

## User Guide for Program CARE-4 (Markov chain models for capture-recapture data)

Hsin-Chou Yang and Anne Chao

Institute of Statistics, National Tsing Hua University, Hsin-Chu, Taiwan

### Table of Contents:

1. [Introduction](#)
2. [Download and Setup](#)
3. [Data Input Format](#)
4. [Analysis without Covariates](#)
  - 4.1 [Models/Estimators Featured](#)
  - 4.2 [Example and Running Procedures](#)
5. [Analysis with Covariates](#)
  - 5.1 [Models/Estimators Featured](#)
  - 5.2 [Example and Running Procedures](#)
6. [References](#)
7. [Appendix](#)

### 1. Introduction

Program CARE-4 calculates population size estimates under Markov chain models with/without covariates for capture-recapture experiments. We use examples to illustrate data inputs, show operation procedures and interpret estimation results. The program is developed based on software GAUSS. The main backgrounds and theoretical development as well as relevant references for the program are provided in the following reference:

Yang, H.-C. and Chao, A. (2005). Modeling Animals' Behavioral Response by Markov Chain Models for Capture-Recapture Experiments. *Biometrics* **61**, 1010-1017.

You are welcome to use CARE-4 for your own research and applications as long as you will not distribute CARE-4 in any commercial form. If you publish your work based on the results from CARE-4, please use the following reference to cite CARE-4.

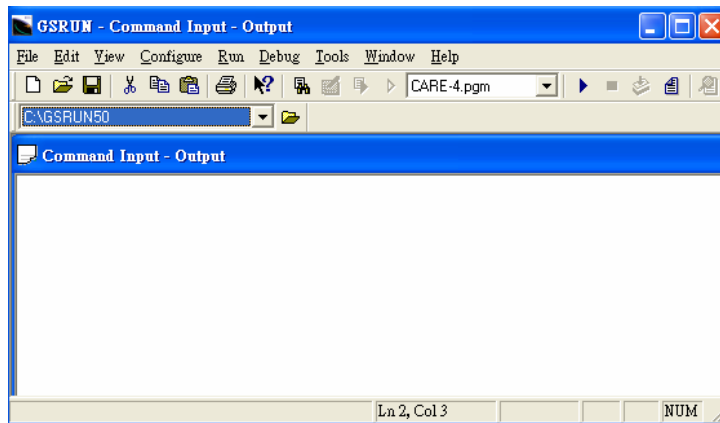
Yang, H.-C. and Chao, A. (2006) Program CARE-4 (for Capture-Recapture Part. 4). Program and User's Guide published at <http://chao.stat.nthu.edu.tw>.

## 2. Download and Setup

Program CARE-4 can be downloaded from Anne Chao's website at <http://chao.stat.nthu.edu.tw/softwareCE.html>. First doubly click the downloaded file "care-4.exe" to unzip all files to a specified folder. The source files along with two illustrative data sets will be stored automatically in the specified folder in your computer.

CARE-4 must be run under an environment of GAUSS. The working environment of GAUSS is provided by the following procedure: first doubly click the "GRTM.exe" to unzip all files of the Gauss Run-Time Module (GRTM) in the previously specified folder. Then doubly click the executable file "setup.exe" to install the Gauss Run-Time Module, which is GUASS free-ware for non-commercial redistribution. (The GRTM allows licensee to redistribute licensee's compiled GAUSS programs free of charge to other users who do not have GAUSS so long as licensee's GAUSS program is distributed free of charge.) Then doubly click the icon "GSRUN50" on the desktop of your computer to initialize the Gauss Run-Time Module and then the interface is shown below.

Figure 1. The interface of CARE-4.



### 3. Data Input Formats

For analysis without covariates, users should provide capture history data; for analysis with covariates, users should provide capture history data along with covariate data. All data must be read from ascii files.

Capture-recapture data are arranged in a matrix, called “individual capture history” matrix, with the rows representing the capture histories of each captured individual and the columns representing the captures on each occasion. The capture history of each captured individual is expressed as a series of 0’s (non-captures) and 1’s (captures). An illustrated example (Example 1 in Section 4.2) for capture history matrix with 104 animals and 5 capture occasions is shown as follows:

Table 1. An illustrated data for capture history data

	<i>Occasion 1</i>	<i>Occasion 2</i>	<i>Occasion 3</i>	<i>Occasion 4</i>	<i>Occasion 5</i>
<i>Animal 1</i>	1	1	0	1	1
<i>Animal 2</i>	0	1	0	1	1
<i>Animal 3</i>	1	1	0	1	0
<i>Animal 4</i>	1	1	0	1	1
<i>Animal 5</i>	1	1	0	1	1
<i>Animal 6</i>	1	1	0	1	1
<i>Animal 7</i>	1	1	0	1	1
<i>Animal 8</i>	1	1	0	1	1
<i>Animal 9</i>	1	1	0	0	1
<i>Animal 10</i>	1	0	1	0	1
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
<i>Animal 104</i>	1	1	1	1	1

For data with covariate information, two types of covariate data can be analyzed in CARE-4. The first type is “individual covariate”, which can be inputted followed by the individual capture history. Individual covariate variables may be discrete (e.g. gender) or continuous (e.g. body weight in kilogram or wing length in centimeter). An illustrated example (Example 2 in Section 5.2) with 171 animals, 10 capture occasions and 3 individual covariates (2 discrete covariates and 1 continuous covariate) is shown as follows:

Table 2. An illustrated data for capture history data and individual covariates data

	<i>Occasion 1</i>	<i>Occasion 2</i>	<i>Occasion 3</i>	.	.	<i>Occasion 10</i>	<i>Gender</i>	<i>Age</i>	<i>Weight</i>
<i>Animal 1</i>	1	1	1	.	.	1	2	3	12
<i>Animal 2</i>	1	1	1	.	.	0	1	3	15
<i>Animal 3</i>	1	0	1	.	.	0	2	3	15
<i>Animal 4</i>	1	1	0	.	.	1	1	1	15
<i>Animal 5</i>	1	0	0	.	.	0	1	3	15
<i>Animal 6</i>	1	0	0	.	.	0	2	2	15
<i>Animal 7</i>	1	1	0	.	.	0	1	3	15
<i>Animal 8</i>	1	0	1	.	.	0	2	2	15
<i>Animal 9</i>	1	1	0	.	.	0	1	3	16
<i>Animal 10</i>	1	0	1	.	.	1	2	3	19
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
<i>Animal 171</i>	0	0	0	.	.	1	2	3	20

The second type is “occasional covariate”, which should be saved in another file. This kind of covariate is used to describe variables that vary with occasions (e.g. catch efforts or environmental variables). If experiments are conducted over time, this type of covariate also describes time-varying variables. Occasional covariate variables may be discrete (e.g. day or night) or continuous (e.g. catch efforts or temperature). These variables should be inputted in different columns. Each column represents an occasional variable. An illustrated example with 10 capture occasions and three occasional variables is shown as follows:

Table 3. An illustrated data for occasional covariate data

	<i>Capture in day or night</i>	<i>Weather</i>	<i>Catch effort</i>
<i>Occasion 1</i>	day	cloudy	25
<i>Occasion 2</i>	night	rainy	10
<i>Occasion 3</i>	day	rainy	11
<i>Occasion 4</i>	night	cloudy	12
<i>Occasion 5</i>	day	cloudy	12
<i>Occasion 6</i>	night	sunny	13
<i>Occasion 7</i>	day	sunny	15
<i>Occasion 8</i>	night	sunny	40
<i>Occasion 9</i>	day	sunny	9

Remark: The individual ID, occasion indices and covariate names (in italic font) in the previous three examples (Table 1 – Table 3) are shown just for illustrations. They should NOT be included in data files. Moreover, continuous type covariates must follow by the categorical type covariates as covariate data are inputted.

#### 4. Analysis without Covariates

##### 4.1 Models/Estimators Featured

CARE-4 considers a class of Markov chain models. The capture history of each animal is modeled as a Markov chain with a bivariate state space with states determined by the capture status (capture/non-capture) and marking status (marked/unmarked). The most general one is the Markov chain model  $\mathbf{MM}_{2tb}$  (where the subscript 2 denotes a two-dimensional model, t and b denote respectively time-varying and behavioral response). However, this model is not identifiable, but it is useful for a conceptual framework; see Yang and Chao (2005) for details. We list in Table 4 a class of Markov chain models discussed in Yang and Chao (2005) and CARE-4. This class of models include three classic models ( $\mathbf{M}_b$ ,  $\mathbf{M}_t$  and  $\mathbf{M}_0$ , proposed in Otis et al., 1978) as special sub-models.

