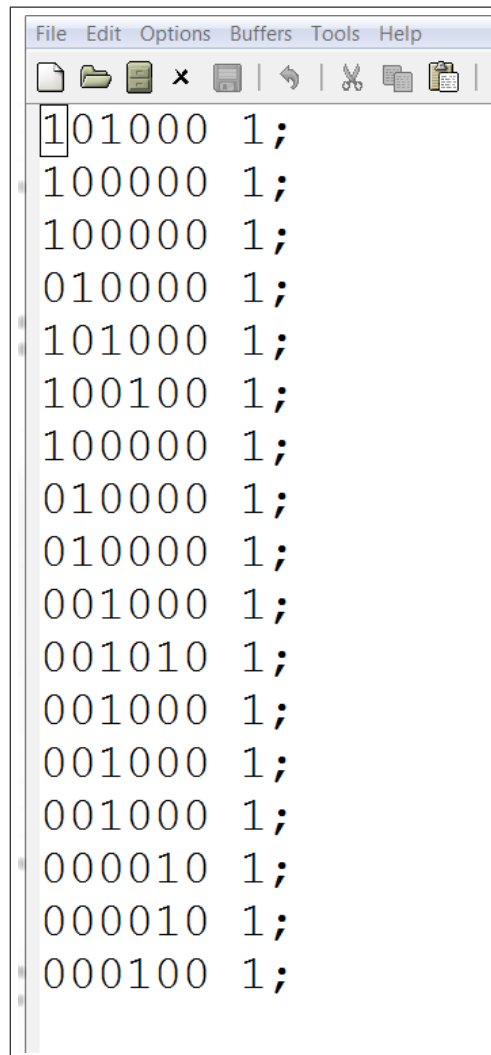| Model	Description
M_0	The most basic model in which p and c are constant
M_t	p differs among sampling occasions and $p_t = c_t$.
M_b	Behavioral response model in which p and c differ. Can describe trap happiness or trap shiness.
M_{tb}	A combination of models M_t and M_b.

Parameter definitions

- p – capture probability. The probability of capturing an individual on a single occasion

- p_t – capture probability on occasion t

- c – recapture probability. The probability of capturing an individual that has been captured previously.

- n – the number of individuals captured

- f_0 – the number of individual not captured

- N – abundance. The number of individuals in the population. $N = n + f_0$.

Closed-population models in program MARK

You can download program MARK here: `http://www.phidot.org/software/mark/downloads/`. Even though it is possible to run MARK on Linux or OS/X, it isn't easy, and I don't recommend it. If you don't have Windows, use R as described below.

The data file (`CH-SO-Andy07.inp`) is a simple text file, formatted as required by program MARK. Each row of the file is a capture history for each of the 17 stinkpots (*Sternotherus odoratus*) captured in 2007 (May 31 - June 5). There were 6 capture occasions, so for every turtle, there are 6 ones and zeros indicating if the turtle was captured on that occasion or not. After each capture history is a space followed by a 1 and a semi-colon to indicate that there was just 1 turtle with this history (Fig. 1).

Figure 1: Stinkpot capture histories in a text file ready to be imported to MARK.

Instructions

(i) Open MARK and create a new project by selecting: `"File > New"`.

(ii) Name the project "Exercise I" and select the encounter history file `"CH-SO-Andy07.inp"` (see Fig. 1)

(iii) Choose `"Closed Captures"` from the list of "data types" on the left and then select `"Huggins' p and c"`.

