Canadian Technical Report of

Fisheries and Aquatic Sciences 2106

1996

COMPUTER ANALYSIS OF DATA FROM STRATIFIED MARK-RECOVERY EXPERIMENTS FOR ESTIMATION OF SALMON ESCAPEMENTS AND OTHER POPULATIONS

by

A. N. Arnason[1], C. W. Kirby[1], C. J. Schwarz[2], and J. R. Irvine

Fisheries and Oceans Canada

Science Branch, Pacific Region

Pacific Biological Station

Nanaimo, British Columbia

V9R 5K6

[1] Department of Computer Science, University of Manitoba, Winnipeg, Manitoba, R3T 2N2
[2] Department of Mathematics and Statistics, Simon Fraser University, Burnaby, B.C. V5A 1S6

Cat. No. Fs 97-6/2106E ISSN 0706-6457

Correct citation for this publication:

Arnason, A. N., C. W. Kirby, C. J. Schwarz, and J. R. Irvine. 1996. Computer analysis of data from stratified mark-recovery experiments for estimation of salmon escapements and other populations. Can. Tech. Rep. Fish. Aquat. Sci. 2106: vi+37 p.

TABLE OF CONTENTS

Section **Page**

LIST OF FIGURES

Figure **Page**

LIST OF TABLES

Table **Page**

ABSTRACT

Arnason, A. N., C. W. Kirby, C. J. Schwarz, and J. R. Irvine. 1996. Computer analysis of data from stratified mark-recovery experiments for estimation of salmon escapements and other populations. Can. Tech. Rep. Fish. Aquat. Sci. 2106: vi+37 p.

This report describes the analysis of 2-sample mark-recovery experiments to estimate the number animals in populations where the marks and recoveries take place over a number of strata. Strata may be defined in time or in space or both, and the s strata in which marking takes place may differ from the t strata in which recoveries take place. The report also describes the use of a program called SPAS (Stratified Population Analysis System) for analyzing this type of data. The program runs under Windows on PC computers and under Xwindows on Sun workstations. SPAS computes the Darroch, Schaefer, and Pooled Petersen estimators as described in Seber (1982) and a new maximum likelihood estimator. Among other advantages, this Darroch-Plante estimator works when $s \neq t$. SPAS permits pooling and deletion of strata and carries out a number of tests for goodness-of-fit and the validity of pooling. It allows the user to carry out very general simulation experiments to investigate the effects of different population and sample sizes, different poolings, and of assumption failures like the occurrence of births and deaths on the precision and bias in the estimates.

Keywords: stratified populations, mark-recapture, escapement, marking, software, simulation

RÉSUMÉ

Arnason, A. N., C. W. Kirby, C. J. Schwarz, and J. R. Irvine. 1996. Computer analysis of data from stratified mark-recovery experiments for estimation of salmon escapements and other populations. Can. Tech. Rep. Fish. Aquat. Sci. 2106: 37 p.

Dans ce rapport, les auteurs analysent deux expériences de marquage et de recapture visant à estimer le nombre de sujets au sein de populations où les opérations de marquage et de recapture se font en plusieurs strates. Celles-ci peuvent être définies par des paramètres temporels, spatiaux ou spatio-temporels; les strates «s», qui correspondent au marquage, peuvent différer des strates «t», qui correspondent à la recapture. Le rapport décrit aussi un programme nommé SPAS (*Stratified Population Analysis System* ou système d'analyse de populations stratifiées) utilisé pour l'analyse du type de données recueillies par les auteurs. Le programme fonctionne sur PC configuré en Windows et sur poste de travail Sun configuré en Xwindows. Le SPAS calcule les valeurs prises par les estimateurs de Darroch et de Schaefer, par les estimateurs groupés de Petersen tels que décrits par Seber (1982), de même que celles prises par un nouvel estimateur du maximum de vraisemblance. Parmi d'autres avantages, il est à mentionner que cet estimateur Darroch-Plante est utilisable lorsque «s» diffère de «t». Le SPAS rend possible le groupage et la suppression de strates, et il effectue un certain nombre de tests pour vérifier la validité de l'ajustement et celle du groupage. Il permet à l'utilisateur de procéder à des simulations d'un caractère très général lui permettant d'examiner les effets, sur le plan de la précision et des erreurs systématiques des estimations, associés à différents niveaux de population et d'échantillonnage, à différents groupages et à l'infirmation d'hypothèses relatives, par exemple, au nombre de naissances et le nombre de décès.

Mots clés : populations stratifiées, marquage-recapture, échappées, marquage, logiciel, simulation

1 INTRODUCTION

This manual describes a program called SPAS (Stratified Population Analysis System) for the analysis of 2-sample mark-recapture experiments in stratified populations. The first sample is called the capture or marking sample. We use *s* and *t* to denote the number of initial and final strata, respectively. A sample is taken in each of the *s* initial strata, and each animal in the sample is marked and returned to its stratum so that its stratum of origin can be identified later. The second (final) sample is called the recovery sample. In each of the *t* final strata, the number of unmarked animals in the sample is recorded, as well as the number of marked animals from each initial stratum. This is a generalization to multiple strata of the Petersen method for estimating population abundance. The initial and final strata may be defined geographically or temporally, and the final strata may be quite distinct from the initial strata. The object of the experiment is to estimate the total number of animals over all strata, and the number of animals per stratum.

The software and manual were developed primarily to assist with the estimation of salmon escapement: the number of mature fish that avoid marine fisheries and return to freshwater to spawn. But this software can be used to estimate other populations as well. Escapement estimates are used in the assessment and management of the fishery resource and they also enable us to monitor effects of habitat change. A fuller description of escapement estimation experiments is given in the Background section below.

1.1 SPAS Capabilities

SPAS will read in and display data on sample sizes and marked recoveries for each initial and final stratum. It uses a scrolling window environment and allows user control over precision and field width so that very large data arrays (large *s* and *t*) can be accommodated. Data are read into the Analysis Data window where the user can temporarily alter the initial or final strata by pooling or deletion of selected strata. The original data array is always preserved so that the user can undo stratum redefinitions, but the data with altered stratum definitions can also be saved as a new permanent data file. SPAS then permits the user to analyze the data in the Data window using a number of analysis and testing methods: Darroch/Plante maximum likelihood estimates, Schaefer estimates, and pooled Petersen estimates (all described more fully below). Results of the analyses are placed in a Analysis Results window where they can be browsed and then saved to print files for sending to a printer or for importing to a word processor or spreadsheet program for formatting and printing. The Data and Results window in fact form the upper and lower pane, respectively, of a single split window, thus ensuring that the two are kept together. Each pane is independently scrollable to facilitate comparisons (for example, between observed and predicted counts).

SPAS also includes a simulation capability that provides a very powerful tool for planning experiments and assessing the properties of the estimates and tests. The simulator works by reading in a file specifying a hypothesized 'true' population. The user must specify *s* and *t*, the initial (true) stratum sizes, the capture and recovery probabilities for each stratum, and

the (vector of) migration rates from each initial stratum to the set of final strata. If this vector sums to less than 1.0, then there is mortality of the animals in that initial stratum between capture and recovery time. Such mortality is allowed in the simulation. The simulation works by reading in a Simulation parameter file into a Simulation Data window in much the same way as a data file is read into the Analysis Data window. That is, the Simulation file can be selected and read in, then browsed in full screen mode; and strata can be designated for pooling or deletion. Once a Simulation file has been loaded into the Simulation Data window, it can be used in one of four ways:

(1) It can be used to generate a data file for a single stochastic realization. That is, simple random samples are drawn with the appropriate sampling probability to represent the capture and recovery process, and animals are distributed over the recovery strata by drawing a sample from the appropriate multinomial distribution for each initial stratum. The result is stored as a data file in exactly the same form as real-world data would be. The program opens a new split window (titled Analysis: Single Replication) and displays the simulated data file in the Data pane where it can be further analyzed or saved just like a real-world data set with the same *s* and *t*.

(2) It can be used to generate a data file of mean values. No random sampling is done in this case; instead, the variables of the data file are replaced by the theoretical means of the sampling distributions in (1). These values will not usually be whole numbers, but otherwise the data file produced has the same form as real-world data with the same *s* and *t*. The mean value data are also placed in the Data pane of a new window (titled Analysis: Mean Values). The mean value data can then be saved or analyzed just like a real-world data set. Mean value analysis provides a quick method of assessing precision and bias in a proposed sampling experiment applied to a population whose size and structure can be guessed.

(3) It can be used to carry out replicated simulations with no stratum alterations (pooling or deletions). In this case, the user selects the analyses to be applied and the number, R, of replications to carry out. The system then carries out R replications of the single sampling experiment described in (1). The results of each replication are used to accumulate means and standard deviations (s.d.s) over the R replications of the estimates of population numbers and their standard errors. These results are formatted as closely as possible to the results of a single analysis (except that each result is replaced by 2 numbers: its mean and standard deviation over R) and placed directly in the Results pane where it may be browsed and/or saved. Replicated simulations can also be used to compute coverages of confidence intervals and powers of tests over the R replications, and some of these capabilities will be added in future releases of the software.

(4) It can be used to carry out replicated simulations with stratum alterations (poolings or deletions) applied to each replicate before the data are analyzed. When this option is selected, the system creates a new Mean Values split window (titled Pooled Simulation) and the user applies the desired alterations to the mean values. The system remembers these alterations and applies them to each of the R replicates. The replicate is then analyzed and used to collect means and s.d.s over the replicates as in (3). This allows the user to explore

the effects of different poolings and violations of the closure assumptions. Note that dropping initial or final strata is like starting the sampling of a run late or terminating sampling early, so violating the assumption that all fish in the run have a chance of being captured and a chance of being recovered.

SPAS uses multiple windows. There are split windows for Analysis and Simulation. The upper and lower panes are always for Data and Results, respectively. There is also a simple scrolling View window for browsing (read-only) of any input or output file. The user may open multiple instances of each window type although only one is active at any time; a Window list is maintained on the main menu to help users keep track of, and switch between, open windows. The system opens new windows automatically when required (cases 1, 2, and 4 above) and adds them to the Window list. There is an Edit menu item that brings up the system editor (Notepad in Windows) so the user can create new files or change existing files (e.g. to make a quick fix to some input data). Edit is handled as a spawned application and the windows created are not added to the Window list.

SPAS does not provide management and editing facilities for raw data. It simply imports a text file with the appropriate summary counts by stratum. Thus for example, SPAS will not accept capture histories for individual animals and produce the summary counts. It is expected that most biologists prefer to do this using some general data management and analysis program such as a spreadsheet program (e.g. Lotus 123, Excel) or a database program (e.g. Paradox, FileMaker Pro), rather than learn to use yet another set of editing and entry conventions. The Edit window in SPAS is an unstructured general full screen editor that is intended to let users create small input (Analysis Data or Simulation Data) files on the fly, or to make minor corrections to existing files that fail to load correctly. SPAS is, however, carefully designed to ensure that it can load data sets that have been exported by such programs, and we give advice and examples for doing this. The output from SPAS is also designed to be easy to re-export into these programs. Thus, for example, a big array of results could be saved and imported into Excel where the user could format it nicely, including adding figure captions, borders and notes, and then print it with the full control these packages offer over fitting the results onto the page or over how to break the array up over pages.

We now give a brief outline of the methods available for analyzing stratified 2-sample data.

1.2 Background

The first application of a stratified method to estimate salmon escapement was by Schaefer (1951). The strata were temporal: each initial stratum sample was the result of sampling for adult sockeye salmon over a week at a fixed point in a river. The $s = 8$ contiguous weeks of sampling covered the period of the spawning run of the salmon, so the method can be used to estimate the size of the run during each week of the run and the total size of the run. The final strata consisted of $t = 9$ weeks of sampling on the spawning grounds where marked animals are identified from the animals recovered dead on the spawning ground ("dead pitches"). This basic experimental design is still widely used for estimating run size of salmonids. We discuss the assumptions required for the analysis of data like this later in the manual.