Assume that there are  $N$  animals in the study area and capture-recapture experiments are conducted over  $t$  occasions. The purpose is to estimate the unknown parameter  $N$ . In Table 4, we define  $X_{it} = I$  [the  $i$ th animal is caught at time  $t$ ], and  $Y_{it} = I$  [the  $i$ th animal is marked at time  $t$ ]. For simplicity, we define three bivariate states as  $a$ ,  $b$  and  $c$ :  $a = (0, 0)$ ,  $b = (0, 1)$  and  $c = (1, 1)$ . Let  $P_{aa}(i, t) = P\{(X_{it}, Y_{it}) = a \mid (X_{i,t-1}, Y_{i,t-1}) = a\}$  be the transition probability from state  $a$  at time  $t-1$  to state  $a$  at time  $t$ . Other transition probabilities  $P_{ab}(i, t)$ ,  $P_{ba}(i, t)$ , ...,  $P_{cc}(i, t)$  are similarly defined.

Except for the non-identifiable model  $\mathbf{MM}_{2tb}$ , CARE-4 provides population size estimates for the other six models in table 4. Basically, there are two types of estimators: unconditional MLE and conditional MLE. The six models as well as the two MLE's featured in CARE-4 and their abbreviations in output (see later sample output for an example) are shown in Table 5. All the formulas for estimators are provided in the Appendix.

Table 4. Models without covariates in CARE-4.

Model	Assumption	Restriction in model <b>MM<sub>2tb</sub></b>
<b>MM<sub>2tb</sub></b>	$P\{(X_{it}, Y_{it}) = s_t   (X_{i,t-1}, Y_{i,t-1}) = s_{t-1}\} = \begin{cases} P_{ac}(t), & s_{t-1} = a, s_t = c \\ P_{bc}(t), & s_{t-1} = b, s_t = c \\ P_{cc}(t), & s_{t-1} = c, s_t = c \end{cases}$	
<b>MM<sub>2b</sub></b>	$P\{(X_{it}, Y_{it}) = s_t   (X_{i,t-1}, Y_{i,t-1}) = s_{t-1}\} = \begin{cases} P_{ac}, s_{t-1} = a, s_t = c \\ P_{bc}, s_{t-1} = b, s_t = c \\ P_{cc}, s_{t-1} = c, s_t = c \end{cases}$	$P_{ac}(t) = P_{ac}$ $P_{bc}(t) = P_{bc}$ $P_{cc}(t) = P_{cc}$
<b>MM<sub>1tb</sub></b>	$P\{(X_{it}, Y_{it}) = s_t   (X_{i,t-1}, Y_{i,t-1}) = s_{t-1}\} = \begin{cases} P_{ac}(t), s_{t-1} = a, s_t = c \\ P_{bc}(t), s_{t-1} = b, s_t = c \\ P_{ac}(t), s_{t-1} = c, s_t = c \end{cases}$	$P_{bc}(t) = P_{ac}(t)$
<b>MM<sub>1b</sub></b>	$P\{(X_{it}, Y_{it}) = s_t   (X_{i,t-1}, Y_{i,t-1}) = s_{t-1}\} = \begin{cases} P_{ac}, s_{t-1} = a, s_t = c \\ P_{bc}, s_{t-1} = b, s_t = c \\ P_{ac}, s_{t-1} = c, s_t = c \end{cases}$	$P_{bc}(t) = P_{ac}(t) = P_{bc} = P_{ac}$
<b>M<sub>b</sub></b>	$P\{(X_{it}, Y_{it}) = s_t   (X_{i,t-1}, Y_{i,t-1}) = s_{t-1}\} = \begin{cases} P_{ac}, s_{t-1} = a, s_t = c \\ P_{bc}, s_{t-1} = b, s_t = c \\ P_{bc}, s_{t-1} = c, s_t = c \end{cases}$	$P_{cc}(t) = P_{bc}(t) = P_{cc} = P_{bc}$
	(Removal model)	
<b>M<sub>t</sub></b>	$P\{(X_{it}, Y_{it}) = s_t   (X_{i,t-1}, Y_{i,t-1}) = s_{t-1}\} = \begin{cases} P(t), s_{t-1} = a, s_t = c \\ P(t), s_{t-1} = b, s_t = c \\ P(t), s_{t-1} = c, s_t = c \end{cases}$	$P_{ac}(t) = P_{bc}(t) = P_{cc}(t) = P_{ac} = P_{bc} = P_{cc} = P(t)$
<b>M<sub>0</sub></b>	$P\{(X_{it}, Y_{it}) = s_t   (X_{i,t-1}, Y_{i,t-1}) = s_{t-1}\} = \begin{cases} P, s_{t-1} = a, s_t = c \\ P, s_{t-1} = b, s_t = c \\ P, s_{t-1} = c, s_t = c \end{cases}$	$P_{ac}(t) = P_{bc}(t) = P_{cc}(t) = P_{ac} = P_{bc} = P_{cc} = P$

We remark that another program, CARE-2, available from the same website as CARE-4, we provide population size estimates for the class of ecological models proposed in Otis et al. (1978). A comparison of all models considered in CARE-2 and CARE-4 is tabulated in Table 5.

Table 5. Estimators for analysis without covariates in CARE-2 and CARE-4.

Model	Estimators/Approaches	Estimators/Approaches
	in CARE-2	in CARE-4
<b>M<sub>0</sub></b>	Unconditional MLE (UMLE) Conditional MLE (CMLE)	Unconditional MLE (UMLE) Conditional MLE (CMLE)

	Estimating equations (EE)	
<b>M<sub>t</sub></b>	Unconditional MLE (UMLE)	Unconditional MLE ( <a href="#">UMLE</a> )
	Conditional MLE (CMLE)	Conditional MLE ( <a href="#">CMLE</a> )
	Estimating equations (EE)	
<b>M<sub>b</sub></b>	Unconditional MLE (UMLE)	Unconditional MLE ( <a href="#">UMLE</a> )
	Conditional MLE (CMLE)	Conditional MLE ( <a href="#">CMLE</a> )
	Estimating equations (EE)	
<b>MM<sub>1b</sub></b>		Unconditional MLE ( <a href="#">UMLE</a> ) Conditional MLE ( <a href="#">CMLE</a> )
<b>MM<sub>2b</sub></b>		Unconditional MLE ( <a href="#">UMLE</a> ) Conditional MLE ( <a href="#">CMLE</a> )
<b>M<sub>tb</sub></b>	Unconditional MLE (UMLE)	
	Conditional MLE (CMLE)	
	Estimating equations (EE)	
<b>MM<sub>1tb</sub></b>		Unconditional MLE ( <a href="#">UMLE</a> ) Conditional MLE ( <a href="#">CMLE</a> )
<b>M<sub>h</sub></b>	Jackknife (JK1, JK2, IntJK)	
	Sample coverage (SC1 & SC2)	
	Estimating equations (EE)	
<b>M<sub>th</sub></b>	Sample coverage (SC1 & SC2)	
	Estimating equations (EE)	
<b>M<sub>bh</sub></b>	Jackknife (JK)	
	Sample coverage (SC)	
	Estimating equations (EE)	
<b>M<sub>tbh</sub></b>	Estimating equations (EE)	

In addition to population size estimates, CARE-4 also calculates the asymptotic standard error (S.E.) estimate by inverting a Fisher information matrix. For interval estimation, CARE-4 provides two types of 95% confidence intervals: one is based on a log-transformation method (Chao, 1987) and the other one is based on asymptotic normality assumption.

Two criteria, likelihood ratio test and AIC, can be performed to select the proper model under a series models considered in CARE-4. For the former, CARE-4 provides the number of parameters and the maximum value of likelihood function under different models, where the unconditional log-likelihood (LL) and conditional log-likelihood (LC) are shown respectively. Then the likelihood ratio test can be performed by comparing the difference of the maximum likelihoods of any two models of interest with the critical value determined by a chi square distribution. For the latter, the AIC value is provided so that the model with the minimum AIC value can be selected.

#### 4.2 Example and Running Procedures

An example is used to demonstrate the use of CARE-4 for analyzing animal capture-recapture data without covariates.

