# LECTURE 8

# CLOSED CAPTURE-RECAPTURE MODELS

More on Closed Models Today

-Overview of Models (Review)

-Cell Structures for Simple Models

-More on Heterogeneity Models

-CATURE & MARK

-Lincoln-Petersen example

-More general examples

Table 3.1. Capture-recapture models for closed populations that allow for unequal capture probabilities.

Monograph with minor changes.

| ***Source*** | ***of*** | ***variation*** | ***in*** | ***capture probability*** |
|---|---|---|---|---|
| ***Model**** | Heterogeneity | Trap response | Time | Estimator availability |
| $M_o$ | | | | yes |
| $M_h$ | X[a] | | | yes |
| $M_b$ | | X | | yes |
| $M_{bh}$ | X | X | | yes |
| $M_t$ | | | X | yes |
| $M_{th}$ | X | | X | yes |
| $M_{tb}$ | | X | X | yes |
| $M_{tbh}$ | X | X | X | no |

*This set of 8 models comes from Otis et al. (1978).

[a]Xs denote the sources of variation in capture probability incorporated in the models.

# 2. $M_0$ : THE EQUAL CATCHABILITY MODEL

* The simplest model but usually unrealistic
* There are two parameters in this model

  N - the population size

  P - the probability of capture, which is constant over all animals over all periods.

* M.L. Estimators found iteratively using the programs CAPTURE or MARK.
* Estimators can be highly biased if heterogeneity or trap response is occurring. Variation in capture probabilities, due to time, are less troublesome.

# 3. $M_b$: TRAP RESPONSE MODEL

- This model makes the following assumptions:

    1. Every unmarked animal in the population has the same probability of capture (p) for all samples.

    2. Every marked animal in the population has the same probability of recapture ( c ) for all samples after it has been captured once.

- There are three parameters in the model:

    N - the population size

    p - the probability of capture for unmarked

    c - the probability of capture for marked

- M.L. estimators found iteratively using programs CAPTURE or MARK.

# 6. $M_t$: THE TIME MODEL

- This is the traditional Schnabel model that only allows for time variation in capture probabilities.

- The parameters in the model are:

  N - the population size

  $p_1, p_2, \ldots . p_k$ - the unmarked capture probability of all animals in each sample.

- Programs CAPTURE or MARK provides the MLEs of N and the ps.

  These estimators are not robust to heterogeneity and trap response.

# Lets stop for a minute and make sure we understand the cell structures

$M_0$

$M_t$

$M_b$

I will write down the capture histories and cell structures of expected values for each model on the board.

The cell structures will define the likelihoods in MARK or CAPTURE.

# 4 $M_h$: THE HETEROGENEITY MODEL

- This model allows capture probabilities to vary by animal, due to heterogeneity, but there is no trap response or time variation.
- The parameters in the model are:

  $N$ - the population size

  $p_j$ - the capture probability of animal $j$ for $j = 1, \ldots, N$

  $p_j$ s are assumed to come from distribution $F(p)$, otherwise the model is overparameterized.

- Estimators include:
  - Burnham's Jackknife
  - Lee and Chao's Coverage Estimator
  - Norris nonparametric MLE (Mixture Model)
- Burnham's estimator is widely used, but it has a questionable theoretical basis. It is given by program CAPTURE.

# MORE ON THE HETEROGENEITY MODEL

- This model allows capture probabilities to vary by animal, due to heterogeneity, but there is no trap response or time variation.
- The parameters in the model are:

  $N$ - the population size

  $p_j$ - the capture probability of animal $j$ for $j = 1, \ldots, N$

  $p_j$ s are assumed to come from distribution $F(p)$, otherwise the model is overparameterized.

- Some of what I say will also apply to th and bh models as well

# MORE ON THE HETEROGENEITY MODEL

- There are several estimators available:

•Jackknife Estimator in CAPTURE (can access from MARK)

Burnham and Overton (1978)

The traditional approach used.

•Coverage Estimator in CAPTURE (from MARK)

There are a whole variety of estimators due to Chao and co-workers

•Finite Mixture Models in MARK

Norris and Pollock (1996), Pledger (2000)

This is a ML method.

# MORE ON THE HETEROGENEITY MODEL

Ways of dealing with heterogeneity.

1. Model it as just described on the previous slide.
2. Stratify it out! Use the groups option in MARK.
3. Model it using covariates (Huggins Approach) in MARK. Analogous to a logistic regression problem in some ways.

$$p_i = \frac{\exp(\beta_0 + \beta_1 x)}{1 + \exp(\beta_0 + \beta_1 x)}$$

$$\ln\left(\frac{p_i}{1 - p_i}\right) = \beta_0 + \beta_1 x$$

# 5. $M_{bh}$: THE HETEROGENEITY AND TRAP RESPONSE MODEL

- This model allows capture probabilities to vary due to heterogeneity and trap response, but not time.