If only the total run size is required, one can obtain an estimate by pooling data over the strata and using the Petersen estimate. The Petersen is formed from 3 statistics: the first sample size (total sample size over all the capture strata), the second sample size (total sample size over all the recovery strata), and the total number of marked animals recovered (in all the recovery strata, regardless of stratum of origin). When the Petersen is applied to stratified data that have been pooled in this way, it is called the pooled Petersen method (Seber, 1982). However, the Petersen estimate can be badly biased when some animals are more catchable than others, especially if marked animals have a different capture rate from unmarked in the recovery sample. Stratified experiments present many opportunities for this violation to occur since the proportion marked may vary across initial strata and animals from different initial strata may have very different chances of moving to and/or being sampled in the different final strata. This can induce differences in overall capture rates between marked and unmarked animals and result in bias.

The advantage of pooling strata is that it reduces the number of parameters that need to be estimated and so (generally) gives a more precise estimate of total population. An intermediate strategy is to pool together only some of the initial or final strata. For example, if one felt that the first 3 weeks had similar sampling fractions and migration rates, it might make sense to consider these 3 strata as a single stratum, thus reducing the number of capture strata (by 2). We call this selective pooling. Selective pooling can be applied to the final strata as well as to the initial strata. Another reason for using selective pooling is to avoid numerical problems created by small sample sizes or by violations and lack of model fit. These problems include failure of the estimation method (for example, some methods involve an iterative search for the estimate, and the search may fail to converge or the formulae may involve division by zero, etc.), or the occurrence of inadmissible estimates (for example, negative stratum sizes, or probabilities of capture greater than 1).

With many samples (e.g. when strata arise from daily samples), there are many possible poolings that may be done. Pooling intervals do not need to be equal; you could pool the first 7 days, then the next 3 days, and so on. The intervals used in pooling the final strata will typically be different from those used for the initial strata. In salmon runs, the recoveries begin later and pooling intervals may or may not coincide with the initial sample intervals; you could, for example pool initial strata by weeks and final strata by 3 day intervals. Strata may be defined using a dual classification of time and area (for example if recoveries are on 2 different spawning areas over two different time intervals) to allow for heterogeneity in sampling intensity on the two areas and to investigate migration rates to the different areas. The only requirement is that the strata be mutually exclusive (an animal can be marked in only one initial stratum and can be recovered in only one final stratum). At the moment, SPAS only lets you pool sequentially numbered strata (e.g. 1 with 2 and 3 but not 1 with 3 and 6), but this will be changed in later releases to allow any pooling. It is not yet clear what criteria should be used to determine the optimal pooling. Clearly, pooling should be done so that animals within a pooled stratum are as homogeneous as possible with respect to capture, migration, and recapture, but some lack of homogeneity (leading to increased bias) can be traded off for higher precision. Strategies for optimizing this trade-off, and tests for bias, are not yet fully developed despite a start being made by Warren and Dempson (1995) and Kirby (1996). Thus a major motivation for developing

SPAS was to let the user try different poolings quickly and easily and to explore the effects of different poolings using simulation.

We now review the main estimation methods for 2-sample stratified experiments and indicate which were implemented in SPAS ; a more detailed description of the estimates and tests and their properties will be given in the section 3 (Analysis and Testing Methods).

(1) Schaefer (1951) developed an estimate of the total population size, N, using ratio and expectation arguments. The Schaefer estimate is biased unless capture probabilities are equal in all initial strata or the recovery probabilities are the same in all the final strata. If either condition holds, the pooled Petersen will also be unbiased and will be more precise (it makes more efficient use of the data). No estimate of standard error (s.e.) is available for the Schaefer estimate. The Schaefer estimate was studied by Warren and Dempson (1995) using simulated sampling experiments. SPAS computes both the Schaefer estimate of N and the individual stratum estimates as given by Warren and Dempson (which they called a "simple daily estimator").

(2) Seber (1982) summarizes the early work of Chapman and Junge (1956) and Darroch (1961). From their work we know that there are 3 cases: (a) $s = t$; (b) $s < t$; and (c) $s > t$. In case (a) both the initial and final stratum (population) sizes can be estimated. In case (b) only the initial stratum (population) sizes can be estimated, and in case (c) only the final stratum sizes can be estimated. The total population size, N, is nevertheless estimable (apart from numerical problems that may arise) in all 3 cases but they gave an estimate and s.e. only for case (a). This estimate is commonly referred to as the Darroch estimate. They gave the necessary and sufficient conditions for the pooled Petersen to be unbiased and developed two chi-square tests for sufficient conditions (i.e., if the tests pass it should be safe to pool…if they fail it may or may not be safe to pool). We implement these tests and the Chapman and Junge estimate for case (a) which we call the Darroch Moment Estimate, since it is essentially obtained by equating the observed counts to the predicted counts and solving for the parameters of the model. This method only works in case (a) where the number of parameters equals the number of independent observations (apart from problems arising from missing observations or redundant strata).

(3) Plante (1990) developed an alternate maximum likelihood method for the Darroch estimate that can be applied to all 3 cases and we have re-implemented her estimates. The method is iterative and uses initial values computed using least squares. We report the least squares estimates (without standard errors). Then, if the iterative method works (and it may not, especially with small sample sizes and large numbers of strata), we report her estimate for N with its s.e., the stratum estimates (for those that are available depending on the case), and a goodness of fit test based on the deviation of observed statistics from their predicted values from the fitted model. We also report the maximized log likelihood which can also be used for testing.

(4) We report the pooled Petersen and its standard error, using the Chapman hypergeometric model as described in Seber (1982, p. 66).

1.3 Manual Outline

The outline for the rest of this manual is as follows: After explaining the notation used in this manual for the data, parameters, and estimates (1.4 Notation), we describe the form of the Analysis Data and Simulation Data input files and give hints on how to export from Lotus 123 or similar spreadsheet program to get valid input files. Readers who are impatient to start using SPAS can skip the rest of this section and go to section 2. You can read the remaining subsections of section 1 when you wish to prepare your own data for analysis, and the remaining sections of the manual when you want to learn more about the program and its methods.

In section 2 (Getting Started with SPAS), we describe how to obtain the files on a distribution disk or through the INTERNET and describe how to install them. We explain (2.2) the Main Menu and sub-menu items and what they do. We then give (2.3) an Introductory Tour that leads you through an example of launching the program, loading and analyzing actual data. Part two of the tour (2.4) leads you through examples of doing a simulation. The introductory tour uses input files supplied on the distribution disk. New users should install the program and carry out the sequence of instructions given in the tour, observing the results pointed out in the tour. By carrying out the tour, the user can get a good preliminary idea of how to use the program and the pitfalls to be avoided.

The remaining sections of the manual give reference material on the program and the methods used. Section 3 (Analysis Methods) describes the analyses implemented, including the assumptions required for valid estimates, methods for testing the assumptions, and the nature of the biases produced by assumption failures. Section 4 (Use of Simulation) gives some further pointers on how to use simulation to examine the properties of estimates and to help plan future experiments so that goals on precision can be met. Section 5 (Discussion) gives some notes on program operation and needs for further developments.

1.4 Notation

We use s and t to denote the number of strata at capture time and recovery time, respectively. We use superscript c and r to indicate statistics or parameters applying to these 2 sample times. For example, when populations are closed the total population size can be designated as N, but when the closure assumption is relaxed, the sizes at the 2 times can be designated by N^c and N^r. Bold is used to indicate vectors and arrays and dot notation indicates summation over strata.

Statistics:

n^c_i the number of animals taken (and released marked) in capture stratum i, i=1…s.

n^r_j the size of sample taken in recovery stratum j, j=1…t.

m_{ij} the number of the n^c_i that are recovered in stratum j.

u^c_i the number of animals marked in capture stratum *i* that are unseen (never recovered); $= n^c_i - m_{i.}$

u^r_j the number of unmarked animals in the sample in recovery stratum *j* $(= n^r_j - m_{.j})$.

Parameters:

N^c_i the size of the population in initial (capture) stratum *i*, $i = 1 \ldots s$.

N^r_j the population size in final (recovery) stratum *j*, $j = 1 \ldots t$.

p^c_i the probability that an animal in initial stratum *i* at capture time is captured in that sample; $i = 1 \ldots s$.

p^r_j the probability that an animal in final stratum *j* at recovery time is recaptured in that sample; $j = 1 \ldots t$.

θ_{ij} the probability that an animal in stratum *i* at capture time is in stratum *j* at recovery time.

1.5 Form of Input Files

The data must be supplied to the program in an ASCII text file (or it can be typed in on the fly from the Edit window) which is most easily prepared using a spreadsheet program like LOTUS 123 or MS Excel and then exported in text form. Any text editor that permits saving of a text file may also be used. The data are laid out in row and column format (Fig. 1); elements in the same row can be separated by blanks or commas (but you cannot mix blanks and comma separators in the same file) and all text strings must be delimited by double quotes. Using blank delimiters allows you to insert extra blanks to align the data for easier reading and to make its meaning clear. Fig. 1 shows the symbolic layout for the data followed by a numeric example with $s = 3$, $t = 4$. The size of *s* and *t* is limited only by available computer memory. Both *s* and *t* must be integers but all other numbers can be integer or real. Reals are useful when counts have been adjusted for marking effects like tag loss and when specifying expected m_{ij} as a planning device.

If you are using LOTUS to create the data you should note the following:

1. Put the title in cell A1. Type ""text" to get quoted text as the first quote is interpreted by 123 as an alignment directive.
2. If you need more than one screen width (80 characters) to enter the data, reset the right margin as far to the right as possible (type /WGDPR then type 240). Make sure that all your columns are wide enough, as the exported file will be exactly as you see it, including truncated entries in columns that are not wide enough.
3. When the data are typed in, export using Print to File (type /PF then type or choose a file name (e.g. MYDATA.DAT), then type R and specify the range for your data (e.g. A1..F7 for the example in Fig. 1), then type OOUQ to specify unformatted output and return to the print menu, then type G to do the save).

DATA File: Symbolic Layout

"Title for the data, enclosed in quotes"

	s	t				
			"Col 1"	"Col 2"	…	"Col t"
"Row 1"	n^c_1		m_{11}	m_{12}	…	m_{1t}
"Row 2"	n^c_2		m_{21}	m_{22}	…	m_{2t}
:	:		:	:		:
"Row s"	n^c_s		m_{s1}	m_{s2}	…	m_{st}
			n^r_1	n^r_2	…	n^r_t

DATA File: Example

"Example where stratification is by time"

3	4				
		"May 14"	"May 21"	"May 28"	"June 14"
"May 4"	1000	45	24	12	3
"May 12"	2000	0	123	42	12
"May 24"	1000	0	0	16	8
		400	892	634	400

Figure 1. The form of the input file required by the Analysis…*Open* or *Load* operation.

If you are using EXCEL to create the data, type the data file as shown into cells A1 to F7 and use Save As from the File menu but be sure to save as type Formatted Text from the "Save File As Type" selection list at the bottom of the dialog box. Older versions of Excel don't have Formatted Text as an option so you will have to use Text; however, the quotes on the quoted strings will be replaced by triple quotes in the text file. You will have to use a text editor to fix these or the file will not import correctly.

If you are using another spreadsheet program to produce input files, refer to your manual for the procedure to produce ASCII files. Note that input files will be more (human) readable if displayed in a fixed pitch font such as Courier and if extra spaces are inserted to make the column labels and column values line up. Don't use tabs for alignment as these will cause problems when you load the file into SPAS.

SIMULATION File: Symbolic Layout

"Title for the Simulation, enclosed in quotes"

s	t					
			"Col 1"	"Col 2"	…	"Col t"
"Row 1"	N^c_1	p^c_1	θ_{11}	θ_{12}	…	θ_{1t}
"Row 2"	N^c_2	p^c_2	θ_{21}	θ_{22}	…	θ_{2t}
:	:	:	:	:		:
"Row s"	N^c_s	p^c_s	θ_{s1}	θ_{s2}	…	θ_{st}
			p^r_1	p^r_2	…	p^r_t

SIMULATION File: Example

"Simulation Example with no death"

3	4					
			"Rec 1"	"Rec 2"	"Rec 3"	"Rec 4"
"Cap 1"	5000	0.3	0.25	0.25	0.25	0.25
"Cap 2"	5000	0.5	0	0.34	0.33	0.33
"Cap 3"	5000	0.7	0	0	0.50	0.50
			0.35	0.50	0.35	0.25

Figure 2. The form of the input file required by the **Simulation** …*Open* or *Load* operation.

