Special Publication No. 92-4

Mark-Recapture Experiments to Estimate the Abundance of Fish

A Short Course Given by the Division of Sport Fish Alaska Department of Fish and Game in 1991

by

David R. Bernard
and
Patricia A. Hansen

November 1992

Alaska Department of Fish and Game

Division of Sport Fish

SPECIAL PUBLICATION NO. 92-4

MARK-RECAPTURE EXPERIMENTS TO ESTIMATE THE ABUNDANCE OF FISH[1]

A Short Course given by Division of Sport Fish
Alaska Department of Fish and Game in 1991

by

David R. Bernard
and
Patricia A. Hansen

Alaska Department of Fish and Game
Division of Sport Fish
Anchorage, Alaska

November 1992

[1] Development and publication of this manuscript were partially financed by the Federal Aid in Sport Fish Restoration Act (16 U.S.C. 777-777K) under Project F-10-7, Job No. RT-8.

The Special Publications series was established in 1991 for the publication of techniques and procedures manuals, informational pamphlets, special subject reports to decision-making bodies, symposia and workshop proceedings, application software documentation, in-house lectures, and other documents that do not fit in another publication series of the Division of Sport Fish. Special Publications are intended for fishery and other technical professionals and lay persons. Distribution is to state and local publication distribution centers, libraries and individuals and, on request, to other libraries, agencies, and individuals. This publication has undergone editorial and peer review.

The Alaska Department of Fish and Game receives federal funding. All of its public programs and activities are operated free from discrimination on the basis of race, religion, sex, color, national origin, age, or handicap. Any person who believes he or she has been discriminated against by this agency should write to:

OEO
U.S. Department of the Interior
Washington, D.C. 20240

TABLE OF CONTENTS[1]

[1] Headings in parentheses are worded exactly as were headings in the original class notes. Other headings were included in this Special Publication to better reflect the content of each section.

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

INTRODUCTION

In January, February, and March, 1991, the Biometrics Section of the Division of Sport Fish, Alaska Department of Fish and Game presented a 3-day short course on mark-recapture experiments. Sessions were given in Anchorage, Juneau, and Fairbanks. Fifty-four people from the Divisions of Sport Fish; Commercial Fisheries; Fisheries Rehabilitation, Enhancement, and Development, and Wildlife Conservation and from the University of Alaska, Fairbanks attended. The short course consisted of six lectures and five laboratories. All laboratories were based on data and situations from Alaska. About 1.2 megabytes of data and software were used in the labs, and there were about 100 pages of written materials given to each student in support of lectures and labs.

This manuscript is a Special Publication of the written materials disseminated during this short course plus some additional material. In this publication, pages with Arabic numerals 4 through 45 were handed to students at the start of the short course to supplement (and be supplemented by) orally presented lectures. Since "oral presentation" can not be used in this manuscript, text has been added to the original written material whenever coordination between written materials and oral lectures was essential to communicate some important concept. For instance, equations describing variations on Petersen's model were originally presented on the blackboard and not on pages. Because of the importance of this model to stock assessment of fish populations in Alaska, descriptions of these equations have been added to the original materials in this Special Publication. All such additions are pages with Arabic numerals 1 through 3. Whenever material presented verbally was of lesser importance, no additions to the "hand-outs" were put in this Special Publication. For example, methods of maximum likelihood used to estimate abundance with multiple-event mark-recapture experiments on closed populations were presented during lecture, but were not described in text. Since experience in Alaska has shown that these methods are not as effective as others, no description of these experiments were added to the original materials for this Special Publication.

The remaining pages of disseminated text concern the five laboratories. Each lab consisted of a series of problems based on mark-recapture experiments in Alaska. Text on all lab materials without the answers were handed to students at the beginning of the short course; at the end of each lab, a copy of the problems with the answers were given to students. Answers (in italics) are included with lab materials in this Special Publication (pages 46-74).

There are some discrepancies in the labs. Tallies of marked and recaptured fish in some tables should correspond to tallies in others, but do not. For instance, 17 northern pike were recaptured in the northwestern end of George Lake in 1987 according to statistics in one table and only 16 according to another table. This discrepancy exists because one of these recaptured fish had lost its individually identifiable tag while retaining its batch mark. Because we could identify where it was recaptured, but not where it had been released, this fish was counted in one table and not in the other. Similar discrepancies arise because other fish lost tags or were not measured. This is the nature of mark-recapture experiments, and the labs reflect this realism. A second family of discrepancies concerns χ^2 statistics on 2x2 contingency tables. Statistics from your software may differ from those in the Special Publication depending on whether your software does or does not contain a correction for continuity (our calculated statistics and software do not include the correction).

The software and data used in the short course can be obtained upon request from the authors. Software to test hypotheses common to mark-recapture experiments is generally included in any comprehensive statistical package for personal computers. However, we designed our own software for hypothesis testing (routines provided through IMSL, Inc.) because we did not want to load several megabytes of a statistical package on a dozen personal computers only to access this small range of procedures. Unless you are faced with a similar problem, you will probably not need our software. With one exception, data for the labs were obtained from the annual reports and archives of the Division of Sport Fish. In the one exception, the lab (on planning) was based on an actual project, even though fictitious, but realistic data were used.

Obviously this Special Publication is not a text on mark-recapture experiments and should not be cited as such in other reports, operational plans, or manuscripts. However, it is a source of information on various methods used in planning and analysis. Therefore, upon request we will send the text file of this Special Publication in Microsoft Word™ to you for dissection and subsequent inclusion of the parts in your reports and operational plans.