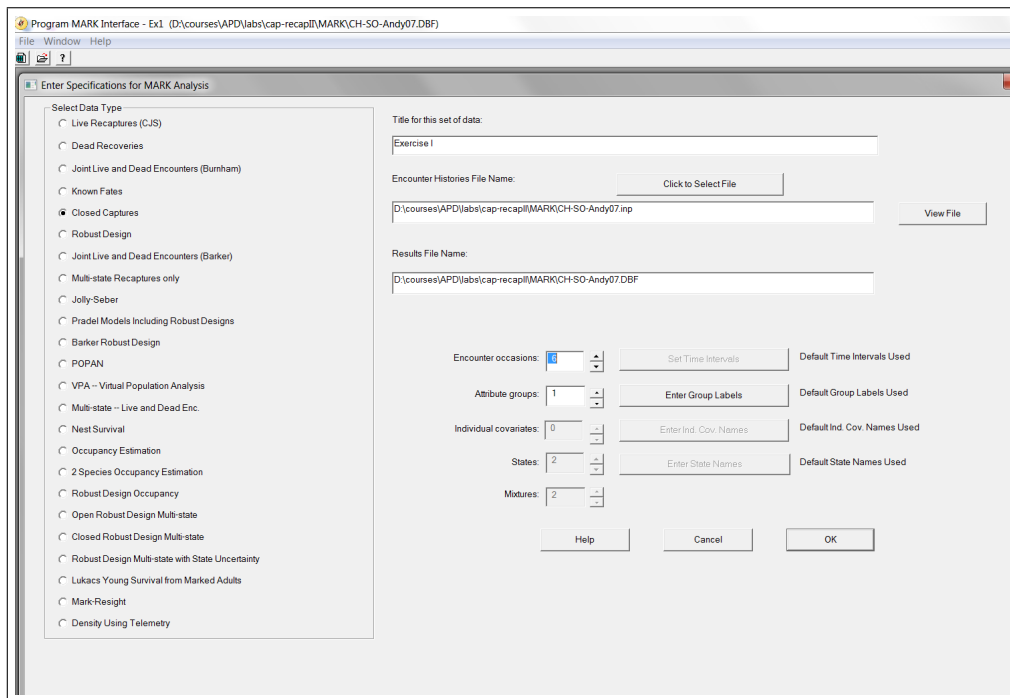(iv) Set the number of encounter occasions to 6, then hit `"OK"`

Figure 2: Setting up the MARK analysis of the stinkpot data.

(v) Run four models differing in their specifications of capture probability (p) and recapture probability (c) by clicking on `"Run > Pre-defined Model(s)"`. Next, click on the `"Select Models"` button and choose models `"M0"`, `"Mt"`, `"Mb"`, and `"Mtb"`. An explanation of these models is shown below in Table 1. Next, click `"OK"`, then hit `"OK to Run"`.

(vi) Inspect the model results, by right-clicking on one of the models in the Results Browser and look at `"Real Estimates"` and `"Derived estimates"`. See Table 1 and the parameter definitions above for help interpreting ther results.

Closed-population models in R

Graduate students and undergraduates without access to Windows, should follow these instructions (instead of the MARK instructions above) for completing the assignment.

Open R (or RStudio) and install the "mra" package using the following command:

```
install.packages("mra")
```

Now load the package like so:

```
library(mra)
```

We can import the data using `read.table`. The only trick is to tell R that the capture histories should be treated as a character string, rather than as a numeric variable. The `colClasses` arguments let's us do that.

```
capture.histories <- read.table("CH-SO-Andy07.inp", sep=" ",
                                colClasses=c("character", "character"),
                                col.names=c("ch", "freq"))
```

It's a small dataset, so we can display it in full:

```
capture.histories

##          ch freq
## 1   101000   1;
## 2   100000   1;
## 3   100000   1;
## 4   010000   1;
## 5   101000   1;
## 6   100100   1;
## 7   100000   1;
## 8   010000   1;
## 9   010000   1;
## 10  001000   1;
## 11  001010   1;
## 12  001000   1;
## 13  001000   1;
## 14  001000   1;
## 15  000010   1;
## 16  000010   1;
## 17  000100   1;
```

We need to convert the capture histories from a character vector to a matrix. The following code does the trick.

```
ch.mat <- t(sapply(capture.histories$ch,
                   function(x) as.integer(strsplit(x, "")[[1]])))
```

```
dimnames(ch.mat) <- list(paste0("Turtle", 1:17), paste0("Time", 1:6))
```