The input file for a simulation (Fig. 2) is similar in form to that for data (Fig. 1). The **Simulation**…*Open* or *Load* operation also requires an ASCII text file that, in this case, specifies the initial stratum population sizes (N^c_i), the capture (p^c_i) and recovery rates (p^r_j), and the migration (and survival) rates (θ_{ij}) between strata. The sum of migration rates in row *i* give the overall survival rate of animals initially in stratum *i* and this may be equal to or less than 1.0. If any row of migration rates sums to more than 1.0, it is an error and the file will be rejected. These input files can also be created using a spreadsheet program and saved as text files as described above. In the case of simulation input files, an advantage of using the spreadsheet is that the user can add rows to compute the sample sizes (the m_{ij}, the n^c_i, and the n^r_j) given the probabilities (the θ_{ij}, the p^c_i and the p^r_j) or one can specify the sample sizes and have the spreadsheet compute the probabilities. The former is accomplished within SPAS using the Mean Value simulation, but there is no capability corresponding to the latter method. We include an example worksheet for users interested in this method (DARROCHB.WK1 on the distribution diskette…it can be loaded by both Lotus 123 and MS Excel).

2 GETTING STARTED WITH SPAS

2.1 Software Acquisition, Installation, and Startup

Currently SPAS versions exist for MS Windows 3.1 (on DOS or OS/2 machines) and for the Xwindows/UNIX environment (on Sun Workstations). The names of the files you need to run SPAS and the introductory tour are the same for both, but we give separate installation instructions for the different platforms. To obtain a copy of the MS Windows software on diskette, send a formatted high density diskette (1.2 Mbyte or 1.44 Mbyte floppy) in a self-addressed, re-useable mailer envelope to the first author. Alternately, you can obtain the distribution files over the INTERNET. For instructions, connect to the population analysis web site maintained by the first author (http://www.cs.umanitoba.ca/~popan) and click on the SPAS link. It is a good idea to check this site from time to time for updates to the software and documentation.

The SPAS distribution files consist of one or more compressed (.ZIP) files and the decompression program PKUNZIP. There will also be a README file. List this file on the screen or print it out: it will give you the instructions to follow for decompressing the files and running the install program. After installation you should have the following files in a directory called SPAS:

- SPAS.EXE The SPAS executable
- SCHAEFER.DAT A sample analysis data file (data from Schaefer, 1951)
- CONNE.DAT Another analysis data file (data from Dempson and Stansbury, 1991)
- SIMTEST.SIM A sample simulation data file
- DARROCHB.SIM A large simulation data file
- DARROCHB.WK1 The Lotus worksheet used to generate DARROCHB.SIM

2.1.1 Installing and de-installing SPAS for Windows

The Windows version of SPAS requires that you have upgraded your Windows 3.1 to Windows 32S (32-bit addressing, flat memory model). We provide the standard upgrade kit (WIN32S.ZIP) with the distribution files, but you do not need to use it if you have already upgraded (check to see if you have a subdirectory called WIN32S in your \WINDOWS\SYSTEM directory). SPAS is installed by decompressing the SPAS.ZIP file and running the winstall program (follow directions in the README file in the distribution file set). This will decompress the SPAS program and files (listed above) and place them on your hard drive. It will also create a Group called SPAS on your desktop and place the program icon in that Group. If you don't want a separate Group for SPAS you can later drag the icon to another existing Group (e.g. if you have a Group folder for Applications), then click on the SPAS Group and select Delete from the File Menu of the Program Manager.

To de-install SPAS, simply delete all the files in the \SPAS directory; then use the Program Manager to delete the SPAS Program Item (icon) and Group. The SPAS de-install is complete because the install does not alter system files or place any files outside its own directory.

To launch SPAS, double click on its icon or type C:\SPAS\SPAS.EXE at the prompt from the File...Run command of the Windows Program Manager.

2.1.2 Installing SPAS for UNIX Graphical Workstations

SPAS is written using standard X11 libraries compiled for a Sun SPARC and should work on any graphical workstation or Xterminal using a GUI (Graphical user interface) that uses these libraries: e.g. Motif, OpenLook. Create a directory called spas and copy the files on the distribution diskette to this directory. To launch, you must have activated an Xwindows session and window manager (mwm, olwm). From your console or Xterm window, locate to the .../spas directory and type, at the command prompt: spas &

2.2 The Main Menu

This description of the menus (and the tour that follows in the next sections) is for the Windows version of SPAS. The menu items and program functionality are the same for other versions but there may be differences in the appearance and management of Windows. Users of other systems can follow through this same tour making the obvious adaptations to meet the windowing conventions of their window manager.

Recall (section 1.1) that SPAS uses multiple windows. There are single windows for editing and viewing files, and split windows for doing an Analysis or Simulation. The top pane is used for Data and the bottom pane for Results. When SPAS is first launched, an empty window containing the main menu appears. This is shown by the outer window in Fig. 3. Note that the title bar contains the title SPAS. We will refer to this main window as the SPAS window.

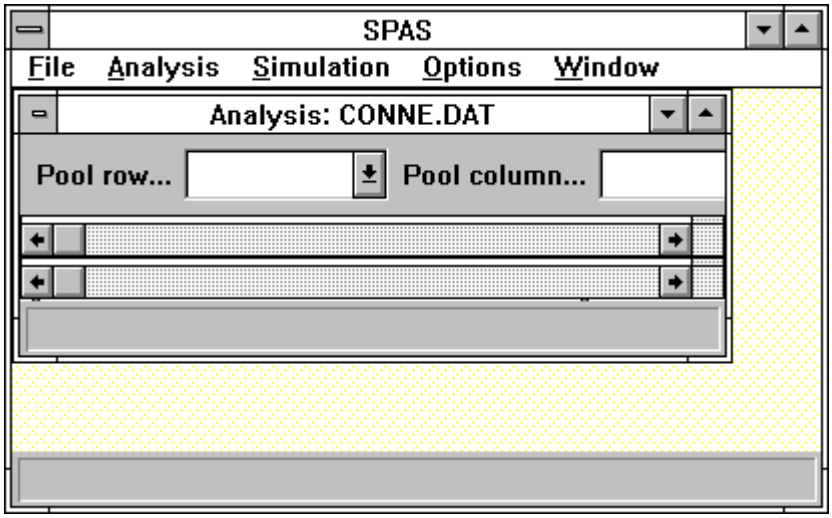

Figure 3. The initial SPAS screen after using the Analysis menu item to Load a data file

The main menu items are File, Analysis, Simulation, Options and Window: Briefly, these drop down menu items do the following:

File allows the user to open Edit and View windows. These are simple scrolling windows for editing and browsing data files. They may use system utilities (e.g. the Note Pad accessory in Windows). The last item on the menu is always Exit, which closes all open windows (contents are lost) and quits SPAS. When a split window is active, the File menu also includes items for Save As, Print Data, Print Results, and Close.

Save As is for saving the contents of the Data pane. It saves the data as a SPAS-readable file that can later be re-loaded. It is particularly useful for saving data that have been manipulated by pooling or deletions of rows and columns so that these operations need not be repeated.

Print is for writing the contents of the Data or Results window to a text file. These files are not SPAS-readable and they may have some very long lines that will make them difficult to read if sent directly to your printer. With a bit of editing, they can be imported to a spreadsheet program if you want to format them and print them nicely.

Close closes the split window, and the contents of both the Data and Results pane are lost.

Analysis initially contains only Open, to let the user open a new Analysis split window. When a new window is Opened, the new split window is created but immediately a standard file navigation window comes up on top of it because a split window should not have an empty Data pane. The user can browse for and select the data file to be loaded, after which the file navigation window goes away and the data are copied (with some re-formatting and addition of summary statistics) from the specified file into the Data pane. A split window is always opened in minimized form as shown in Fig. 3 and its contents won't be visible until it is resized. The use of an Analysis window, once data have been successfully loaded, is described more fully in the next section.

When an existing Analysis window is the active window, this menu item shows both Open and Load and a list of Analysis options. When you choose one of the Analysis options, that analysis is applied to the current data in the Data pane and the results are written to the Results pane of the window (obliterating any previous contents). Load lets you reload data into the Data pane, obliterating the current contents of the Data pane. Note that the Load does not change the contents of the Results pane, so the Data and Results may be "out of sync" until the next Analysis is run.

You can close the currently active window (contents of both the Data and Results pane are lost) by choosing Close from the File menu or by using Close on the control menu for the window (in Windows, the control menu drops down by clicking the Control Menu Box icon that looks like a file drawer at the upper left corner of the window; in Xwindows, you may need to right click in the window or menu bar).

Simulation operates like Analysis: that is, if no Simulation split window is Open, the menu shows only Open and choosing it brings up the standard file dialog box. When a Simulation window is the active window, this menu shows both Open and Load and a number of Simulation options. Only the Replicated option causes results to be written directly to the Results pane. All the others cause simulated data to be written to the Data pane of a new split window that is opened on top of the current window.

Options is a context sensitive menu that contains various program options. They typically affect only the currently active window and all subsequent operations. Precision sets the display precision in the Data or Results pane (it has no effect on the internally stored precision). If the active window is an Analysis window, the Restore option lets you go back to the originally loaded data before poolings or deletions. If the active window is a Simulation window, the Seeds option lets you manually set the seed used in the random number generators (see section 4 for more on use of the seeds). If either type of split window is active, the Display Covariance option gives the complete variance-covariance array for the parameter estimates.

Windows gives a list of the currently open windows by type (View, Analysis, Simulation) and the data file currently loaded to it. Edit windows, being independently launched applications, may not be included in the list. The list is automatically updated whenever a new window is opened or new data is loaded to an existing window. The user can select a window from the list to change the active window. The window to be made active can also be selected by clicking on it, if it is visible.

2.3 The Tour Part I: Using Analysis

This section will lead you through some analyses of the distribution data sets to illustrate window management and the use of Analysis in SPAS. Start SPAS up and obtain the main SPAS window as described in the previous section. Throughout the rest of this section, main menu items will be printed in **bold** and sub-menu items will be printed in *italic* text. The displayed text in a window or dialog box will be indicated in `fixed pitch font`.

2.3.1 Windows and Files

Before loading a data file, it is a good idea to preview it with View to check that it conforms to the correct input format. Select *View…* from the **File** menu, and select the `conne.dat` file from the standard file dialog. You should then get a window within the main window whose title is `View:CONNE.DAT`. Scroll through the contents of this window vertically (and horizontally if necessary). You should see that the file contains a title `Conne River 1992` on the first line, followed by the dimensions of the data arrays on the second line (6 rows by 6 columns). Following that are some simple column labels on line 3, followed by 6 rows of row labels and data. Finally, on the last line of the file, are the total numbers of animals recovered in each of the 6 recapture strata. You can compare this to Fig. 1 and see that this is the correct layout. These data are taken from Table 1 of Dempson and Stansbury (1991) and were

also analysed in Plante's thesis (1990). (The data are in fact from 1987; the title is in error). When you are through, close the View window (using **File**…*Close*).

Now we will load these data into a split window in order to do an Analysis. Click on the **Analysis** menu item and then on *Open…* A split window will be placed within the main window and immediately a standard file dialog box will open up on top of it. Choose the `conne.dat` file. When the dialog box goes away, your SPAS window should look exactly like Fig. 3: inside the main window is the minimized split window containing the conne.dat data, but the data are not visible until you resize the windows. First resize the main window by clicking on its maximize button (the up-arrow button at the top right corner of the outer window) or, as has been done in Fig. 4, by dragging the corners to enlarge the window. Then click on the maximize button of the split window whose title bar is Analysis: CONNE.DAT. Notice (Fig. 4) that when you maximize a child window it takes over the entire window of the parent window and its title is incorporated into that window's title bar. The two panes of the window should now be clearly visible, with the data in the upper (Data) pane and the lower (Results) pane empty. It is a good idea to maximize the window you are currently working with and minimize it when you are finished and ready to switch to another window. The minimize button is the double-arrow button on the upper right (Fig. 4). If both the main and the Analysis windows are maximized, then minimize only the Analysis window (the lower of the two minimize buttons). Try maximizing, minimizing and resizing both windows, and scroll through the Data pane, until you are familiar with their appearance and operation. Note that you cannot resize the individual panes of the split window: each one will always occupy half the available window height, even if one pane is empty.