*Example 1: Mouse data*

The data set used in this example is distributed with CARE-4 and stored by default in the directory c:\program files\CARE-4\data. The mouse (*Microtus pennsylvanicus*) data were first discussed in Nichols, Pollock and Hines (1984). The original live-trapping experiment was conducted monthly from June to December, 1980. During each month, the capture-recapture procedure was repeated for 5 consecutive days. The detailed data are given in Williams, Nichols and Conroy (2002, pp. 525-528). We use the data collected in June. A total of 104 distinct mice were caught in the experiment.

We describe the procedures for analyzing mouse data. The output will be shown and briefly described. The following procedure must be executed in a GAUSS environment.

- (1) Provoke GAUSS environment either by doubly clicking GSRUN50 on your desktop as described in Download and Setup or by clicking the executable file GSRUN.exe stored in the directory GSRUN50.
- (2) Click "File" on the top menu of GAUSS and subsequently click "Run Program" and select the program CARE-4.gcg which is stored in a pre-specified working directory (The default is c:\program files\CARE-4\). It prompts you subsequently the following input steps:
- (3) "Please choose the method for analysis: 1. Analysis without covariates. 2. Analysis with covariates." In this example, we input **1**.
- (4) "Please input the number of distinct individuals:" In this example, we input **104**.
- (5) "Please input the number of sampling occasions:" Input **5**.
- (6) "Please input the filename for capture history:" Input **c:\program files\CARE-4\data\example1.dat**.
- (7) "Please input the filename to save the output:" Input for example **c:\program files\CARE-4\output.out**. Please wait a moment and the results will be shown in the GAUSS window. Moreover, the output is also saved in c:\program files\CARE-4\output.out. The standard output for CARE-4 with this example with the above input is shown in Table 6. Interpretations follow by the output.



Table 6. The output of mouse data analysis.

```
#####
### CARE-4 for Markov chain capture-recapture study ###
### Authors: Hsin-Chou Yang and Anne Chao          ###
### Version: 1.1 (September 2005)                  ###
#####

You have chosen the analysis without covariates.

Starting time of data analysis (Year/Month/Day/Hour/Minute/Second):
2005 / 7 / 5 / 23 / 28 / 45

Ending time of data analysis (Year/Month/Day/Hour/Minute/Second):
2005 / 7 / 5 / 23 / 28 / 45

=====
== 1. Summary Statistics ==
=====

-----
Filename of input data : c:\CARE-4\data\program files\example1.dat
Total number of distinct animals : 104
Number of capture samples : 5
-----

  j | u[j]  m[j]  n[j]  M[j]  ft[j]  fl[j]  n00[j]  n01[j]  n10[j]  n11[j]
-----|-----
  1 | 63    0    63    0    20    63    41    63    0    0
  2 | 19    52   71    63    14    30    22    19    11   52
  3 | 9     65   74    82    21    18    15    18    15   56
  4 | 7     56   63    91    22    18    13    17    28   46
  5 | 6     57   63    98    27    20    19    22    22   41
  6 |      104
-----

  j | naa[j]  nab[j]  nac[j]  nba[j]  nbb[j]  nbc[j]  nca[j]  ncb[j]  ncc[j]
-----|-----
  1 | 41     0    63     0     0     0     0     0     0
  2 | 22     0    19     0     0     0     0    11    52
  3 | 13     0     9     0     2     9     0    15    56
  4 | 6      0     7     0     7    10     0    28    46
  5 | 0      0     6     0    19    16     0    22    41
  Sum | 82     0   104     0    28    35     0    76   195
-----

=====
== 2. Model Description ==
=====

CARE-4 analyzes a two-dimensional Markov chain model (MM2(b)), two one-dimensional Markov chain models (MM1(b), MM1(tb)), and three sub-models (M0, Mt, Mb). See Yang and Chao (2005) for relevant likelihood functions.

N      : refers to the population size;
j      : refers to the jth sample or jth capture occasion;
n00(j) : number of observed animals that are not captured on samples j-1 and j;
n01(j) : number of observed animals that are not captured on sample j-1 but captured on sample j;
Similar definitions for n10(j), n11(j).
```

p00(j) : transition probability from non-capture on sample j-1 to non-capture on sample j;  
 p01(j) : transition probability from non-capture on sample j-1 to capture on sample j;  
 Similar definitions for p10(j) and p11(j).

State 'a': the state that an animal has never been caught on and prior to sample j;  
 State 'b': the state that an animal has been caught prior to sample j but not caught on sample j;  
 State 'c': the state that an animal has been caught prior to sample j and caught again on sample j;

naa : number of transitions from state a to state a in the experiment;  
 Similar definitions for nab, nac, ..., etc.

pbb : number of transitions from state b to state b in the experiment;  
 Similar definitions for pbc, pca, ..., etc.

==== 3. The Fit & Estimation of Capture Recapture Models ====

Model	Estimate	S.E.	95%CI (Log-transf.)	95%CI (Asy-normal)
MO(CMLE)	104.65	0.83	( 104.09 , 108.47 )	( 103.02 , 106.28 )
MO(UMLE)	104.06	0.81	( 104.00 , 109.22 )	( 102.48 , 105.64 )
Mt(CMLE)	104.62	0.81	( 104.09 , 108.40 )	( 103.03 , 106.21 )
Mt(UMLE)	104.03	0.79	( 104.00 , 108.44 )	( 102.48 , 105.57 )
Mb(CMLE)	106.67	2.09	( 104.69 , 114.37 )	( 102.57 , 110.77 )
Mb(UMLE)	105.86	1.93	( 104.35 , 113.96 )	( 102.07 , 109.65 )
MM1b(CMLE)	106.36	1.80	( 104.63 , 112.91 )	( 102.83 , 109.90 )
MM1b(UMLE)	105.66	1.71	( 104.31 , 112.81 )	( 102.30 , 109.02 )
MM1tb(CMLE)	107.05	2.20	( 104.85 , 114.87 )	( 102.73 , 111.37 )
MM1tb(UMLE)	106.26	2.07	( 104.49 , 114.44 )	( 102.20 , 110.32 )
MM2b(CMLE)	106.67	2.09	( 104.69 , 114.37 )	( 102.57 , 110.77 )
MM2b(UMLE)	105.86	1.93	( 104.35 , 113.96 )	( 102.07 , 109.65 )

Model	MIN(-LL)	MIN(-Lc)	#[para.]	AIC
MO(CMLE)	-42.99	338.45	1.00	678.91
MO(UMLE)	-43.18	338.47	2.00	-82.36
Mt(CMLE)	-45.37	336.09	5.00	682.18
Mt(UMLE)	-45.57	336.10	6.00	-79.14
Mb(CMLE)	-48.18	332.66	2.00	669.31
Mb(UMLE)	-48.25	332.68	3.00	-90.50
MM1b(CMLE)	-51.20	329.69	2.00	663.38 *
MM1b(UMLE)	-51.27	329.71	3.00	-96.54 #
MM1tb(CMLE)	-57.96	322.81	9.00	663.63
MM1tb(UMLE)	-58.02	322.84	10.00	-96.04
MM2b(CMLE)	-51.24	329.59	3.00	665.18
MM2b(UMLE)	-51.31	329.62	4.00	-94.63

\* the model with minimum AIC among all conditional estimators  
 # the model with minimum AIC among all unconditional estimators

==== 4. The MLEs of Transition Probabilities ====

\*\*\* Model MO(CMLE) \*\*\*

j	p00[j]	p01[j]	p10[j]	p11[j]
1.0000	0.3617	0.6383	0.6383	0.3617
2.0000	0.3617	0.6383	0.6383	0.3617
3.0000	0.3617	0.6383	0.6383	0.3617
4.0000	0.3617	0.6383	0.6383	0.3617
5.0000	0.3617	0.6383	0.6383	0.3617

*** Model MO(UMLE) ***				
j	p00[j]	p01[j]	p10[j]	p11[j]
1.0000	0.3581	0.6419	0.6419	0.3581
2.0000	0.3581	0.6419	0.6419	0.3581
3.0000	0.3581	0.6419	0.6419	0.3581
4.0000	0.3581	0.6419	0.6419	0.3581
5.0000	0.3581	0.6419	0.6419	0.3581