Now the capture histories look like this:

```
ch.mat
```

```
##          Time1 Time2 Time3 Time4 Time5 Time6
## Turtle1      1     0     1     0     0     0
## Turtle2      1     0     0     0     0     0
## Turtle3      1     0     0     0     0     0
## Turtle4      0     1     0     0     0     0
## Turtle5      1     0     1     0     0     0
## Turtle6      1     0     0     1     0     0
## Turtle7      1     0     0     0     0     0
## Turtle8      0     1     0     0     0     0
## Turtle9      0     1     0     0     0     0
## Turtle10     0     0     1     0     0     0
## Turtle11     0     0     1     0     1     0
## Turtle12     0     0     1     0     0     0
## Turtle13     0     0     1     0     0     0
## Turtle14     0     0     1     0     0     0
## Turtle15     0     0     0     0     1     0
## Turtle16     0     0     0     0     1     0
## Turtle17     0     0     0     1     0     0
```

We will use the `F.huggins.estim` function to fit the closed population models. Model M_0 can be fit like this:

```
M0 <- F.huggins.estim(capture=~1, recapture=NULL, histories=ch.mat)
```

Model M_b like this:

```
Mb <- F.huggins.estim(capture=~1, recapture=~1, histories=ch.mat)
```

To fit model M_t, we have to create a time variable.

```
time <- tvar(factor(1:6), nan=17) ## 6 time periods.  17 animals.
Mt <- F.huggins.estim(capture=~time, recapture=NULL, histories=ch.mat)
```

You should be able to figure out how to fit model M_{tb} by extending the code above.

Estimates of abundance along with AICc values can be found by typing the name of the fitted model object. For example:

```
M0
```

```
## Call:
## F.huggins.estim(capture = ~1, recapture = NULL, histories = ch.mat)
##
```

```
##  Capture and Recapture model:
##  Variable     Est        SE
##  (Intercept)  -2.37255  0.49911
##
## Population Size Estimate (se): 41.0364 (16.6626)
## 95% confidence interval for population size: 24.04 to 99.05
## Individuals observed: 17
## Effective sample size: 102
##
## Message =  SUCCESS: Convergence criterion met
## Number of estimable coefficients (estimated) =  1
## Log likelihood =  -43.9354578026216
## Deviance =  87.8709156052432
## AIC =  89.8709156052432
## AICc =  89.9109156052432
```

You can extract information about capture probability (which isn't needed to do the assignment) using code like this:

```
round(M0$p.hat, 3)
```

```
##         [,1]  [,2]  [,3]  [,4]  [,5]  [,6]
##  [1,] 0.085 0.085 0.085 0.085 0.085 0.085
##  [2,] 0.085 0.085 0.085 0.085 0.085 0.085
##  [3,] 0.085 0.085 0.085 0.085 0.085 0.085
##  [4,] 0.085 0.085 0.085 0.085 0.085 0.085
##  [5,] 0.085 0.085 0.085 0.085 0.085 0.085
##  [6,] 0.085 0.085 0.085 0.085 0.085 0.085
##  [7,] 0.085 0.085 0.085 0.085 0.085 0.085
##  [8,] 0.085 0.085 0.085 0.085 0.085 0.085
##  [9,] 0.085 0.085 0.085 0.085 0.085 0.085
## [10,] 0.085 0.085 0.085 0.085 0.085 0.085
## [11,] 0.085 0.085 0.085 0.085 0.085 0.085
## [12,] 0.085 0.085 0.085 0.085 0.085 0.085
## [13,] 0.085 0.085 0.085 0.085 0.085 0.085
## [14,] 0.085 0.085 0.085 0.085 0.085 0.085
## [15,] 0.085 0.085 0.085 0.085 0.085 0.085
## [16,] 0.085 0.085 0.085 0.085 0.085 0.085
## [17,] 0.085 0.085 0.085 0.085 0.085 0.085
```

For model M_0, capture probability is the same for all individuals during all time periods. This won't be the case for the other models.

Assignment

(a) Summarize your results by creating a table in which each row is a model, and include the following columns: the estimates of N (abundance), the standard errors of N (SE), and the AICc values.

(b) The model with the lowest AICc is considered the best in the set of models. Which model has the lowest AICc? Do your results suggest that it is important to account for variation in capture (and recapture) probability over time? Explain.

(c) Why do you think model M_0 is the worst model in the set? Look at the capture histories when answering this question.

(d) Given what you know about turtles and turtle trapping, what sources of variation in capture probability (other than time) do you think we might want to account for to obtain more reliable abundance estimates?