As further practice with window manipulation, minimize the split window and re-open the View window (using **File**…*View*). Compare the contents of the raw data in the View window with its formatted look in the Data pane (use the Window menu to switch between windows and use resizing to make the contents visible). Notice that while the loaded Analysis Data closely resembles that in the View window, there are some differences. For one, we can see by scrolling over to the far right of the Data pane, that SPAS has calculated the number of animals released and never recovered, and scrolling down to the bottom, that SPAS has placed in the second last row the number of unmarked animals caught in each recovery stratum.

2.3.2 A First Analysis

Now restore your main window so that it looks more or less like Fig. 4 (by maximizing the Analysis split window) and we'll analyze the data by selecting the *All* option from the **Analysis** menu. Since we'll be running an iterative Darroch analysis, you will be presented with a dialog box to enter some parameters for the analysis. For this example, simply accept the defaults by selecting the OK button, or hitting Enter (see section 3.2.1 for discussion of these defaults). You should then see the word `Running…` in the status bar of the Analysis window (the Results window may look messed up and the scroll bars may go crazy while this is happening); wait until the status bar (Fig. 4) is clear before proceeding.

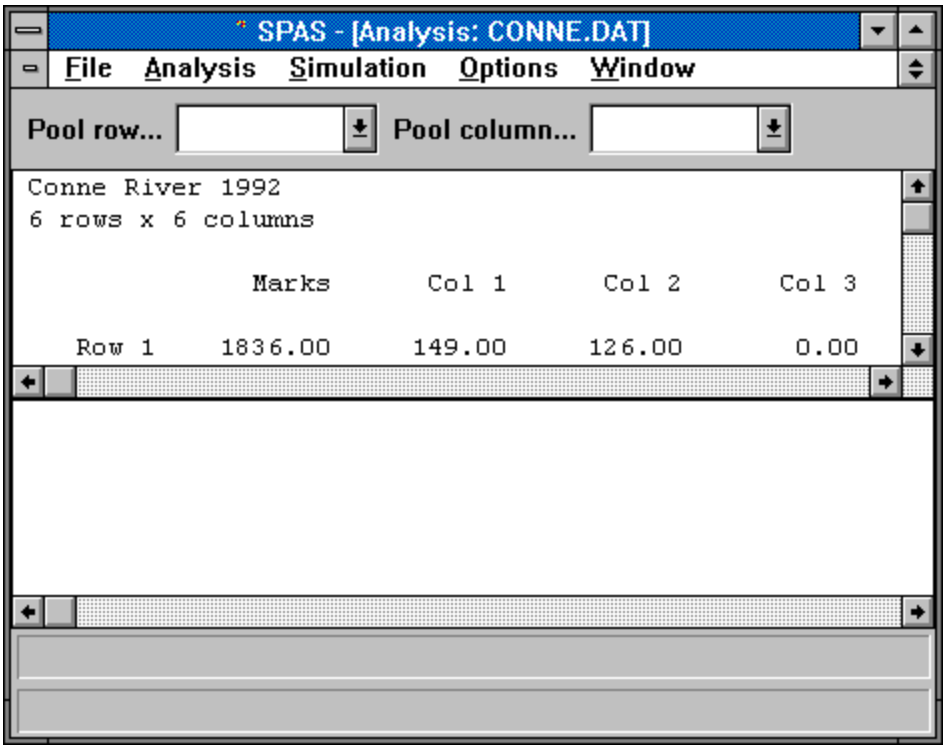

Figure 4. The initial SPAS screen from Fig. 3 after resizing the main (SPAS) window and maximizing the ANALYSIS window. Note the vertical scroll bar in the upper (Data) pane and the two status bars at the bottom of the window: the upper one belongs to the Analysis and the lower to the main SPAS window.

Once the analyses have completed, the Results window will contain the results of the analyses. Your Results window may still look empty but that is because it is positioned at some blank lines at the end of the Results. Scroll back to the top of the Results pane (drag the box on the vertical scroll bar back up to the top as in Fig. 4) and scroll through the results line by line (by clicking repeatedly on the down arrow). There are five analyses, each one beginning with the title line: in this case `Conne River 1992.` We will not go into these analyses in detail as this is done in section 3. However, we can note the following:

1. The Results pane begins by giving chi-square tests for complete mixing of animals across final strata independent of their initial stratum and for equal proportions of marks in the final strata. Passing either of these tests (i.e. $p > 0.05$) is sufficient (but not necessary) for the validity of full pooling. The significance value gives the probability of observing chi-square larger than the calculated value given validity of the pooling hypothesis. Since these probabilities are very small here ($p < 0.01$), it is possible that partial or fully pooled (Pooled Petersen) estimates are biased.
2. The ML Darroch estimate section shows the iterative search succeeded and produced a very precise estimate of total population size (71,127 with an s.e. of 2,246). This section includes Plante's Goodness of fit test (G-square, and its equivalent chi-square form) and a table of predicted counts (for both marked and unmarked animals). With a square data array (here 6 x 6) the model has as many parameters as free observations so there are 0 d.f. for the test and

the expected counts are identical to the observed. You can align the two arrays in the two panes and check this for yourself. We report estimates of the initial stratum sizes and capture probabilities when $s \leq t$ although we have not yet worked out the s.e.'s for these.

3. The Darroch moment estimate section is only available for square data arrays. It gives the point estimates for initial and final stratum sizes and their capture probabilities. The array elements are the N_{ij}, the estimated numbers of the N animals that are in initial stratum *i* and final stratum *j*. The point estimate is generally very close to the ML value but its s.e. is likely less accurate than the ML estimate.
4. The least squares estimator is the starting iteration for the ML estimator. It is again identical to the ML value in this case (because the data are square and non-singular).
5. The Schaefer estimate (73,433) is somewhat higher than the ML estimate (71,127) and no s.e. is available. Initial and final stratum size estimates and their capture probabilities are given. These are available for any *s, t* but are often inadmissible and s.e.'s are not known.
6. The Pooled Petersen estimate generated a total population estimate of 71,956 with a s.e. of 1,969. These are both close to the ML estimate so perhaps, despite the failure of the tests, there are other reasons why complete pooling is valid in this case. Usually the Pooled Petersen is much more precise than the ML estimate, though more prone to bias.

Next, we will attempt some partial poolings of the data and see if they appear to be justified.

2.3.3 Pooling Data

Pooling rows or columns may be necessary to avoid small sample and numeric problems that cause the ML iterations to fail to converge. That is not the case with the Conne River data but it may still be worthwhile to attempt poolings to increase precision and explore possible sources of difference among capture rates. The criteria for pooling without introducing serious bias are not yet fully understood, although it is likely that the criteria for pooling rows are different from the criteria for pooling columns. As discussed in section 3, it appears to be safe to pool columns if the recovery strata have very similar marked fractions or recapture rates. If you look back at the Darroch moment estimators, the last row in the Table of Stratum Estimates indicates that the recapture rates are similar for the strata labeled Col 2 through Col 5, but very different in the first and last stratum. We will attempt some pooling over various column combinations to show the effect of pooling.

Pooling and deletion of rows and columns are done from the `Pool row…` and `Pool column…` drop down menus above the Analysis Data pane (Fig. 4). Follow these steps to pool columns 2 and 3:

- Click on the arrow to the right of the `Pool columns…` box and a drop down list of the column labels appears.
- Click on the second label (Col 2). A dialog box appears with radio buttons for Expand, Drop, and Pool.
- Click on the Pool button. Then click in the `With next…` window and type 1. Click the OK box to dismiss the dialog box.
- Another dialog box appears prompting you for a label for the pooled column. Although the dialog box already contains a default value (in our example, it will be "Col 2"), it is recommended that you change the label to indicate that the column is pooled data. Click at the end of the default label and type `+3` so the label reads `Col 2+3.` Click the OK box.

- You'll notice that after you complete the pooling, the data array size below the data title automatically updates to reflect the pooling (to 6 rows by 5 columns). Now choose the *ML Darroch* method from the **Analysis** menu and examine the resulting analysis, again using the default parameters.

The pooled analysis results are relatively unchanged from the unpooled analysis. You should have an estimate (s.e.) of 72,068 (2,216) and the G2 goodness of fit test gives a value of 3.01 on 1 d.f. for an actual significance of p=0.08. The pooling appears to be acceptable.

Now continue by using the methods above to pool columns 4 and 5. Don't forget to choose the *ML Darroch* method from the **Analysis** menu after doing the pooling to regenerate the corresponding results in the Results pane. This 6 row by 4 column array should give an estimate of 71,916 (2,196) and a G2 of 3.27 (2 d.f.) for a p= 0.195. Recall that you can change the displayed precision of the data or results by selecting **Options**...*Precision* from the main menu.

Finally, note that you can pool previously pooled columns: try pooling the middle 2 columns of the 6 by 4 array to give a 6 by 3 array. The middle column now represents the pooling of the original Col 2 through Col 5 and you should get an estimate of 71,752 (2,056) and a G2 of 3.31 (3 d.f.) with p=0.346. If you make a mistake while pooling the data, just select *Restore* from the **Options** menu and start again.

It appears that all these poolings are acceptable. To see an example of an unacceptable pooling *Restore* the data to its unpooled form, pool `Col 1` with the next (1) column (call it `Col 1+2`) and re-run the **Analysis**. You should see that the estimate jumps up considerably (to 75,134 and an s.e. of 2314) and the G2 indicates a lack of fit (34.83 with 1 d.f., which is significant at $p < 0.01$).

2.3.4 Expanding pooled rows and columns

The pooling operations described above can also be applied to any set of rows. You can apply several different poolings to rows and columns in any order before doing the Analysis. You can also back out of a given pooling by selecting a pooled row or column and choosing expand. (Note that it is important that you keep track of what poolings have been done by using meaningful row and column labels.) This lets you try different combinations of poolings to see which yield the best results. In general, the user will want to find a pooling that gives admissible (non-negative) estimates, and then to continue pooling to increase precision provided there is no evidence of unacceptable bias or lack of fit.

Note that while pooled rows and columns can be pooled with other rows, including other pooled rows, expanding pooled data always reverts back to the component rows of the original data, not to any intermediate poolings. As an example, pool the column labeled `Col 1+2` with the next 1 columns. Name the pooled result `Col 1+2+3`; bring up the pooling dialog box again and expand the column. The pooled column will be replaced in the data set by the original columns 1 through 3, not `Col 1+2` and `Col 3` as might be expected.

2.3.5 Removing, Restoring and Saving Data

In addition to allowing you to easily pool data, SPAS also allows you to remove unwanted strata from a loaded data set. You might do this if a stratum sample is small and has an atypical capture rate, or as a means of avoiding singularity in the data matrix. Both rows and columns, pooled or not, can be dropped one at a time by selecting the row or column and clicking on *Drop* in the dialog box. The effect of this is equivalent to not having carried out the marking sample (if a row is dropped) or not having carried out the recovery sample (if the column is dropped). If there are substantial numbers in the deleted strata and these animals could not have been captured or recovered in any other (non-deleted) stratum, then you run the risk of introducing substantial non-closure bias (see section 3 for more on sources of bias).

If you wish to recover dropped rows or columns or revert the data set to its original form, you can simply select the *Restore* option from the **Options** menu. When this option is selected, the data set is restored to its original form and all drops, as well as poolings, are lost. Unlike poolings, drops cannot be undone selectively.

SPAS lets you save the current data array. This is done by choosing *Save As* from the **File** menu when you have an Analysis split window active. This brings up the standard file dialog box allowing you to select an existing file name or to type a new name. The data are then written out to that file, overwriting the contents of the file without warning (!!) if the name chosen already exists. We suggest you use the extension `.dat` when choosing a file name to remind you that this is a SPAS-loadable data file. Only the contents of the Data pane are saved. The name on the title bar of the active window changes to reflect the new name.