*** Model Mt(CMLE) ***				
j	p00[j]	p01[j]	p10[j]	p11[j]
1.0000	0.3978	0.6022	0.3978	0.6022
2.0000	0.3214	0.6786	0.3214	0.6786
3.0000	0.2927	0.7073	0.2927	0.7073
4.0000	0.3978	0.6022	0.3978	0.6022
5.0000	0.3978	0.6022	0.3978	0.6022

*** Model Mt(UMLE) ***				
j	p00[j]	p01[j]	p10[j]	p11[j]
1.0000	0.3944	0.6056	0.3944	0.6056
2.0000	0.3175	0.6825	0.3175	0.6825
3.0000	0.2887	0.7113	0.2887	0.7113
4.0000	0.3944	0.6056	0.3944	0.6056
5.0000	0.3944	0.6056	0.3944	0.6056

*** Model Mb(CMLE) ***			
j	pac[j]	pbc[j]	pcc[j]
1.0000	0.5217	0.6886	0.6886
2.0000	0.5217	0.6886	0.6886
3.0000	0.5217	0.6886	0.6886
4.0000	0.5217	0.6886	0.6886
5.0000	0.5217	0.6886	0.6886

*** Model Mb(UMLE) ***			
j	pac[j]	pbc[j]	pcc[j]
1.0000	0.5326	0.6886	0.6886

2.0000		0.5326	0.6886	0.6886
3.0000		0.5326	0.6886	0.6886
4.0000		0.5326	0.6886	0.6886
5.0000		0.5326	0.6886	0.6886

\*\*\* Model MM1b(CMLE) \*\*\*

j		p00[j]	p01[j]	p10[j]	p11[j]
1.0000		0.4671	0.5329	0.2804	0.7196
2.0000		0.4671	0.5329	0.2804	0.7196
3.0000		0.4671	0.5329	0.2804	0.7196
4.0000		0.4671	0.5329	0.2804	0.7196
5.0000		0.4671	0.5329	0.2804	0.7196

\*\*\* Model MM1b(UMLE) \*\*\*

j		p00[j]	p01[j]	p10[j]	p11[j]
1.0000		0.4598	0.5402	0.2804	0.7196
2.0000		0.4598	0.5402	0.2804	0.7196
3.0000		0.4598	0.5402	0.2804	0.7196
4.0000		0.4598	0.5402	0.2804	0.7196
5.0000		0.4598	0.5402	0.2804	0.7196

\*\*\* Model MM1tb(CMLE) \*\*\*

j		p00[j]	p01[j]	p10[j]	p11[j]
1.0000		0.4115	0.5885	0.0000	0.0000
2.0000		0.5687	0.4313	0.1746	0.8254
3.0000		0.5007	0.4993	0.2113	0.7887
4.0000		0.4856	0.5144	0.3784	0.6216
5.0000		0.5005	0.4995	0.3492	0.6508

\*\*\* Model MM1tb(UMLE) \*\*\*

j		p00[j]	p01[j]	p10[j]	p11[j]
1.0000		0.4071	0.5929	0.0000	0.0000
2.0000		0.5608	0.4392	0.1746	0.8254
3.0000		0.4895	0.5105	0.2113	0.7887
4.0000		0.4730	0.5270	0.3784	0.6216
5.0000		0.4914	0.5086	0.3492	0.6508

\*\*\* Model MM2b(CMLE) \*\*\*

j		pac[j]	pbc[j]	pcc[j]
1.0000		0.5217	0.5556	0.7196
2.0000		0.5217	0.5556	0.7196
3.0000		0.5217	0.5556	0.7196
4.0000		0.5217	0.5556	0.7196

5.0000		0.5217	0.5556	0.7196
----- -----				
*** Model MM2b(UMLE) ***				
j		pac[j]	pbc[j]	pcc[j]
----- -----				
1.0000		0.5326	0.5556	0.7196
2.0000		0.5326	0.5556	0.7196
3.0000		0.5326	0.5556	0.7196
4.0000		0.5326	0.5556	0.7196
5.0000		0.5326	0.5556	0.7196
----- -----				

In the outputs of analysis without covariates, the version of software CARE-4, starting time and ending time of data analysis are shown first and then followed by the four main parts of CARE-4 outputs. They are (1) summary statistics; (2) model description; (3) the fit and estimation of capture recapture models; (4) the MLEs of transition probabilities.

The first part of the output shows basic information including the data filename, (c:\program files\CARE-4\data\example1.dat for this example), the number of distinct animals caught in the experiment (104 in this case), the number of trapping occasions (5 in this case). Then the summary statistics are listed in order. We use these data to introduce some notation. The numbers of captures for the six occasions are  $(n_1, n_2, \dots, n_5) = (63, 71, 74, 63, 63)$ . Out of the  $n_j$  animals, there are  $u_j$  first-captures and  $m_j$  recaptures, so that  $u_j + m_j = n_j$ , with  $(u_1, u_2, \dots, u_5) = (63, 19, 9, 7, 6)$  and  $(m_1, m_2, \dots, m_6) = (0, 52, 65, 56, 57)$ . The statistic  $M_j$  denotes the number of marked animals just before the  $j$ th occasion. Thus  $M_j = u_1 + u_2 + \dots + u_{j-1}$  and  $(M_1, M_2, \dots, M_6) = (0, 63, 82, 91, 98, 104)$  for these data. That is, the number of marked individuals in the population progressively increased from  $M_1 = 0$  to  $M_6 = 104$ . Here  $M_{t+1}$  denotes the total number of distinct animals caught in the experiment. The frequency counts for the five occasions are  $(f_5(1), f_5(2), \dots, f_5(5)) = (20, 14, 21, 22, 27)$ , where  $fT(j)$  denotes the number of animals captured exactly  $j$  times on occasions 1, 2, ...,  $T$ . Since singleton information is usually important, we also list  $(f1(1), f1(2), \dots, f1(5)) = (63, 30, 18, 18, 20)$  where  $f1(j)$  denotes the number of singletons on occasions 1, 2, ...,  $J$ .

Besides the capture-recapture records, the transition data based on the capture status for each trapping occasion  $\{n_{00}(t), n_{01}(t), n_{10}(t), n_{11}(t); t = 1, \dots, 5\}$  are also shown. The numbers of transitions from state 0 (non-capture) to state 0 (non-capture) at

occasions 1, ..., 5 are 41, 22, 15, 13 and 19; The numbers of transitions from state 0 (non-capture) to state 1 (capture) from occasion 1 to occasion 5 are respectively, 63, 19, 18, 17 and 22; Similar interpretations hold for other numbers of transitions. The transition data based on both the capture status and marking status,  $\{n_{aa}(t), n_{ab}(t), n_{ac}(t), n_{ba}(t), n_{bb}(t), n_{bc}(t), n_{ca}(t), n_{cb}(t), n_{cc}(t)\}$  as defined in Table 4 are provided too. The second part shows the description of models and definitions of variables. The theoretical backgrounds of these models can refer to Yang (2002) and Yang and Chao (2005).

The third part shows the results of model fitting and estimation for six models. Under each model, the MLE of  $N$ , its estimated standard error (SE), the 95% confidence interval based on the log-transformation method (Chao, 1987) and asymptotic normality method are shown in order. Moreover, the logarithms of the maximum unconditional likelihood (LL), the maximum Binomial likelihood (LB) and the maximum conditional likelihood (LC) along with the number of parameters and AIC are presented. The fourth part shows the results of the estimated transition probabilities.

We now interpret the output for Example 1. The classical model  $\mathbf{M}_b$  fits the data well based on the goodness of fit test using the counts  $\{u_1, u_2, \dots, u_5\}$ , where  $u_j$  is the number of new captures at time  $j$ . The three Markov models  $\mathbf{MM}_{1b}$ ,  $\mathbf{MM}_{2b}$  and  $\mathbf{MM}_{1tb}$  also provide adequate fits to the transition counts. Thus, we focus on the results for the classical model  $\mathbf{M}_b$ , and three Markov chain models:  $\mathbf{MM}_{1b}$ ,  $\mathbf{MM}_{2b}$  and  $\mathbf{MM}_{1tb}$  based on conditional MLE method. All behavioral models yield similar population size estimates (both the UMLE and CMLE are 106 or 107) and similar estimated SEs. However, the interpretations of the animal's behavior are different under different models.

On the basis of the classical model  $\mathbf{M}_b$ , we perform the following three likelihood ratio tests (LRT) for two nested models based on the maximized conditional likelihood. Since this is a case of multiple-comparison, the significance level for each test is fixed to be  $1.67\% = (5\%)/3$  in order to control a Bonferroni-type overall level of 5%.