- The parameters in the model are:

  N - the population size

  $p_j$ - the unmarked capture probability for the jth animal

  $c_j$- the marked capture probability for the jth animal

The ($p_j$, $c_j$) are assumed to come from some bivariate distribution G(p,c).

# The Heterogeneity And Trap Response Model (continued)

- Program CAPTURE provides a generalized removal estimator.

Basically the first, second, third ..samples are ignored consecutively to try to reduce the influence of the heterogeneity. Unfortunately the precision of the estimators gets worse as more samples are ignored.

- An alternative jackknife estimator, given by Pollock and Otto (1983) is also in CAPTURE.

- Norris has provided a nonparametric MLE of N and G

# 7. OTHER TIME DEPENDENT MODELS

- Model $M_{tb}$ - now has estimator available in CAPTURE
- Model $M_{th}$ - now has estimator available in CAPTURE
- Model $M_{tbh}$ - is only a conceptual model and has no estimators of N available.

# Reminder of Last Lecture Example

# 9. EXAMPLE

- Meadow vole study by James Nichols
- Five sampling periods
- Traps prebaited with corn
- Will show Model Selection Output
- Will show Model $M_h$: The Heterogeneity Model output, because it was the chosen model.

Precision of the estimator is quite good because of the high capture probabilities.

# MODEL SELECTION

## CAPTURE

- There is an old procedure in CAPTURE, which is quite complex, also it only works well if data are very good, i.e., high capture probabilities. The method is based on a whole series of tests which are summarised into one overall criteria between 0 and 1.

- Reduce the number of models to be chosen from if possible.

[Sometimes there may be biological reasons to eliminate some models, e.g, trap response. We will see this in the taxi cab example].

## MARK

If one is using ML models in MARK then one can use AIC methods to chose among models

Table 3.3. Model selection procedure from program CAPTURE for the meadow vole data collected by J.D. Nichols at Patuxent Wildlife Research Center, Laurel, Maryland, in October 1981.

| Model | $M_0$ | $M_h$ | $M_b$ | $M_{bh}$ | $M_t$ | $M_{th}$ | $M_{tb}$ | $M_{tbh}$ |
|---|---|---|---|---|---|---|---|---|
| Criteria | 0.80 | 1.00* | 0.38 | 0.59 | 0.00 | 0.32 | 0.52 | 0.98 |

This suggests one should use the $M_h$ estimator although there is some evidence of trap response and time being present as well.

In the interests of simplicity and getting an estimate we need to use $M_h$ estimator.

Table 3.4. Selected statistics and parameter estimated from program CAPTURE for meadow vole data collected at Patuxent Wildlife Research Center, Laurel, Maryland, in October 1981 by J.D. Nichols. Model $M_h$, the heterogeneity model, is used.

| | Frequencies of capture[a] | | | | |
|---|---|---|---|---|---|
| *i* | 1 | 2 | 3 | 4 | 5 |
| *F( i)* | 29 | 15 | 15 | 16 | 27 |

Number of animals captured = 102[b]

Average P-HAT = 0.44

**Interpolated population estimate = 139, with Standard Error = 10.85**

Approximate 95% Confidence Interval from 177 to 161

[a] These are the numbers of animals caught from 1 to 5 times.

[b]This is the number of distinct animals captured at least once.

p = 0.44 is a very high probability

44% of animals were captured on each occasion

# MANY COMPUTER PROGRAMS

**MARK**

**CAPTURE**

JOLLY

JOLLYAGE

POPAN

RELEASE

SURVIV

MSSURVIV

RDSURVIV

TMSURVIV

# CAPTURE

- Classic closed population models of Otis et al. (1978), but with just some updating for new estimators.
- Contains a model selection procedure (not v good).
- Can be run on the web from Patuxent software archive.
- Copy of program can also be downloaded from Patuxent software archive.
- Can also run from MARK.

# MARK

- User friendly windows based program for capture-recapture,telemetry and band return models.
- Many options (under statement)
- Can run CAPTURE and POPAN from MARK
- Uses AIC for model selection
- Allows multiple groups, age classes, multi-state extension, covariates
- Can download from their web site.Can also download an online book and other resources.

http://welcome.warnercnr.colostate.edu/~gwhite/mark/mark.htm

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

# USE of CAPTURE on the Web

- http://www.mbr-pwrc.usgs.gov/software
- http://www.mbr-pwrc.usgs.gov/software/capture.html
- I will demonstrate how to do this in class today
- See hardcopy handout from class on the input formats and output for three examples.
  - Rabbit data-Darroch (Mt)
  - Microtus data-Jacknife (Mh)
  - Removal data-Zippen (Related to Mb-I will discuss later)

# Use of CAPTURE from MARK

- I will demonstrate this in class using the taxi cab data as an example
- You have to access the output in an unusual way by going to the tests tab in a MARK output window.
- It will be more natural to look at MARK first.