SPAS also lets you save the contents of either the Data pane or the Results pane of an Analysis split window (using **File**…*Print data* and **File**…*Print Results*, respectively). The data, saved as straight text just as they appear on screen, are intended for sending to a printer (if the lines are not too long) or for importing to a spreadsheet if you want to format the text before printing. Saving the data using *Print…* does not change the window name and the data file is not re-loadable. We suggest using the extension `.prt` or `.txt` to distinguish these files. The saved Data or Results print files can be browsed later from a **File**…*View* window. As an exercise in saving and using multiple windows, try the following exercise:

1. Maximize the main SPAS window to give yourself the whole screen area. Make the Conne River Analysis split window active. Use **Options**…*Restore* to return to the unpooled data or re-load the data (**Analysis**…*Load*) into the current window.
2. Pool the middle columns (Col 2 with the next 3) and do an ML analysis (**Analysis**…*ML Darroch*).
3. Save the pooled data (**File**…*Save As*) to a file called `pconne.dat`. Use **File**…*View* to browse the file (lack of alignment and unprintable tab marks between fields may make the file a bit difficult to read). Close the View window (**File**…*Close*).
4. Open a second analysis window (**Analysis**…*Open*) and load the `pconne.dat` data. Maximize the window and scroll through the Data pane. Do an ML analysis as in step 2.

5. Using the **Window** menu to switch between active windows, compare the Data and Results panes of the `conne.dat` window with the corresponding panes of the `pconne.dat` window: they should be identical. When you are done, close the split windows.

At this point, you may either continue on with the tour, or exit SPAS and continue the tour at another time. To quit SPAS, select the *Exit* option from the **File** menu. Be sure you have saved or printed any data or results you want to keep from *any* of the open windows, because the contents of *all* windows is lost (without further warning!!) on Exit.

2.3.6 Analyzing the Schaefer Data

The Conne River data is very well behaved in the sense that almost every possible pooling gives ML estimates and for most of these, the model exhibits good fit. This is probably because of the strongly diagonally dominant data array and the almost constant recovery probabilities. In addition, this experiment encompassed the entire run and is probably not subject to serious bias due to lack of closure. The closure assumption is discussed in section 3. A much more problematic data set is the Schaefer data supplied in the file `schaefer.dat` These data were extensively analyzed by both Darroch (1961) and Seber (1982), although their analyses produced inadmissible estimates. In this (optional) part of the tour, we will use pooling to get around the numeric problems and examine the tests to see if the pooled estimates are likely biased.

1. Close all windows and maximize the main SPAS window. Open a new split window (**Analysis**…*Open*) with the `schaefer.dat` file and maximize the split window. Scroll through the data pane and notice that there are some quite low counts in this 8 row by 9 column data array.
2. Choose (**Analysis**…*All*) and wait for the results to appear in the Results pane (may take several minutes on 386 and slower machines). Use **Options**…*Precision* to set the field width in the Data pane equal to that for the Results panes to facilitate some of the comparisons below. (NOTE: Changes to Data precision take place immediately and the Data pane is re-written; changes to the Results pane do not take effect until the next analysis is done.) Scroll through the results and note the following:
   - The iterative Darroch estimate failed to produce an estimate.
   - The complete mixing test gives a low chi-square value (18.71 with 7 d.f.); even though this is significant ($p = 0.01$), the low value indicates there may be redundant strata.
   - The equal proportions test fails (chi-square of 141.93 with 8 d.f., $p < 0.001$) indicating that the pooled Petersen and Schaefer estimates may be significantly biased.
   - The least squares estimator did produce an estimate that was, unfortunately, inadmissible, since it produced negative initial stratum size estimates. The negative stratum size estimates were caused by the negative capture probability estimates, directly to the right of the stratum size estimates.
   - As we scroll further to the right, we can see that the expected number of m_{ij}'s don't differ greatly from the observed numbers but that the predicted unmarked animals do. The first few (column) strata account for most of the contribution to the (significant) G2 test value.
   - The Schaefer estimate generated a total population estimate of 47,886.

- The pooled Petersen estimate generated a total population estimate of 47,278 with a standard error of 1,779.8.

Next, we will attempt to eliminate the failure and inadmissible estimate problems by pooling the data. Carry out the following poolings:

- Pool row 1 with the next 2 rows; label it PR1-3.
- Pool row 6 with the next 2 rows; label it PR6-8.
- Pool column 1 with the next 1 columns; label it PC1-2;
- Pool column 7 with the next 2 columns; label it PC7-9

Check that you have reduced the data array to 4 rows by 6 columns before running the complete analysis again. Scroll through the results and note the following:

- The Complete Mixing test has gone up relative to its d.f. (16.91 with 3 d.f.), indicating that some redundancy has been removed by pooling the data array.
- The Equal Proportions test is still highly significant (141.6 with 5 d.f.), indicating that pooling may not be valid.
- The ML Darroch estimate now converges and gives admissible estimates. The population estimate and its s.e.: 54,080 (4,659) are both higher than those for the Pooled Petersen (which, of course, remains unchanged from the unpooled analysis).
- The G2 test indicates a poor fit; notice by comparison of the observed counts (Data pane) with the predicted counts (Results pane), that the lack of fit is mostly due to a failure to predict the unmarked numbers in the first 2 and last recovery stratum.

In summary, use of the ML estimator and exploratory pooling has allowed us to get an admissible set of initial stratum estimates. There are still indications of lack of fit which could be due to several causes: marked animals may not be representative of the migration patterns of unmarked fish, or there may be closure or tag recognition problems. There are several other poolings that also give admissible ML estimates (can you find them?); they all seem to require pooling of the smaller initial and final strata, but always give lack of fit in these same strata. Without knowledge of the field collection methods, it's hard to come to conclusions about the true cause of these problems. In any case, there is clear heterogeneity that is sufficiently large to cause us to reject the Pooled Petersen. The ML estimate is no doubt less biased, but it is clearly not entirely reliable either, given the lack of fit.

2.4 The Tour Part II: Running a Simulation

In addition to providing numerous analysis methods for data, SPAS allows you to analyze hypothetical populations to explore the properties of the estimates and to aid in experiment planning and design. This functionality is provided through the **Simulation** menu. To begin, load the sample simulation file provided with the SPAS distribution by selecting the *Open…* option from the **Simulation** menu and choosing `simtest.sim` from the standard file dialog box. SPAS opens a Simulation split window, which should now contain the sample simulation data set. Your screen should look more or less like Fig. 3, except that the inner window title will be `Simulation: simtest.sim`. Maximize or re-size this window so that you can scroll through its Data pane.

Inspecting the loaded simulation data set, we can see that it closely corresponds to the format laid out in Fig. 2, with the exception that SPAS has added row and column titles to some of the parameters (notably the capture probabilities, initial stratum size, and the recapture probabilities). Also, by scrolling to the far right of the data set, we can see that SPAS has calculated the survivorship for each of the initial strata. Although values less than one are easily accommodated by SPAS, values greater than one are errors which SPAS will inform you about should you attempt to run a simulation with survivorship greater than one hundred percent.

2.4.1 A Mean Value Simulation

We will carry out a mean value analysis for this set of parameters and then do a replicated simulation and compare the results of the two simulations. Carry out the mean value analysis as follows:

1. With the `Simulation: simtest.sim` split window active, choose *Mean Values* from the **Simulation** menu. SPAS opens a new split window whose title is `Analysis: Mean Values`. Re-size this and scroll through the Data pane. Note that it is just like a real data set except that the counts are not whole numbers. You can see that the expected counts are derived from the parameters in the simulation window (try to size and arrange the two windows so that you can compare the two Data panes). For example, the capture probability in the first marking stratum was 0.3, and the initial stratum size was 5000 animals, giving 1500.00 expected marks in the first marking stratum. Of these, a fraction 0.25 go to recovery stratum 1 where they are recovered at rate 0.5, giving an expected value for m_{11} of 187.50.
2. With the Analysis: split-window active, select *All* from the **Analysis** menu.
3. Scroll through the Results pane. Note that most estimates correspond exactly to the theoretical values, as one would expect. The interesting result is the expected precision. For the ML estimate, the s.e. is 139.16 while it is 115.4 for the Pooled Petersen.

2.4.2 A Simulation with Replication

By clicking or using the **Windows** menu, switch to the Simulation window. We will now do a replicated simulation. This can help confirm if the mean value results are accurate and provides more detailed insight into the distribution of the statistics, estimates and test results. Before continuing, it is necessary to explain the concepts of replications and good replications. All the estimates available from Simulation can fail in some way (e.g. the iterative Darroch can encounter a singular matrix, a division by zero can occur in the Schaefer, etc.). When failures occur, the estimates produced are invalid or non-existent and therefore cannot be added to the overall statistics for the simulation. A replication where such a failure takes place is a bad replication, since its results are unusable. The opposite of such a replication is a good replication, i.e. a replication where all the estimators succeed in producing valid estimates (note that valid does not imply admissible). Since you request that all of the statistics produced by the simulation have a base of a given number of simulations, SPAS will attempt to produce that number of good simulations. If it runs out of allowed attempts, results will be reported based on the number of good simulations achieved.

Replicated simulations do not open a new split window, but write their results directly to the Results pane. Start the simulation now by selecting the *Replicated…* option from the **Simulation** menu. When this option is selected, you will be presented with a dialog box that contains these items: five switch boxes with yes/no buttons that allow you to select analyses to run during the simulation, a type-in box for the number of replications to perform, and a type-in box for the maximum number of replications to perform (i.e. the maximum allowed attempts, good and bad, before giving up on obtaining the requested number of good replications). Accept the default analyses (all analyses will be performed) and type in 10 replications and 20 maximum. Click in the OK box.

As the simulation runs, you will be appraised of its progress (the number of replications attempted, the number of good replications performed, etc.) in the status box at the bottom of the simulation window. Once the simulation is complete, scroll through the Results pane and note the following:

1. The initial and final seeds for the random number generator used in generating the hypothetical populations: 123456789, 123456789 and 396367697, 1235695893 respectively. SPAS will use the final seeds as the initial seeds for any subsequent simulation during the same session to ensure independence of results. If your initial and final seeds don't agree with these, your numeric results may not be quite the same as those reported below. You can manually set initial seeds by using *Seeds…* under the **Options** menu and re-do the simulation to get the identical results.
2. The statistics for the simulated populations. Two tables give the mean and standard deviation over reps of the generated data. The form of each table corresponds to the table of raw input data in the Data pane of a simple Analysis: it gives the initial (n^c_{i}) and final (n^r_{j}) sample sizes in each stratum and the m_{ij}. You should compare these values with those in the Data pane of the `Analysis: Mean Value` window to see how closely the mean statistics track to the expected values. For example, we saw from the Mean Value analysis that the expected value for n^c_1 was 1500.0 and for m_{11} was 187.50. The average of these statistics over the 10 replications are 1491.8, and 180.6: we see that the simulated values do not differ significantly from the expected results (the s.d. of the counts are reported in the next table as 33.5 and 8.69, respectively; thus the s.e. of the means are obtained by dividing by $\sqrt{10}$ giving 10.6 and 2.75 respectively, and the means are 1 and 2.5 s.e. away from their expected values, neither of which is extreme, although the latter deviation is significant at the 5% level).
3. The mean and s.d over reps of the selected estimators (and their standard errors, if available). In our example, the ML Darroch estimate had an average value (over the 10 reps) of 20,013.81 with a standard deviation (over reps) of 125.14. Since the true population size was 20,000 we see that the estimate was not biased: the s.e. of the average is around 40 ($125.14/\sqrt{10}$) and so the average is less than 1 s.e. from its true value. The simulation also lets us check for bias in the standard error formula for the estimate: the average (over reps) of the standard error of the estimate was 142.91 which, if unbiased should be close to the observed standard deviation of the estimate over reps (125.14). The mean value analysis predicted a precision of 139.16. The observed discrepancy is probably just due to the small number of reps used. The comparison should normally be based on 100 or more replications to get a precise value for the standard deviation of the estimate.

4. After the replicated results for the estimates, SPAS prints out tables of the mean and s.d. over replicates of the estimates of stratum sizes and capture rates and for the predicted counts for each estimate (the ML, moment, and Least Squares). These tables correspond to the Table of Stratum Estimates and Predicted Counts in the Results pane of a simple Analysis. When all the assumptions are satisfied the average predicted counts will not differ significantly from the predicted values obtained by doing the *Mean Value* simulation (Results pane). However, this comparison can be revealing of the behaviour of the stratified estimates when assumptions such as closure are not satisfied (see next section).

You may now either quit SPAS, or, if you have a fast computer or lots of time, try repeating the simulation with 100 reps and see how this affects the results. To compare the two runs, you can open a new simulation window (Simulation…Open) and reload the same .sim file. Carry out the 100 reps in this window and switch between the two windows to compare with the results of the original 10 reps. Your Windows list will now show two windows with the identical name (`Simulation: simtest.sim`) but the window numbers will indicate which was opened first.