- (1) Classical  $\mathbf{M}_b$  vs.  $\mathbf{MM}_{2b}$ : The chi-squared-based LRT with one degree of freedom yields a value of 6.14 and P-value of 1.32%, which is significant. This indicates that the Markov chain model  $\mathbf{MM}_{2b}$  provides significant improvement over the classical behavioral model.
- (2)  $\mathbf{MM}_{2b}$  vs.  $\mathbf{MM}_{1b}$ : There is little difference between the maxima of the likelihoods. The two models are thus not statistically different and the simpler model  $\mathbf{MM}_{1b}$  is

preferable. This can also be seen from the fitted model  $\mathbf{MM}_{2b}$ , where  $\hat{P}_{ac}$  is close to  $\hat{P}_{bc}$ , which indicates no enduring effect.

(3)  $\mathbf{MM}_{1b}$  vs.  $\mathbf{MM}_{1tb}$ : The LRT for 7 degrees of freedom gives a value of 13.76 and a P-value of 5.56%, which is not significant under our multiple comparisons.

Based on the above multiple comparisons, we select the Markov chain model  $\mathbf{MM}_{1b}$ . This conclusion is consistent if model selection based on AIC is performed. Under model  $\mathbf{MM}_{1b}$ , the AICs for unconditional and conditional MLE are -96.54 and 663.38 respectively. Under this model, the transition probabilities are estimated to be  $\hat{p}_{01} = 0.53$  (SE 0.04),  $\hat{p}_{11} = 0.72$  (SE 0.03), which indicates an ephemeral “trap-happy” case. If an animal has not been captured on the current occasion, then the probabilities for capture and non-capture on the next occasion are about the same (0.53 vs. 0.47). However, if captured, the probability of recapture on the next occasion becomes 0.72. These data provide an example that the Markov model with an ephemeral effect gives a better description of an animal’s behavioral response.

## 5. Analysis with Covariates

### 5.1 Models/Estimators Featured

In program CARE-4, we distinguish covariates as two types: individual covariates and occasional covariates as in Huggins (1989, 1991). Individual covariates include individual’s characteristics (age, sex, body weight or wing length) and occasional covariates could be environmental variables (temperature on each occasion) or known catch-effort expended in trapping method (e.g., number of traps on each capture occasion).

Suppose for each animal, there are  $s$  individual covariates. Let the individual covariates for the  $i$ th animal be denoted as  $\mathbf{W}_i' = (W_{i1}, W_{i2}, \dots, W_{is})$  and  $\boldsymbol{\beta}' = (\beta_1, \beta_2, \dots, \beta_s)$  denotes the effects of these covariates. It is necessary to assume that the individual covariates are constant across the  $t$  capture occasions in the experiment, as they cannot be measured on an occasion if the individual is not captured. If heterogeneity is fully explained by individuals’ covariates, then the heterogeneity effect can be expressed conveniently as  $\boldsymbol{\beta}' \mathbf{W}_i = \beta_1 W_{i1} + \beta_2 W_{i2} + \dots + \beta_s W_{is}$ .

Assume that there are  $b$  occasional covariates:  $\{R_{11}, R_{12}, \dots, R_{1t}\}, \{R_{21}, R_{22}, \dots,$

$R_{2t}$  }, ...,  $\{R_{b1}, R_{b2}, \dots, R_{bt}\}$ . For example,  $\{R_{11}, R_{12}, \dots, R_{1t}\}$  may represent the temperature on each occasion, and  $\{R_{b1}, R_{b2}, \dots, R_{bt}\}$  may represent the capture effort on each occasion. Let  $\mathbf{r}' = (r_1, r_2, \dots, r_b)$  denote the effects of the occasional covariates. Define  $\mathbf{R}_j' = (R_{1j}, R_{2j}, \dots, R_{bj})$ . Then the occasional effect for the  $j$ th occasion is  $\mathbf{r}' \mathbf{R}_j = r_1 R_{1j} + r_2 R_{2j} + \dots + r_b R_{bj}$ .

The logistic-type Markov chain model incorporating covariates is expressed as

$$\begin{cases} \text{logit } \{P_{01}(ij)\} = a + c_j + \boldsymbol{\beta}' \mathbf{W}_i + \mathbf{r}' \mathbf{R}_j \\ \text{logit } \{P_{11}(ij)\} = a + d_j + v + \boldsymbol{\beta}' \mathbf{W}_i + \mathbf{r}' \mathbf{R}_j \end{cases} \quad (1)$$

where  $a$  denotes the baseline intercept and  $\{c_1, c_2, \dots, c_{t-1}\}$  and  $\{d_1, d_2, \dots, d_{t-1}\}$  represent the unknown occasional or time effects ( $c_t = 0, d_t = 0$ ). The parameter  $v$  represents the effect of a capture on the previous occasion, which implies that  $v > 0$  corresponds to a case of trap-happy and  $v < 0$  corresponds to a case of trap-shy. The interpretation of the coefficient of any  $\beta$  is based on the fact that when  $\beta > 0$ , the larger the covariate is, the larger the capture probability is, while if  $\beta < 0$  then the larger the covariate is, the smaller the capture probability is. A similar interpretation pertains to the coefficient of any  $r$  for occasional covariates. Note that if there are no covariates in the previous logistic-type Markov chain model, then the model is only a re-parameterization of the Markov chain model **MM**<sub>1tb</sub> discussed in Section 4. To distinguish the models with covariates from those without covariates, we use a superscript  $*$  to denote Markov chain models with covariates.

The parameters in the logistic models are estimated by a conditional approach based on the captured individuals (Huggins, 1989, 1991). The conditional likelihood on which inferences are based is given by

$$L_c(\boldsymbol{\xi}) = \sum_{i=1}^M \prod_{j=1}^t \frac{1}{(1 - Q_j)} \left\{ [P_{00}(ij)]^{I[X_{i,j-1}=0, X_{ij}=0]} [P_{01}(ij)]^{I[X_{i,j-1}=0, X_{ij}=1]} [P_{10}(ij)]^{I[X_{i,j-1}=1, X_{ij}=0]} [P_{11}(ij)]^{I[X_{i,j-1}=1, X_{ij}=1]} \right\}.$$

Therefore, the conditional likelihood approach provides a unified method if there is no unknown individual heterogeneity. The population size is estimated by the Horvitz-Thompson estimator, which is  $\hat{N}_{HT} = \sum_{i=1}^M \{1 - \prod_{j=1}^t [1 - \hat{P}_{01}(ij)]\}^{-1}$ , where  $\hat{P}_{01}(ij)$  is the estimated capture probability evaluated at the conditional MLE. The variance of the resulting estimator can be estimated by an asymptotic variance formula derived in Huggins (1989, 1991). The current convergence criterion in GAUSS, as we know, is that a relative gradient  $< 10^{-5}$ . The default of maximum number of iterations in CARE-4 is



Table 7. Models with covariates in CARE-4.