# Use of MARK: For Closed Models

- Starting Today and continuing Tuesday I will demonstrate MARK for closed population capture-recapture models
- First for the rabbit data (2 periods LP) and then for the taxi cab data (10 periods)
- I will emphasize the input format and syntax
- I will show you how to use the parameter index matrices (PIMS) to create specific models to run.(M0, Mt, Mb)
- I will show you how the AIC is used to compare models and select the best one.
- I will show you how to look at the output files for a chosen model
- I will show how to switch to looking at CAPTURE output when you are in a MARK analysis if you want to use both at once.
- For Tuesday I will try and compile some detailed summary notes on some key points of using the program but you will also be learning by trying it.

# Taxi Cab Example from Edinburough (Carrothers1973)

## CAN ACCESS CAPTURE FROM MARK
## CAN USE MARK DIRECTLY

-Closed Population N=420

-k=10 occasions on 10 days close together

-No trap response ( :-))).

-Constant sampling effort so perhaps no time variation either.

-Heterogeneity likely.

## Model selection criteria. Model selected has maximum value.

| Model | M(o) | M(h) | M(b) | M(bh) | M(t) | M(th) | M(tb) | M(tbh) |
|---|---|---|---|---|---|---|---|---|
| Criteria | 0.91 | 1.00 | 0.45 | 0.61 | 0.00 | 0.51 | 0.39 | 0.6 |

Appropriate model probably is M(h)

Suggested estimator is Jackknife or Chao Estimator for M(h)

## Model M(h) Suggested for use here

Jacknife 471 with standard error 36.32

Chao 407 with standard error 27.42

Finite Mixture approach did not work here. Huge SE!

## Model M(0) Null Model not to be used

MLE 368 with standard error 14.4896

Always underestimates when there is heterogeneity.

# Closed Capture-Recapture Analysis and Use of Programs

**Usually best to Use MARK!!**

1. **Use CAPTURE (from MARK).** Old program but allows heterogeneity models to be fit in one analysis.*Does not use AIC* and does not do *multiple groups*.
2. **Use MARK closed options directly**- good *interface*, *AIC*, *multiple groups*

- **Standard Closed Captures** –the non heterogeneity models
- **Huggins model** –*covariates* approach to fitting heterogeneity models (not shown in class but related to logistic regression)
- **Norris and Pollock, Pledger** –*finite mixture models* for heterogeneity.

# Summary Slides (Will revisit)

# Summary Closed Capture-Recapture Design Issues

## Precision Issues

- Need adequate capture probabilities and numbers of samples to estimate standard errors that are small enough (ie. RSE ~20%).

- Look at Tables in Otis et al. (1978).Note that good model selection requires much larger capture probs than just estimation under one (assumed correct) model.

- Full Simulation Study

- Use Expected Values for guesses of what the data might be like and do analysis on that data using MARK or CAPTURE.

# Summary Closed Capture-Recapture Design Issues

**Minimise Model Bias-** Satisfy Assumptions

1. **Closure**- Short studies,no mortality, no recruitment, no immigration or emigration. Check with telemetry sometimes?

2. **Equal Catchability**

**Heterogeneity**-often hard to avoid unless one can use different methods of capture in each sample which is not usually feasible. Rerandomise trap locations each time?

Collect covariate data for Huggins method or to stratify on.

**Trap Response**- often hard to avoid unless one can use different methods of capture in each sample which is not usually feasible.

**Time Variation**-try to eliminate so that simpler models can be used.

3. **No Tag Loss** – Obviously avoid, check out in pilot studies. Use double tagging method to estimate tag loss if it is a problem.

# Summary Closed Capture-Recapture Programs

**CAPTURE**

**- old but still useful, run from web or inside MARK. Which is preferred depends on the situation**

**-to access from MARK look under the tests tab (illogical but that's what they did)**

**MARK**

- **Very powerful but complex to learn to use, There are many procedures and we will only cover a few.**
- **Key issues, input format, PIMS and their manipulation.**

# Key References

Williams et al. (2002). Analysis and Management of Vertebrate Populations. Academic Press.

Amstrup et al. (2005). Handbook of Capture-Recapture Methods. Princeton University Press.

Pollock et al. (1990). Statistical Inference for Capture-Recapture Models. Wildlife Society Monograph. (pdf available). Old but still useful for the basics.