2.4.3 Pooling Simulated Data

Since it is quite common to pool real data, it is desirable to be able to predict the outcome of such poolings on the precision and accuracy of the estimates. This is accomplished with the Pooled Simulation window. To create a poolable simulation data set, make sure the `Simulation: simtest.sim` window is active and select the *Replicated w/Pooling* option from the **Simulation** menu. SPAS will open a split window titled: `Pooled Simulation: simtest.sim`. Its data pane will contain the expected values (just as if you had chosen a Mean Values simulation). The data in this pane can be pooled and dropped just as in analysis of a real or Mean Value data set to simulate the effects of pooling on the simulated population. Although SPAS applies the poolings and deletions you specify to the mean values in the Data pane, it is also remembering them, and will apply the identical set of operations to each replicated data set when you come to carry out the simulation.

For example, let's simulate the effect of starting our sampling too late into a run by dropping the first row (labeled “Row 1”): bring up the Pool Rows... dialog, select Row 1 and click the *Drop* button. Note that the first stratum of 5000 unmarked animals is still modeled and these animals migrate and contribute to the unmarked recoveries. The effect of dropping the first row is to pretend that the first sample was never taken so that none of these animals will be marked. Next, perform 100 replications of this experiment by selecting the *Replicated* option from the **Simulation** menu. In the dialog box that then appears, turn off the choice of Moment estimates (the data array is no longer square) and Least Squares estimates (uninteresting) and set the number of replications (100) and maximum attempts (120). Hit OK to close the dialog and start the simulation. As before, SPAS will update you on the simulation's progress and write the results to the Results window once it is complete. What effect did the late start time have on the results of the estimates? You should find that, for the population estimates, both the Schaeffer and Petersen are unbiased for the true population (20,000). This occurs because the conditions for pooling are satisfied (the recovery rates are constant over strata) and it is known that if the only violation of closure is through births or new entries, then the Petersen estimates the population including the new additions (i.e., the population at recovery time). The ML Darroch,

however, is biased. You should get an average around 19,400 and an average s.e. of about 160. This is an unusual case where the Petersen is more robust to failure of the closure assumption than the ML. You can do the Mean Value analysis to confirm these results by following these steps:

- use **File**…*Save As* to save the contents of the Data pane of your Pooled Simulation window (say to `mvsim.dat`). Note that the window title does not change when you do a *Save As* as it does when using *Save As* from an Analysis window. This is because the data saved can be re-loaded to an Analysis window but not to a Pooled Simulation window *[N.B. the Save As does nothing at the moment...this feature is not yet implemented]*.
- use **Analysis**…*Open* to load your fixed up `mvsim.dat` and carry out the same analyses on the mean value data as you did in the replicated simulations.

As another example, let's see what effect pooling the first two re-capture strata would have on our estimates. Switch to the Simulation Pooling window using the **Windows** menu. Next, restore the deleted first row using the *Restore* option from the **Options** menu. Finally, pool the first two columns and run the simulation as before, scrolling through the Results window to determine the effects of the pooling. This gives unbiased results for all population estimates because of the constant recovery rates over strata.

This concludes our guided tour of SPAS. Although you are not familiar with all of its features, you are 90% on your way to using it to its full capacity. At this point, exit SPAS by selecting the *Exit* option from the **File** menu.

3 ANALYSIS AND TESTING METHODS

SPAS implements the methods and tests outlined in Seber (1982, Chapter 11) with some improvements and extensions and replacement of obsolete methods (Darroch's formulation of the ML method, especially for $s \neq t$) with more reliable ones (Plante's formulation). To understand and properly use these methods it is important to understand their assumptions and limitations. We review the assumptions and then describe the estimates and tests with some pointers on their use in detecting assumption failures and protecting against them.

3.1 Assumptions

The key to understanding all the methods is to be aware of the assumptions and effects of their failures on the Petersen estimate and how they generalize to stratified estimates. The main assumptions of the Petersen method are:

1. *Closure*: no animals enter or leave between the two sample times. This assumption can be relaxed. If the loss rate is the same for marked and unmarked animals, the Petersen estimate is still a consistent estimator of N. Consistency is a statistical term that means, roughly, unbiased in large sample experiments. If there are no losses but B new animals enter, then it is a consistent estimator of population size at time 2, N+B. If both births and losses occur, it

is an overestimate for both N and N+B; in fact it is consistent for $N+B/\phi$ where ϕ is the proportion surviving from the first to the second sample.

2. *No tag loss*: if the tag retention rate between the two times is θ ($0 \leq \theta \leq 1$), then the Petersen estimate is consistent for N/θ (Arnason and Mills, 1981). This assumption also requires that tagged and untagged recoveries are correctly identified as such on recapture. It is important to estimate the magnitude of these and other tag-related effects and adjust the effective number of tags released or recovered. This can remove bias in the Petersen estimate but variability added to the estimate by these effects will not be reflected in the usual standard error. The tag-related effects that commonly occur in fisheries work are:
   a) *Tagging-induced mortality*: this is estimated from holding experiments on a sample of tagged fish and used to reduce the effective number of tag releases.
   b) *Tag loss*: this is estimated by double tagging experiments. Methods and adjustments are discussed by Seber (1982, Chapter 3).
   c) *Tag non-reporting*: this is especially a problem when the recovery sample is taken by a commercial or sport fishery rather than a designed sampling effort. Estimation and adjustments for tag-reporting are discussed by Bowen and Sargent (1983).
   d) *Tag identification*: tags that are not identified as such have the same effect as if the tag was lost. Steps can be taken to estimate the non-identification rate by double sampling. For example, when recoveries are by "dead pitching", a second crew can follow to re-pitch.
3. *Equal catchability*: the probability of recovering an animal is independent of its marked/unmarked status. This is the Achilles heel of the Petersen estimate. The estimate is sensitive to failure of this assumption and failure can occur in many ways (trap-happiness, trap avoidance, behavioural or attribute differences causing unequal catchability among animals, etc.…see Seber, 1982 for further discussion) most of which cannot be tested for using the sampling data alone. It can produce overestimates, if the effect is to give a lower ratio of marks than expected under equal catchability, or underestimates if this ratio is higher than expected.

To form estimates from the stratified experiment (Darroch ML and moment estimators, and the stratum estimators of the Schaefer method) the three assumptions of the Petersen experiment are again required in slightly extended form:

1. *Closure*: animals that make up the population of the capture strata have a non-zero probability of recovery in one of the final strata and all animals in the final strata were also present in one of the initial strata. Thus if dead animals are sampled, as in the salmon run example, physical death does not imply lack of closure. In salmon runs, closure is achieved by ensuring that sampling starts at the beginning of the run and that sampling for carcasses continues until all animals have spawned and died. The effects of failure of the closure assumption in stratified experiments depends on whether the assumptions for pooling are met and the nature of the closure violation. The effects are summarised in Table 1.
2. *No tag loss*: animals retain their tags and are correctly identified as marked or unmarked and, if marked, by initial stratum. Adjustments for the other tag effects listed above may require distribution of the adjustments across strata since if a tag is lost, the capture stratum is unknown and must be assigned proportional to the distribution of initial strata in the marked recoveries that retain their tags.

3. *Equal catchability*: all animals in a given final stratum, whether marked or unmarked, have the same probability of being sampled.

Table 1. Closure effects on the consistency of the Schaefer (Sch.), Pooled Petersen (PPE) and Darroch (Dar.) estimators, as reported in Arnason et al. 1996. Three cases of the Darroch are considered: estimates of total size N when $s = t$; estimates of N and initial stratum sizes N^c when $s < t$; estimates of N and final stratum sizes N^r when $s > t$. Consistency results are shown for closure assumptions for death and birth (Pr = Present; Ab = Absent) and when initial (p^c) and final (p^r) sampling probabilities are equal (E) or unequal (U) across strata. A blank cell indicates the estimate is consistent; + or – indicate positive or negative bias, x indicates badly biased and inadmissible estimates, and o indicates cases that were not investigated.

Closure Effect		Sampling Prob.		N			Dar. $s<t$		Dar. $s>t$	
Death	Birth	p^c	p^r	Sch.	PPE	Dar. $s=t$	N	N^c	N	N^r
Ab	Ab	E	E							
		E	U							
		U	E							
		U	U	+	+					
Pr	Ab	E	E							x
		E	U							x
		U	E	+	+				–	x
		U	U	+	+				–	x
Ab	Pr	E	E					x		
		E	U	+	+			x		
		U	E				–	x		
		U	U	–	–		–	x		
Pr	Pr	E	E	+	+	+	o	o	o	o
		E	U	+	+	+	o	o	o	o
		U	E	+	+	+	o	o	o	o
		U	U	+	+	+	o	o	o	o

To these assumptions, we need to add:

4. All *marked* animals released in a given initial stratum have the same probability distribution of movement to the final strata. Darroch (1961) points out that this assumption is sufficient when $s \geq t$; when $s < t$, it is also necessary to assume that the marked and unmarked animals move with the same probability distribution.

3.1.1 Pooling and the pooling tests:

To carry out complete pooling requires further assumptions (for validity of the pooled Petersen and Schaefer estimates). It is known (Chapman and Junge 1956) that the Schaefer estimate is biased unless the capture probabilities, p^c_i , are equal across all initial strata, i , or all the recovery probabilities, p^r_j , are equal across final strata, j. The conditions for consistency of the pooled Petersen estimate (PPE) are not so stringent but cannot be stated so succinctly. The general condition involves a combination of the migration, capture and recovery rates, but a number of special cases can be identified and some of these can be tested for.

If any of the following conditions is satisfied, then the PPE is consistent (i.e. unbiased in large samples) if:

1. the recovery probabilities are constant across strata (i.e., $p^r_j = p^r$ for all j).
2. the (expected) ratio of marked to unmarked is constant across all recovery strata. This can be achieved in one of several ways. Two possibilities are:
   (a) the proportion of each initial stratum marked is constant across all capture strata and marked and unmarked animals experience the same migration patterns;
   (b) the migration pattern of marked and unmarked animals across final strata is independent of their initial strata; i.e. $\theta_{ij} = \theta_j$. If this holds, the $\boldsymbol{m}$ matrix is likely to be singular.

The test labelled "Equal Proportions" at the beginning of the Analysis results tests for Condition 2 using a chi-square test of the 2-by-t table with rows given by the $m_{\cdot j}$ and the u^r_j. If a low or non-significant test result occurs, it means that full or partial pooling is probably acceptable. If the low result is due to similarity of migration patterns, as in 2(b), the pooling is probably necessary because there are redundant strata and the Darroch estimator will probably fail to converge.

The test labelled "Complete Mixing" is a 2-by-s table with rows $m_{i\cdot}$ and u^c_i. This is a test of the hypothesis that the probability of resighting a released animal is independent of its stratum of origin. This condition will certainly hold, regardless of any assumption about migration, if condition (1) holds.

If either test passes (i.e. $p > 0.05$), it should be safe to use the PPE. In practice, few biological populations satisfy either condition and the tests usually indicate rejection, but this does not mean that partial or complete pooling is invalid. Other criteria should be examined, including seeing if pooling produces big changes in the estimates.

Other ways of determining whether pooling is legitimate have been discussed before (section 2.3.2). These involve examining the estimates of capture probabilities and the effect of pooling on the G2 Goodness of fit test. A large increase in the G2 value or a large change in the population estimate (say, larger than the s.e.) as a result of a partial pooling should certainly raise suspicions that heterogeneity has been introduced. A likelihood ratio test for pooling is possible but is not yet implemented. The current likelihood value that is printed out by SPAS cannot be used for this purpose because it is evaluated at the current data values and thus different poolings result in different data and non-comparable likelihood values. If it were

always evaluated using the unpooled data, the change in likelihood value or the AIC (Akaike Information Criterion) could be used to help select a pooled model. If the user can get reasonable looking estimates of capture rates, stratum sizes, and migration rates from the data, then a simulation study as described in the next section can tell whether a given pooling is worthwhile: the user must trade off the risk of introducing bias against the possible gains in precision. Studies by Kirby (1996) have shown that the Darroch or partial pooling of the Darroch can be much better than the PPE because of very large biases (up to 30%) in the PPE due to heterogeneity in capture and recovery rates.