Model	Assumption	Restriction in model $\mathbf{M}^*_{tbh}$
$\mathbf{MM}^*_{tbh}$	$\begin{cases} \text{logit} \{P_{01}(ij)\} = a + c_j + \boldsymbol{\beta}' \mathbf{W}_i + \mathbf{r}' \mathbf{R}_j \\ \text{logit} \{P_{11}(ij)\} = a + d_j + v + \boldsymbol{\beta}' \mathbf{W}_i + \mathbf{r}' \mathbf{R}_j \end{cases}$	
$\mathbf{MM}^*_{bh}$	$\begin{cases} \text{logit} \{P_{01}(ij)\} = a + \boldsymbol{\beta}' \mathbf{W}_i + \mathbf{r}' \mathbf{R}_j \\ \text{logit} \{P_{11}(ij)\} = a + v + \boldsymbol{\beta}' \mathbf{W}_i + \mathbf{r}' \mathbf{R}_j \end{cases}$	set $c_j = 0, d_j = 0, \mathbf{r} = \mathbf{0}$
$\mathbf{MM}^*_{tb}$	$\begin{cases} \text{logit} \{P_{01}(ij)\} = a + c_j + \mathbf{r}' \mathbf{R}_j \\ \text{logit} \{P_{11}(ij)\} = a + d_j + v + \mathbf{r}' \mathbf{R}_j \end{cases}$	set $\boldsymbol{\beta} = \mathbf{0}$
$\mathbf{MM}^*_{th}$	$\text{logit}(P_{ij}) = a + c_j + \boldsymbol{\beta}' \mathbf{W}_i + \mathbf{r}' \mathbf{R}_j$	set $v = 0$
$\mathbf{MM}^*_h$	$\text{logit}(P_{ij}) = a + \boldsymbol{\beta}' \mathbf{W}_i$	set $c_j = 0, \mathbf{r} = \mathbf{0}, v = 0$
$\mathbf{MM}^*_b$	$\begin{cases} \text{logit} \{P_{01}(ij)\} = a \\ \text{logit} \{P_{11}(ij)\} = a + v \end{cases}$	set $\boldsymbol{\beta} = \mathbf{0}, c_j = 0, d_j = 0, \mathbf{r} = \mathbf{0}$
$\mathbf{MM}^*_t$	$\text{logit}(P_{ij}) = a + c_j + \mathbf{r}' \mathbf{R}_j$	set $\boldsymbol{\beta} = \mathbf{0}, v = 0$
$\mathbf{MM}^*_0$	$\text{logit}(P_{ij}) = a$	set $\boldsymbol{\beta} = \mathbf{0}, c_j = 0, d_j = 0, \mathbf{r} = \mathbf{0}, v = 0$

### 5.2 Example and Running Procedures

In the following section, we provide one example to demonstrate the procedure of CARE-4 for covariate analysis.

#### Example 2: House mice data (with three individual covariates)

The data for the house mouse (*Mus musculus*) were originally collected by Coulombe and analyzed by Otis et al. (1978, pp. 62-67) and Huggins (1989). This data are distributed with CARE-4 and saved by default in c:\program files\CARE-4\data\. The experiment was carried out for 10 occasions (in the morning and night daily for five

days). Two individual covariates are collected for each mouse: gender (male and female) and age (young, semi-adult and adult), and one occasional covariate (capture in day or night) is recorded. A total of 173 mice were caught. We excluded two records in the following analysis because the covariates for two of the mice were missing. For the complete data, the user can view capture history data and individual covariates data saved by default in c:\program files\CARE-4\data\example2-1.dat and occasional covariate data saved in c:\program files\CARE-4\data\example2-2.dat. Part of the complete data including capture history and individual covariates are shown in Table 8 and occasional covariate is shown in Table 9.

Otis et al. (1978) concluded there was no behavioral response effect but time variation and individual heterogeneity were strong. No suitable estimators were available at the time, and thus they suggested the use of the number of distinct animals caught in the experiment.

Table 8. Individual capture history of deer mice with three covariates: Gender (0: male, 1: female) and Age (y: young, a: adult).

<i>Occ 1</i>	<i>Occ 2</i>	<i>Occ 3</i>	<i>Occ 4</i>	<i>Occ 5</i>	<i>Occ 6</i>	<i>Occ 7</i>	<i>Occ 8</i>	<i>Occ 9</i>	<i>Occ 10</i>	<i>Gender</i>	<i>Age</i>
1	1	1	1	1	1	1	1	1	1	0	y
1	0	0	1	1	1	1	1	1	1	1	y
1	1	0	0	1	1	1	1	1	1	0	y
.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.
0	0	0	0	0	1	1	1	1	1	0	a
0	0	0	0	0	1	1	1	1	1	1	a

Table 9. Occasional covariate data: Capture in day or night.

<i>Capture in day or night</i>	
<i>Occasion 1</i>	morning
<i>Occasion 2</i>	night
<i>Occasion 3</i>	morning
<i>Occasion 4</i>	night
<i>Occasion 5</i>	morning
<i>Occasion 6</i>	night
<i>Occasion 7</i>	morning
<i>Occasion 8</i>	night
<i>Occasion 9</i>	morning

In this example, the experiment time (morning or night) was treated as an occasional covariate and gender and age were regarded as individual covariates. We selected female and adult as reference groups and defined the following dummy variables:

$W_{i1} = 1$  if the  $i$ th animal is a male, and 0 for a female;

$W_{i2(1)} = 1$  if the  $i$ th animal is a young, and 0 otherwise;

$W_{i2(2)} = 1$  if the  $i$ th animal is a semi-adult, and 0 otherwise.

For the occasional covariate, night-time serves as a reference group and we define

$R_{1j} = 1$  if it is morning at time  $j$ , and 0 for night-time.

We consider the following model and all sub-models:

$$\begin{cases} \text{logit } \{P_{01}(ij)\} = a + r_1 R_{1j} + \beta_1 W_{i1} + \beta_{2(1)} W_{i2(1)} + \beta_{2(2)} W_{i2(2)} \\ \text{logit } \{P_{11}(ij)\} = a + r_1 R_{1j} + v + \beta_1 W_{i1} + \beta_{2(1)} W_{i2(1)} + \beta_{2(2)} W_{i2(2)} \end{cases}$$

We describe the procedures for analyzing house mice data with covariates. The following procedure must be executed in a GAUSS environment.

- (1) Provoke GAUSS environment either by doubly clicking GSRUN50 on your desktop as described in Download and Setup or by clicking the executable file GSRUN.exe stored in the directory GSRUN50.
- (2) Click "File" on the top menu of GAUSS and subsequently click "Run Program" and select the program CARE-4.gcg which is stored in a pre-specified working directory (The default is c:\program files\CARE-4\). It prompts you subsequently the following input steps:
- (3) "Please choose the method for analysis: 1. Analysis without covariates. 2. Analysis with covariates." In this example, we input **2**.
- (4) "Please input the number of distinct individuals:" In this example, we input **171**.
- (5) "Please input the number of sampling occasions:" Input **10**.
- (6) "Please input the number of categorical individual covariates:" Input **2**.
- (7) "Please input the number of continuous individual covariates:" Input **0**.
- (8) "Please input the filename containing the capture history and individual covariates (continuous type covariates must follow by the categorical type covariates):" Input **c:\program files\CARE-4\data\example2-1.dat**.

- (9) "Please input the number of categorical occasional covariates:" Input **1**.
- (10) "Please input the number of continuous occasional covariates:" Input **0**.
- (11) "Please input the filename containing the occasional covariates (continuous covariates must follow by the categorical covariates):" Input **c:\program files\CARE-4\data\example2-2.dat**.
- (12) "Do you want to include the unknown time effects (y or n)?" (This means that whether the effects  $\{c_1, c_2, \dots, c_{t-1}\}$  are needed in the logistic model). We input **n**.
- (13) "Please input the filename to save the output:" Input for example **c:\program files\CARE-4\output.out**. Please wait a moment and the results will be shown in the GAUSS window. Moreover, the output is also saved in c:\program files\CARE-4\output.out. The standard output for CARE-4 in this example with the above input is shown in Table 10.

格式化: 編號 + 階層: 1 + 編號  
 樣式: 1, 2, 3, ... + 起始號  
 碼: 1 + 對齊方式: 左 + 對齊:  
 0 pt + 定位點之後: 18 pt + 縮  
 排: 18 pt, 定位點: 不在 1.5  
 字元

Remark: If you have abundant data, it may take a long time to get your output due to complicated iterative estimation in GAUSS program operating on a large array or high-dimensional matrix.

Table 10. The output of covariate analysis for house mice data.

```
#####
### CARE-4 for Markov chain capture-recapture study ###
### Authors: Hsin-Chou Yang and Anne Chao          ###
### Version: 1.1 (September 2005)                  ###
#####

You have chosen the analysis with covariates.

Starting time of data analysis (Year/Month/Day/Hour/Minute/Second):
2005 / 7 / 6 / 0 / 30 / 44
Ending time of data analysis (Year/Month/Day/Hour/Minute/Second):
2005 / 7 / 6 / 0 / 30 / 47

=====
== Summary Statistics ==
=====

-----
Filename of capture history/individual covariate data:  d:\software(no cov)\data\examp161.dat
Filename of occasional covariate data:  d:\software(no cov)\data\examp162.dat
Total number of distinct animals : 171
Number of capture samples : 10
-----

-----
j | u[j]  m[j]  n[j]  M[j]  ft[j]  f1[j]  n00[j] n01[j] n10[j] n11[j]
-----|-----
1 | 68    0    68    0    2    68    103    68    0    0
-----
```

2		33	27	60	68	62	74	70	33	41	27
3		26	36	62	101	40	74	68	43	41	19
4		12	40	52	127	31	65	76	33	43	19
5		15	58	73	139	16	54	60	59	38	14
6		3	38	41	154	13	45	79	19	51	22
7		12	64	76	157	5	41	70	60	25	16
8		0	35	35	169	1	26	77	18	59	17
9		2	74	76	169	0	9	72	64	23	12
10		0	38	38	171	1	2	75	20	58	18
11					171						

j		naa[j]	nab[j]	nac[j]	nba[j]	nbb[j]	nbc[j]	nca[j]	ncb[j]	ncc[j]
1		103	0	68	0	0	0	0	0	0
2		70	0	33	0	0	0	0	41	27
3		44	0	26	0	24	17	0	41	19
4		32	0	12	0	44	21	0	43	19
5		17	0	15	0	43	44	0	38	14
6		14	0	3	0	65	16	0	51	22
7		2	0	12	0	68	48	0	25	16
8		2	0	0	0	75	18	0	59	17
9		0	0	2	0	72	62	0	23	12
10		0	0	0	0	75	20	0	58	18
Sum		284	0	171	0	466	246	0	379	164

==== The Fit & Estimation of all models ====

Model	Estimate	S.E.	MIN(-LL)	AIC	95%CI (Log-transf.)	95%CI (Asy-normal)	Status
MM*0	173.99	1.83	1093.07	2188.14	( 171.99 , 180.02 )	( 170.41 , 177.57 )	Converge
MM*t	173.79	1.76	1071.43	2146.86	( 171.90 , 179.68 )	( 170.34 , 177.24 )	Converge
MM*b	173.32	1.62	1091.17	2186.34	( 171.67 , 178.99 )	( 170.14 , 176.50 )	Converge
MM*h	175.38	2.33	1080.36	2168.72	( 172.65 , 182.64 )	( 170.82 , 179.94 )	Converge
MM*tb	173.59	1.71	1071.25	2148.50	( 171.79 , 179.42 )	( 170.23 , 176.94 )	Converge
MM*th	175.14	2.26	1058.44	2126.89	( 172.52 , 182.26 )	( 170.71 , 179.57 )	Converge
MM*bh	174.40	2.05	1077.45	2164.90	( 172.14 , 181.14 )	( 170.37 , 178.42 )	Converge
MM*tbh	174.72	2.16	1057.90	2127.80	( 172.29 , 181.69 )	( 170.49 , 178.95 )	Converge

==== Model Description ====

The general logistic model MM\*tbh is

$$\text{logit}(P_{ij})=a + c_j + v * Y_{ij} + \text{beta} * W_i + r * R_j$$

where

- i : refers to the ith individual;
- j : refers to the jth sample or jth capture occasion;
- a : baseline intercept;
- c\_j : the unknown time or occasional effect of the jth capture occasion (set c\_t=0, where t: the number of capture occasions;
- v : (behavioral response) the effect w.r.t. the past capture history indicator Y\_ij;

```
beta      : the effect of individual covariates W_i;
r         : the effect of occasional covariate R_j;
```

```
==== The MLEs of Regression Coefficients ====
```

```
*** Model MM*0 ***
      a
MLE   -0.69
S.E.  0.05

*** Model MM*t ***
      a      r1(1)
MLE   0.31   -0.68
S.E.  0.16   0.10

*** Model MM*b ***
      a      v
MLE   -0.62  -0.22
S.E.  0.06   0.12

*** Model MM*h ***
      a      beta1(1) beta2(1) beta2(2)
MLE   -0.38  -0.28   -0.02   -0.46
S.E.  0.08   0.11    0.13    0.11

*** Model MM*tb ***
      a      v      r1(1)
MLE   0.31  -0.07  -0.66
S.E.  0.17   0.08   0.11

*** Model MM*th ***
      a      beta1(1) beta2(1) beta2(2) r1(1)
MLE   0.63  -0.28   -0.02   -0.47  -0.68
S.E.  0.18   0.11    0.14    0.12   0.11

*** Model MM*bh ***
      a      v      beta1(1) beta2(1) beta2(2)
MLE   -0.28  -0.28  -0.29   -0.03  -0.48
S.E.  0.09   0.11   0.10     .     0.11

*** Model MM*tbh ***
      a      v      beta1(1) beta2(1) beta2(2) r1(1)
MLE   0.65  -0.12  -0.29   -0.02  -0.48  -0.66
S.E.  0.18   0.12   0.11    0.10   0.12   0.11
```

In the outputs of covariates analysis, the version of software CARE-4, starting time and ending time of data analysis are shown first and then followed by the four main parts of CARE-4 outputs. They are (1) summary statistics, (2) model description, (3) the fit and estimation of capture recapture models, (4) the MLEs of transition probabilities.

The first part of the output provides basic information: (1) The data filename of capture history matrix and individual covariate data is c:\program

files\CARE-4\data\example2-1.dat for this example; (2) The data filename of occasional covariate data is c:\program files\CARE-4\data\example2-2.dat; (3) The number of distinct animals caught in the experiment is 171 in this case; (4) The number of trapping occasions is 10 in this case. Then the summary statistics are listed in order and the format is the same as the output of analysis without covariates.

For the data of Example 2, heterogeneity is explained by individual covariates. This is because adding the unknown time effects to our model in Equation (1) results in almost no change for the maximum likelihood; thus time effects can be fully explained by the observable characteristics of the trapping occasions. Huggins (1989) found that the behavioral response effect on any previous capture was not significant. Table 10 shows that model  $\mathbf{MM}^*_{th}$  has the lowest AIC under a framework of Markov chain model. The Markov chain models  $\mathbf{MM}^*_{tbh}$  and  $\mathbf{MM}^*_{th}$  have comparable AIC and all estimation results are similar under these two models. The fitted coefficient  $\hat{\gamma} = -0.12$  (s.e. 0.12) under model  $\mathbf{MM}^*_{tbh}$  indicates that there is no evidence to show that the animals exhibit a response to the capture on the most recent occasion. Therefore, we select model  $\mathbf{MM}^*_{th}$  and the conclusion is consistent with the finding of Huggins (1989) from a conventional approach.

For the time or occasional covariate under model  $\mathbf{MM}^*_{th}$ , the coefficient  $\hat{\gamma}_1 = -0.68$  (s.e. 0.11) shows that the capture probabilities are higher at night because night time is treated as a reference group. For gender effect, the negative regression coefficient  $\hat{\beta}_1 = -0.28$  (s.e. 0.11) demonstrates that the females have larger capture probabilities than the males. For age effect, the regression coefficient  $\hat{\beta}_{2(1)} = -0.02$  (s.e. 0.14) is not significant; hence there is no significant difference of capture probabilities between the young and the adults. However, the regression coefficient  $\hat{\beta}_{2(2)} = -0.47$  (s.e. 0.12) is significantly different from 0, which implies that adults have higher capture probabilities than semi-adults. The population size estimate under model  $\mathbf{MM}^*_{th}$  is 175.1 with an estimated s.e. of 2.3 and a 95% confidence interval of (173, 182). The results are slightly different from those obtained in Huggins (1989) due to different ways of treating animals with missing covariates. Although the Markovian dependence is not significant in this example, CARE-4 provides an alternative framework to assess the effects of animals' behavioral responses and relevant covariates.

## References

- Chao, A. (1987). Estimating the population size for capture-recapture data with unequal catchability. *Biometrics* **43**, 783-91.
- Chao, A. (2001). An overview of closed capture-recapture models. *Journal of Agricultural, Biological and Environmental Statistics* **6**, 158-75.
- Darroch, J.N. (1958). The multiple-recapture census I. Estimation of a closed population. *Biometrika* **45**, 343-59.
- Huggins, R.M. (1989). On the statistical analysis of capture experiments. *Biometrika* **76**, 133-40.
- Huggins, R.M. (1991). Some practical aspects of a conditional likelihood approach to capture experiments. *Biometrics* **47**, 725-32.
- Nichols, J. D., Pollock, K. H. and Hines, J. E. (1984). The use of a robust capture-recapture design in small mammal population studies: a field example with *Microtus pennsylvanicus*. *Acta Theriologica* **29**, 357-365.
- Otis, D.L., Burnham, K.P., White, G.C. and Anderson, D.R. (1978). Statistical inference from capture data on closed animal populations. *Wildlife Monographs* **62**, 1-135.
- Williams, B.K., Nichols, J.D. and Conroy, M.J. (2002). *Analysis and Management of Animal Populations*. Academic Press, San Diego, CA.
- Yang, H.-C. (2002). *The Applications of Markov Chain Models and Kernel Smoothing in Capture-Recapture Experiments*, PhD Dissertation, National Tsing Hua University, Hsin-Chu, Taiwan.
- Yang, H.-C. and Chao, A. (2005). Modeling Animals' Behavioral Response by Markov Chain Models for Capture-Recapture Experiments. *Biometrics*, **61**, 1010-1017.
- Zippin, C. (1956). An evaluation of the removal method of estimating animal populations. *Biometrics* **12**, 163-89.