3.2 Estimation methods

We give a brief review of the estimates used in SPAS emphasizing how they were implemented. We list sources of further information on their properties. Most of the estimates produce an estimate of population size and its s.e. and are then followed by a table of stratum estimates and fitted values (ML Darroch, Least Squares) or a table of stratum estimates (Darroch moment, Schaefer). Recall that when $s < t$, only the initial-stratum population sizes can be estimated. In this case, the fitted values are for the m_{ij} and the u^r_j. When $s > t$, only the final-stratum population sizes can be estimated and the fitted values are for the m_{ij} and the u^c_i. Typically, it is the unmarked animals that provide most of the lack of fit. When $s = t$, either set of unmarks can be used for the fit, but we treat this the same as the $s < t$ case; this case is a full rank model (number of statistics equals the number of parameters) so the fit is exact (and the ML and moment estimators are equivalent in large samples).

3.2.1 ML Darroch and Least Squares estimates

We have implemented the estimator as described by Plante (1990). It involves an iterative search starting at the Least Squares (LS) estimates. The formula for the LS estimator is also given by Plante (1990). If the LS estimator gives inadmissible estimates (which will be apparent from negative stratum estimates or capture probabilities greater than 1), it is almost impossible to get the ML to converge. The ML is obtained by an unconstrained search starting at the LS estimators.

When this estimate is chosen, a dialog box asks the user to set the search criteria: maximum number of iterations and convergence value (iteration stops when the change in the likelihood is below this value). The defaults (25, and 0.001, respectively) can usually be accepted. You can increase the number of iterations if convergence is slow and the reported iterations used equals the maximum set. Decrease the convergence value if you suspect the iteration has halted prematurely, getting stuck before maximizing the likelihood.

The s.e for N is obtained using standard likelihood methods (using the information matrix) and is used to form a normal 95% c.i.. The inverse information matrix at the ML estimates can be printed out (**Options**…*Display Cov*), but to interpret and use it (e.g. to test hypotheses about the parameters), you need to understand Plante's parameterisation. Ideally, one would like to constrain the estimators to admissible values (by using a constrained search or by reparameterizing in terms of parameters that are valid for all real values) but we have not yet found a practicable way to do this. The LS estimate may itself fail because it involves inversion

of the [**m m′**] matrix, which may be singular. Failure of either estimate can be due to small sample problems (which may be resolved by trying poolings). The m_{ij} array should be well behaved: not too many zero or very small values (<5); no linear dependencies among rows and columns leading to singularity (a strongly diagonal dominant or upper triangular structure to the array avoids this problem); there must be no row or column that is all zeros. If this is not the cause, and sample sizes are large, then it may be due to model failure. The most probable culprits are heterogeneity (radically different behaviour of animals from the same stratum) and non-closure and tagging assumption violations.

We did not attempt to implement the ML method as developed by Darroch and outlined in Seber (1982) as it is only valid for $s = t$ (the methods given in Seber for $s \neq t$ are not ML and are not practical as general methods). The Plante estimator is the only method that gives ML estimates for all 3 cases of s and t and is more robust to some forms of non-closure than the pooled estimates (Table 1). It will be less biased but also less precise than the PPE if there is variability in catchability over both the initial and final strata, but overall, the unbiasedness outweighs the loss of precision making it the better estimator (Arnason et al. 1996). We give a demonstration of this in the next section using simulation.

3.2.2 Darroch moment estimate

This estimate is only available if $s = t$. In this case, the estimates are equivalent to the ML so the stratum estimates it provides for both initial and final strata can be considered ML. The population estimate, though calculated by a different formula, (Chapman & Junge, eq. 9) should be virtually identical to the ML Darroch. The s.e. is calculated by an asymptotic method (Chapman & Junge, eq. 19) and is probably less reliable than that for the ML estimate. The estimate is provided because it provides a convenient way to present the initial and final stratum estimates when both are available.

3.2.3 Schaefer estimate

This estimate requires either constant capture or constant recovery rates over initial and final strata, respectively. If either condition holds, then the PPE is also unbiased and will generally be the more precise estimator. Although no s.e. is available for the Schaefer, one could easily be developed using a resampling method (e.g. bootstrap) but it is not worth doing, since the better PPE gives a s.e. Warren and Dempson (1995) studied the Schaefer estimate (they call it “a simple daily estimate”), but only in cases where the recovery rates were constant and so they found that the PPE was the superior estimator. Their conclusion that the Darroch estimator provides no improvement is, of course, a result of these constant rates and does not take into account the fact that the Darroch estimate reduces bias when rates are not constant. If *both* sets of rates are constant, it is possible to form estimates for both initial and final stratum sizes using simple ratio arguments (Warren and Dempson 1995). Note that this is a stronger assumption than that required for the overall population estimate alone. These estimates have the advantage that they will always produce non-negative estimates, but they may not always be admissible (if the estimate of stratum size is less than the sample size). We suspect that these estimates are very non-robust to any source of assumption violation and should be investigated thoroughly, using the simulation methods of the next section, before any faith is placed in them. We included them in SPAS because it was felt further investigation of their properties was in order.

3.2.4 The pooled Petersen estimate (PPE)

The properties of this estimator are discussed in detail by Seber (1982) in both unstratified (Chapter 3) and stratified (Chapter 11) experiments. The discussion in Seber is brought up to date by Arnason et al (1996). There are many ways to form the estimate, its s.e. and the c.i. There are essentially two methods for forming the estimate, $\hat{N}$, and s.e., depending on whether the recovery sample is assumed to be taken with or without replacement. The latter leads to the "hypergeometric" form (Seber 1982, eq. for N* and v* on p. 60) which we use in SPAS and which is the generally accepted method based on ML principles (Bartmann et al. 1987). When the marked recoveries, *m*, are low relative to the total capture and recovery samples, there is little numeric difference between the two methods. If *m* exceeds the initial sample size (as can happen when there is sampling with replacement) the s.e. formula will fail (negative variance) and the "binomial" form (Seber 1982, eq. for N_1 and v_1 on p. 61) must be used. The 95% c.i. can be formed using the usual normal theory as $\hat{N} \pm 1.96$*s.e. If sample sizes, especially for *m*, are small, $\hat{N}$ can have a strongly skewed distribution, in which case a c.i. based on a (inverse cube root) transform of the estimate produces a c.i. that is not symmetric about $\hat{N}$ and is known to have better coverage (Sprott 1981). The method is outlined in Arnason et al (1991). We provide both the normal- and transform-based c.i.; in large samples there will be little numeric difference between the two. Later versions of the simulator may permit investigation of the coverage of the c.i. both with and without violation of assumptions.

4 USING SIMULATION

Many of the points about using simulation have been made in the Tour (part 2, section 2.4). In this section we add a number of general points on the use of simulation and do a final example on the use of simulation to plan an experiment and investigate the trade-off in precision and accuracy between different pooling schemes.

All the **Simulation** options except *Mean Values* are stochastic. That is, they draw random samples to reflect the sampling and migration processes. The random number generators (L'Ecuyer and Côté 1991) and random variate generator package (Brown et al. 1993) used are high quality, portable, and reliable. The user must be careful not to re-use the streams of random numbers inadvertently by re-setting the seeds as this will induce non-independence between sets of results. As long as SPAS is running it uses the final seeds from one run as the initial seeds for the next, no matter what window the simulation is run from. The problem can arise however if SPAS is restarted, as it always starts from the fixed seed values. It is a good idea to record the final values before exiting SPAS and use **Options**…*Seeds* to reset the initial seeds to these final values the next time SPAS is started up.

4.1 Planning experiments.

The steps in using the simulator to help plan an experiment are the following:

1. Form guessed estimates of the probable population size and its distribution across strata (the N^c_i), and guesses as to how each initial stratum will distribute itself across final strata (θ_{ij}), and the capture (p^c_i)and recovery rates (p^r_j). The use of a spreadsheet like `darrochb.wk1` can be very helpful here because it lets you specify sample sizes rather than sampling probabilities, and to try many different "guesstimate" combinations quickly to arrive at something reasonable. Alternatively, one can use the editor to create and change an initial .sim file of parameters and, for each new set of parameters chosen, select *Mean Values* and examine the Data pane of the Analysis: Mean Values split window to see the resultant sample sizes. The split window can be closed without running any analysis, and this procedure can then be repeated until a reasonable set of values is found.
2. Run a Mean Values simulation on the guestimated parameters. The Mean Values simulation will tell you the expected value (MV) of the estimate and of the s.e. If there are no assumption violations, the MV of the estimate should be equal to the true value. The difference (MV − true value) is the bias in the estimate. Negative bias means the estimate is an under-estimate of the true value. If the bias is expressed as a percent of the true value, it is called the relative bias (RB). The MV of the s.e. is generally a reliable guide to the precision you can expect. The MV of the s.e. expressed as a percentage of the MV of the estimate is the predicted (percent) CV. However, if there are assumption violations, the s.e. may be biased and there is no way to be sure from the MV analysis alone. This will be checked later using replicated simulations. As described in step 1, try different versions of the .sim file parameters until a satisfactory overall precision is achieved. Look at the PPE for this, even though it may be biased, since this is the highest precision you can get for a given two-sample experiment.
3. You can now use the Analysis: Mean Values split window to explore the effect of different poolings. Generally, if you are assuming non-constant capture and recovery rates over strata, the Darroch estimates will be unbiased and the PPE will be biased but more precise. Try different pooling levels to see how much pooling can be done to gain precision in the Darroch estimate without introducing effective bias. It is known (Cochran 1977) that, if the RB does not exceed half the CV, then the bias is negligible in the sense that it has little effect on the coverage of the 95% c.i. so this can be used as a criterion for acceptable poolings. Once a promising sampling and pooling strategy is arrived at, you are ready to investigate it in more detail using replication.
4. Some questions cannot be answered by MV analysis but can be investigated from replicated stochastic simulations (given the guessed population structure and sampling and migration rates). These include:
    - what are the chances that the Darroch estimator(s) will fail?
    - is the s.e. formula biased by the presence of assumption failures?
    - how much variability in sample sizes can be expected?
    - are the stratum size estimates (available from the Darroch moment and Schaeffer estimates) unbiased?

    These are investigated by doing a **Simulation**…*Replicated* or *Replicated w/Pooling* on the .sim file. Note that if you are investigating a pooling arrived at using MV analysis, you will

have to specify the same pooling again in the Data pane of the Pooled Simulation split window that comes up when you select *Replicated w/Pooling*. In either case, the Results pane reports your initial and final seeds and the number of attempts made to obtain the specified number of good reps (for all estimates: if you want to know the failure rate of a specific estimator, select only that estimate in the replication dialog box). If you reset the seeds to the initial values reported and re-run a simulation, SPAS will re-generate the identical data sets and you should get the identical tables of statistics over replications unless the number of attempts or good reps changes. If you run out of attempts but have at least 1 good rep, results will be reported over the achieved number of good reps. If many of the attempts give bad results, you can expect the estimates and their s.e. to be biased because the estimates are now conditional estimates, given they did not fail.

Use a minimum of 100 replications to get a precise value for the s.d. over reps of the estimate (opposite the estimate name under the Std. Dev. column in the simulation results): this is what you use as the true s.e. and increasing the number of replications makes it a more precise estimate of the true s.e. (it is always unbiased for the true s.e. of the estimate). If your machine is very fast, it is best to do 1000 replications. You compare this with the mean over reps of the s.e. (under the Mean column opposite the row labeled `Std. Error` beneath the row labeled by the estimate name) to form the bias (mean – true). Generally, you will find that this mean is close to the s.e. you got in the corresponding MV analysis but it can be very different from the true s.e. If the s.e. shows effective negative bias it means that a c.i. formed in this situation will be very misleading: it will be too narrow and will have much less than 95% coverage. Positive bias is less serious: you just end up with a wider interval and higher coverage than is really the case, but if the precision is nevertheless satisfactory, it doesn't matter.

5. Carry out further investigations by repeating the steps above to examine sensitivity to parameter changes and other assumption failures such as:
    - What happens if capture rates are increased by a small amount?
    - Are there other poolings that give more precise but effectively unbiased estimates?
    - What happens if the population is half the size originally guessed (halve the N^c_i), or if mortality reduces the population at recovery time by half (halve the θ_{ij})?
    - What happens if sampling starts late (drop the first row in the Pooled Simulation pane) or ends early (drop the last column…be careful not to leave a θ_{ij} array with an empty row or column).