## Appendix 1

In this Appendix, we give formulas for the estimators featured in CARE-4 under various models. Refer to Tables 4 and 5 for definitions and references.

### 1. Model $M_0$ (Otis et al., 1978; Darroch, 1958):

- Unconditional MLE:  $M_0$ (UMLE) [Back to Table5](#)

$$\text{Equation for } N: \frac{\partial \log L}{\partial N} = \sum_{j=1}^M (N - j + 1)^{-1} + t \log(1 - p) = 0,$$

$$\text{Equation for } p: \frac{\partial \log L}{\partial p} = \frac{n_{\bullet}}{p} - \frac{Nt - n_{\bullet}}{1 - p} = 0, \text{ where } n_{\bullet} = \sum_{j=1}^t n_j.$$

- Conditional MLE:  $M_0$ (CMLE) [Back to Table5](#)

$$\text{Equation for } N: 1 - \frac{M}{N} = (1 - p)^t,$$

$$\text{Equation for } p: \frac{\partial \log L}{\partial p} = \frac{n_{\bullet}}{p} - \frac{Nt - n_{\bullet}}{1 - p} = 0, \text{ where } n_{\bullet} = \sum_{j=1}^t n_j.$$

### 2. Model $M_t$ (Otis et al., 1978; Darroch, 1958):

- Unconditional MLE:  $M_t$ (UMLE) [Back to Table5](#)

$$\text{Equation for } N: \frac{\partial \log L}{\partial N} = \sum_{j=1}^M (N - j + 1)^{-1} + \sum_{j=1}^t \log(1 - e_j) = 0,$$

$$\text{Equation for } e_j: \frac{\partial \log L}{\partial e_j} = \frac{n_j}{e_j} - \frac{N - n_j}{1 - e_j} = 0, \quad j = 1, 2, \dots, t.$$

- Conditional MLE:  $M_t$ (CMLE) [Back to Table5](#)

$$\text{Equation for } N: 1 - \frac{M}{N} = \prod_{j=1}^t (1 - e_j),$$

$$\text{Equation for } e_j: e_j = \frac{n_j}{N}, \quad j = 1, 2, \dots, t.$$

### 3. Model $M_b$ (Otis et al., 1978; Zippin, 1956):

- Unconditional MLE:  $M_b$ (UMLE) [Back to Table5](#)

$$\text{Equation for } N: \frac{\partial \log L}{\partial N} = \sum_{j=1}^M (N-j+1)^{-1} + \sum_{j=1}^t \log(1-p) = 0,$$

$$\text{Equation for } \phi: \frac{\partial \log L}{\partial \phi} = \frac{m_{\cdot}}{\phi} - \frac{(M_{\cdot} - m_{\cdot})p}{1-\phi p} = 0,$$

$$\text{Equation for } p: \frac{\partial \log L}{\partial p} = \frac{n_{\cdot}}{p} - \frac{Nt - M - M_{\cdot}}{1-p} - \frac{(M_{\cdot} - m_{\cdot})\phi}{1-\phi p} = 0, \text{ where}$$

$$n_{\cdot} = \sum_{j=1}^t n_j, m_{\cdot} = \sum_{j=1}^t m_j \text{ and } M_{\cdot} = \sum_{j=1}^t M_j.$$

- Conditional MLE: Mb(CMLE) [Back to Table5](#)

$$\text{Equation for } N: N = \frac{M}{1-(1-p)^t},$$

$$\text{Equation for } \phi: \frac{\partial \log L}{\partial \phi} = \frac{m_{\cdot}}{\phi} - \frac{(M_{\cdot} - m_{\cdot})p}{1-\phi p} = 0,$$

$$\text{Equation for } p: \frac{\partial \log L}{\partial p} = \frac{n_{\cdot}}{p} - \frac{Nt - M - M_{\cdot}}{1-p} - \frac{(M_{\cdot} - m_{\cdot})\phi}{1-\phi p} = 0, \text{ where}$$

$$n_{\cdot} = \sum_{j=1}^t n_j, m_{\cdot} = \sum_{j=1}^t m_j \text{ and } M_{\cdot} = \sum_{j=1}^t M_j.$$

#### 4. Model MM<sub>1b</sub> (Yang and Chao, 2005):

- Unconditional MLE: MM1b(UMLE) [Back to Table5](#)

$$\text{Equation for } N: \sum_{j=1}^M \frac{1}{N-M+j} + t \ln(1-P_{01}) = 0$$

$$\text{Equation for } P_{01}: P_{01} = \frac{n_{01}}{t(N-M) + n_{0+}},$$

$$\text{Equation for } P_{11}: P_{11} = \frac{n_{11}}{n_{1+}}.$$

- Conditional MLE: MM1b(CMLE) [Back to Table5](#)

$$\text{Equation for } N: 1 - \frac{M}{N} = (1 - P_{01})^t = \left( 1 - \frac{n_{01}}{t(N-M) + n_{0+}} \right)^t,$$

$$\text{Equation for } P_{01}: P_{01} = \frac{n_{01}}{t(N-M) + n_{0+}},$$

$$\text{Equation for } P_{11}: P_{11} = \frac{n_{11}}{n_{1+}}.$$

### 5. Model MM<sub>1tb</sub> (Yang and Chao, 2005):

- Unconditional MLE: MM1tb(UMLE) [Back to Table5](#)

$$\text{Equation for } N: \sum_{i=1}^M \frac{1}{N-M+i} + \sum_{j=1}^t \ln[1-P_{01}(j)] = 0$$

$$\text{Equation for } P_{01}(j): P_{01}(j) = \frac{n_{01}(j)}{N-M+n_{0+}(j)},$$

$$\text{Equation for } P_{11}(j): P_{11}(j) = \frac{n_{11}(j)}{n_{1+}(j)}.$$

- Conditional MLE: MM1tb(CMLE) [Back to Table5](#)

$$\text{Equation for } N: 1 - \frac{M}{N} = \prod_{j=1}^t [1 - P_{01}(j)] = \prod_{j=1}^t \left( 1 - \frac{n_{01}(j)}{N-M+n_{0+}(j)} \right),$$

$$\text{Equation for } P_{01}(j): P_{01}(j) = \frac{n_{01}(j)}{N-M+n_{0+}(j)},$$

$$\text{Equation for } P_{11}(j): P_{11}(j) = \frac{n_{11}(j)}{n_{1+}(j)}.$$

### 6. Model MM<sub>2b</sub> (Yang and Chao, 2005):

- Unconditional MLE: MM2b(UMLE) [Back to Table5](#)

$$\text{Equation for } N: \sum_{j=1}^M \frac{1}{N-M+j} + t \ln(1-P_{ac}) = 0,$$

$$\text{Equation for } P_{ac}: P_{ac} = \frac{n_{ac}}{t(\hat{N}-M) + n_{a+}},$$

$$\text{Equation for } P_{bc}: P_{bc} = \frac{n_{bc}}{n_{b+}},$$

$$\text{Equation for } P_{cc}: P_{cc} = \frac{n_{cc}}{n_{c+}}.$$

- Conditional MLE: MM2b(CMLE) [Back to Table5](#)

Equation for  $N$ :  $1 - \frac{M}{N} = \left(1 - \frac{n_{ac}}{t(N-M) + n_{a+}}\right)^t$ ,

Equation for  $P_{ac}$ :  $P_{ac} = \frac{n_{ac}}{t(\hat{N}-M) + n_{a+}}$ ,

Equation for  $P_{bc}$ :  $P_{bc} = \frac{n_{bc}}{n_{b+}}$ ,

Equation for  $P_{cc}$ :  $P_{cc} = \frac{n_{cc}}{n_{c+}}$ .