4.2 An example using `Darrochb.sim`

Start by using **Simulation**…*Open* to load the `darrochb.sim` file. This file is an example of what you might arrive at the end of Step 1 above. Scroll through the Data pane and look at the parameters chosen for this 8 x 10 stratified population. We have chosen a moderate sized population of 130,000 fish and assumed that the fish pass through the initial (weekly) strata following a unimodal entry curve that peaks at 33,000 fish in week 4. The capture rates vary inversely with initial stratum size as happens when sample sizes are limited by fixed trap capacity or personnel hours available to do the marking. If you look at any row of the θ_{ij} array, you see that fish move to the recovery area following a unimodal distribution also, taking about 3 to 4 weeks to get to the recovery area (if we assume "col1" is the same week as "row 1") but the

distribution is quite diffuse as might happen if the two sites are far apart. There is a high and variable loss rate of animals, partly due to cutting off recoveries too early: note that in the last 5 initial strata, there are non-zero values in the last column of the θ_{ij} array.

Now choose **Simulation**…*Mean Values* as in Step 2 above. The Analysis: Mean Values split window opens. Notice the roughly constant capture sample sizes (under the Marks column) and the very small m_{ij} array values resulting from the low marking rate and high mortalities. Do an **Analysis**…*All* and browse through the Results pane. We note that we can expect significant results from both the tests, though because of the low capture rates and recovery sample sizes, not an enormously significant result for the Equal Proportions test. The Darroch estimator (and initial stratum size estimates) is unbiased, as expected and has a CV of 19.2% (100*25,029/130,000). The PPE has a 13% negative bias (100 *[113,444-130,000]/130,000) and a CV of 6.2% (100*7,212/113,444), so we should investigate partial poolings. Before doing this, you might switch back to the `Simulation: Darrochb.sim` window and try **Simulation**…*Replicated*. Choose a low number of good reps (2) and total attempts (5) because most of the attempts will be bad and the simulation is very slow. This is because the low expected m_{ij} values often lead to unanalysable results or slow convergence. This is another reason to investigate pooling: clearly there is a high probability that pooling will be necessary just to get analysable statistics. We will skip Step 3 above and go directly to a stochastic examination of the results of pooling rows and columns in pairs.

From the `Simulation: Darrochb.sim` window choose **Simulation**…*Replicated w/Pooling*. A Pooled Simulation window will be opened and its Data pane will contain exactly the same values as you saw when you did the **Simulation**….*Mean Values*. Now pool rows and columns in pairs to reduce this to a 4 x 5 array. Note that the non-zero expected m_{ij} values are now mostly greater than 10 so there is little chance of getting an unexpected zero row or column by chance in one of the replications. After you have pooled the data, choose *Replicated* from the **Simulation** menu, and from the dialog box, choose 100 good replications and 125 total replications. When the results have been written to the results pane, browse through them and note that the MV analysis correctly predicted the average of its estimate and s.e. (if you switch back to the MV results pane you will see these are 113,444 and 7,212 respectively); the PPE estimate is biased (for the true value of 130,000) as noted from the MV analysis but the s.e. is not. It doesn't help that the s.e. is unbiased because the bias in the estimate itself means that the c.i. will have almost zero coverage. Pooling has improved the precision in the ML Darroch estimate (to a CV of less than 9%) without introducing effective bias (RB is under 2.5% which is less than half the CV). The s.e. is also unbiased (the average s.e. over reps is close to the true s.d. over reps of the estimate), so despite the presence of non-closure, and some degree of heterogeneity due to pooling, the coverage of the ML Darroch estimator in this situation would be close to the nominal 95%. Note that for both the PPE and the ML Darroch, heterogeneity leads to underestimates of the true population value. As a caution, you might note that the individual (initial stratum) estimates are not very accurate for some of the strata.

The reader is encouraged to try other experiments, such as those suggested in step 5 above.

5 DISCUSSION

SPAS has proved useful in demonstrating that the Darroch estimator can be superior to the PPE; it allows for unequal capture and recovery rates among strata and, although less precise than the PPE, can be tried with various partial poolings that recover some of this precision. The development of an easily-used interactive tool is important to support this sort of exploratory analysis. The simulation and analysis tools in SPAS have also allowed us to see that the Darroch model can still be used when applied in situations for which it was not, strictly designed: e.g. when there is mortality or new recruits (but not both), and when there is some heterogeneity within strata (as occurs when strata are pooled). We saw that both the estimate and its s.e. were usually (but not always) robust to these violations of assumptions. The use of mean value and stochastic simulations is an indispensable tool for sorting out the degree of assumption violation that can be tolerated.

The analyses in SPAS can also be applied to data that present problems due to non-standard tagging results. We have already mentioned that it is common to adjust for tagging mortality, tag loss, and tag non-reporting where there is ancillary information about the magnitude of these effects. It is also possible to adjust for multiple sightings: if an animal is seen more than once at capture time, only its first sighting is used (unless the animal is known to have died at a later sighting). If animals are returned to the final strata after a recapture (i.e. the final strata are recapture rather than dead recovery samples), there is a possibility that it may be recaptured more than once. Again only the first recapture is used. If unmarked animals are returned without marking them individually then the u^r_j may also be inflated by multiple captures of animals and must be reduced by the average recapture frequency of the marked animals. If marked animals are not individually identifiable, it will not be possible to do this adjustment.

A number of improvements could be made to SPAS. The likelihood method opens up the possibility of better tests for pooling and the fitting of constrained models (e.g. to allow some of the strata to have common capture rates; to ensure admissible estimates, etc.). Analyses could be extended to adjust for tag loss, non-reporting, etc., although this complicates the data entry procedures. The simulation methods could also be extended to allow exploration of various marking effects (by letting marked and unmarked animals be subject to different rates) for the degree of bias they produce. We are continuing work on SPAS and some of these improvements will be incorporated in later releases of SPAS (check the web site listed in section 2.1).

However, there is a limit to what the Darroch and other 2-sample estimators can yield, even from the most careful planning and sophisticated analysis. These estimators are very limited. It is difficult to test for assumption violations and if violations are known or suspected, there are no alternatives once the data are in. Mortality often is present and means that estimates are not available of the population at recovery time, which is often what is really wanted (how many animals reach the spawning ground rather than how many pass the sampling station). It is impossible to separate migration from survival rates or to obtain the required population estimates unless at least three samples are taken (Arnason 1973).

There are a variety of analytical approaches that can be used to generate salmon escapement and other population estimates. Irvine et al (1992) describe software developed to convert periodic approximations of fish abundance to escapement estimates. Schwarz et al (1993) describe a method of calculating escapement when multiple recaptures of individually tagged fish occurs. Labelle (1994) documents a likelihood method designed for situations when a fish counting fence is used but a complete count is not obtained. These and other methods are reviewed in Irvine and Nelson (1995). It is up to the investigator to decide which analytical approach is best suited for his/her situation. The two-sample mark-recovery approach described in the current report is widely used in ecology, including fisheries: Simpson (1984) and the conclusion of Dempson and Stansbury (1991) provide numerous references to (and additional datasets for) the use of these methods. Two-sample experiments are not always the best choice of experiment, but when they are used, for better or worse, SPAS can make the most of them. It gives biologists better experiment-planning tools and a comprehensive and easily-used suite of the most powerful analysis methods to estimate population sizes with better accuracy and precision.

ACKNOWLEDGMENTS

This work was supported by subvention grants to A.N.A. and C.J.S. from the Canada Department of Fisheries and Oceans and by Operating Grants from the Natural Sciences and Engineering Research Council of Canada. We acknowledge the contribution of Lai Shar in programming the final revisions to the SPAS software and developing the distribution package. Thanks are owed to Brian Dempson for reviewing the final manuscript, including a careful test of the software; and not least for re-invigorating the study of 2-sample stratified estimates by enlisting the aid of statisticians to help resolve the questions that arose in analysing the Conne River data.

REFERENCES

Arnason, A. N. 1973. The estimation of population size, migration rates, and survival in a stratified population. Res. Popul. Ecol. 15:1-8.

Arnason, A. N., C. W. Kirby, and C. J. Schwarz. 1996. Stratification can improve the accuracy of mark-recapture estimates of salmon run size. *Submitted to:* N. Am. J. Fish. Manage.

Arnason, A. N., and K. H. Mills. 1981. The bias and loss of precision due to tag loss in Jolly-Seber estimates for mark-recapture experiments. Can. J. Fish. Aquat. Sci. 38:1077-1095.

Arnason, A. N., C. J. Schwarz, and J. M. Gerrard. 1991. Estimating closed population size and number of marked animals from sighting data. J. Wildl. Manage. 55:716-730.

Bartmann, R. M., G. C. White, L.H. Carpenter, and R. A. Garrott. 1987. Aerial mark-recapture estimates of confined mule deer in pinyon-juniper woodland. J. Wildl. Manage. 51:41-46.

Bowen, W. D., and D. E. Sargent. 1983. Mark-recapture estimates of harp seal pup (*Phoca groenlandica*) production in the northwest Atlantic. Can. J. Fish. Aquat. Sci. 40:728-742.

Brown, B. W., J. Lovato, and K. Russell. 1993 The RANLIB library. Distributed by anonymous ftp from Internet site: odin.mda.uth.tmc.edu

Chapman, D. G., and C. O. Junge. 1956. The estimation of the size of a stratified population. Ann. Math. Statist. 27:375-389.

Cochran, G. 1977. Sampling techniques. 3rd ed. John Wiley and Sons, New York.

Darroch, J. N. 1961. The two-sample capture-recapture census when tagging and sampling are stratified. Biometrika 48:241-260.

Dempson, J. B., and D. E. Stansbury. 1991. Using partial counting fences and a two-sample stratified design for mark-recapture estimation of an Atlantic smolt population. N. Am. J. Fish. Manage. 11:27-37.

L'Ecuyer, P., and S. Côté. 1991. Implementing a random number package with splitting facilities. A.C.M. Trans. Math. Software 17:98-111.

Irvine, J. R., J. F. T. Morris, and L. M. Cobb. 1993. Area-under-the-curve salmon escapement estimation manual. Can. Tech. Rep. Fish. Aquat. Sci. 1932: 84 p.

Irvine, J. R., and T. C. Nelson (editors). 1995. Proceedings of the 1994 salmon escapement workshop plus an annotated bibliography on escapement estimation. Can. Manuscr. Rep. Fish. Aquat. Sci. 2305: 97 p.

Kirby, C. W. 1996. An investigation of two-sample stratified population estimates. M. Sc. Thesis, University of Manitoba, Winnipeg (in prep.).

Labelle, M. 1994. A likelihood method for estimating Pacific salmon escapement based on fence counts and mark-recapture data. Can. J. Fish. Aquat. Sci. 51: 552-566

Plante, N. 1990. Estimation de la taille d'une population animale à l'aide d'un modèle de capture-recapture avec stratification. M.Sc. thesis, Université Laval, Quebec.

Schaefer, M. B. 1951. Estimation of the size of animal populations by marking experiments. U.S. Fish Wildl. Serv. Fish. Bull. 69:191-203.

Schwarz, C. J., R. E. Bailey, J. R. Irvine, and F.C. Dalziel. 1993. Estimating salmon spawning escapements using capture-recapture methods. Can. J. Fish. Aquat. Sci. 50:1181-1197.

Seber, G. A. F. 1982. The estimation of animal abundance. 2nd ed. Griffin, London.

Simpson, K. 1984. The accuracy of mark-recapture estimates of escapement. pp. 209-225, *In*: P. E. K. Symons and M. Waldchuk [eds.] Proceedings of the workshop on stream indexing for salmon escapement estimation. Can. Tech. Rep. Fish. Aquat. Sci. 1326.

Sprott, D. A. 1981. Maximum likelihood applied to a capture-recapture model. Biometrics 37:371-375.

Warren, W. G., and J. B. Dempson. 1995. Does temporal stratification improve the accuracy of mark-recapture estimates of smolt production: a case study based on the Conne River, Newfoundland. N. Am. J. Fish. Manage. 15:126-136.