This report would not be complete without acknowledgment and thanks to those persons whose assistance was crucial to the success of this endeavor. Many thanks to Sandy Sonnichsen and Gail Heineman for their help with data and software; to Keith Webster and Bob Clark for their assistance with machinery; to Doug McBride, Al Didier, and Peggy Merritt for their coordination in Anchorage, Juneau, and Fairbanks; to Gwyn Karcz, Allen Bingham, and Al Howe for their editing of this manuscript; and to the biologists who attended the short course for their ideas and patience. And finally, we are specially grateful to the many biologists over the years with whom we have had the pleasure of exploring mark-recapture experiments.

PETERSEN'S MODEL: CHAPMAN'S AND BAILEY'S MODIFICATIONS

Petersen's model for two-event mark-recapture experiments on closed populations is:

$$\hat{N} = \frac{M\,C}{R}$$

where N = abundance, M = number of marked fish released alive during the first sampling event, C = number of fish captured during the second sampling event, and R = number of fish marked and released alive during the first event that were recaptured during the second. Estimates from this model contain some *statistical bias*, especially when sample sizes (most notably number recaptured) are small.

Chapman's and Bailey's modifications of Petersen's model (Seber 1982) are used to reduce statistical bias of estimated abundance. Whenever sampling during the second event is done without replacement or early sampling during this event affects the fraction of marked fish in the population, Chapman's modification:

$$\hat{N} = \frac{(M + 1)(C + 1)}{(R + 1)} - 1$$

$$V[\hat{N}] = \frac{(M+1)(C+1)(M-R)(C-R)}{(R + 1)^2(R + 2)} \approx \frac{\hat{N}(M - R)(C - R)}{(R + 1)(R + 2)}$$

based on the hypergeometric probability distribution is the appropriate model to use. Estimated abundance from Chapman's model has no statistical bias if $M + C > N$, and statistical bias in the estimate is most likely negligible if $R > 7$. If sampling during the second event is done with replacement or early sampling during this event does not affect the fraction of marked fish in the population, Bailey's modification:

$$\hat{N} = \frac{(M)(C + 1)}{(R + 1)}$$

$$V[\hat{N}] = \frac{M^2(C+1)(C-R)}{(R + 1)^2(R + 2)} = \frac{\hat{N}^2(C - R)}{(C+1)(R+2)}$$

based on the binomial probability distribution is the more appropriate model. As before, statistical bias in estimated abundance is most likely negligible if $R > 7$. Estimates from both Chapman's and Bailey's modifications are similar for large sample sizes, especially with large numbers of recaptured fish. Under these conditions, either model could be used to estimate abundance.

If mark-recapture experiments do not meet certain conditions, estimated abundance from either modification of Petersen's model will contain *structural bias*. Descriptions of these conditions, how to plan experiments to meet them, consequences of not meeting them, and how to remove and measure structural bias in estimates comprise the remainder of this Special Publication.

COURSE OUTLINE

I. Hypothesis testing
   A. Probability distributions, α, β, and power

   B. Kolmogorov-Smirnov two-sample test

   C. Contingency tables and χ^2

II. Two-event experiments on closed populations - Petersen's method
   A. Chapman's Model/Bailey's Model
      1. Equations
      2. Statistical Bias
      3. Requirements
         a. "AND" conditions
         b. "OR" conditions

      4. Planning
         a. Sample sizes
         b. Fulfilling requirements

      5. Analysis
         a. Hypothesis tests/responses
            1. Handling-induced effects on fish
            2. Loss of marks
            3. Recruitment
            4. Size-selective sampling
            5. Unequal probabilities of capture and incomplete mixing
         b. Estimated age/size compositions with size-selective sampling
         c. Estimated abundance with growth recruitment

   B. Darroch's method

III. Multiple-event experiments on closed populations - maximum likelihood methods

IV. Multiple-event experiments on open populations - Jolly-Seber methods

DATA AND SOFTWARE

There are several programs and data files associated with this short course, all of which are in the directory C:\MREXP. During the laboratory, the name of each relevant data file will be given with the problem. Sometimes the software to conduct analysis of these data will be named also, however, most of the time selection of software will be left up to your discretion. All software is interactive in nature; just invoke the program and you will be prompted for input. Some input will be given directly in the problems, but other inputs will require some reworking of the data files. With one exception, software designed for this short course can be found in any good statistical package available for use on personal computers. Text files of this document and the lab problems are in the directory also. Feel free to take any or all of the data, software, and/or text files with you when the short course is over.

Currently, text files are available only in Microsoft Word™. If you wish a copy of the lecture and lab materials in WordPerfect™, please indicate so on the sign-up sheets, and copies will be sent to you when available.

AGE1.WK1	Lotus 123™ spreadsheet for the first problem in the laboratory on ADJUSTMENTS IN ESTIMATES OF COMPOSITION (data).
DARROCH1.WK1	Lotus 123™ spreadsheet for the first problem in the laboratory on PARTIALLY STRATIFIED EXPERIMENTS: DARROCH (data).
DARROCH2.WK1	Lotus 123™ spreadsheet for the first problem in the laboratory on PARTIALLY STRATIFIED EXPERIMENTS: DARROCH (data).
FLICK.WK1	Lotus 123™ spreadsheet to cull recruits from mark-recapture data using the method of Robson and Flick (1965) (requires one file for input) (software).
HYPOP2.WK1	Lotus 123™ spreadsheet for the second problem in the laboratory on HYPOTHESIS TESTING (data).
HYPOP3.WK1	Lotus 123™ spreadsheet for the third problem in the laboratory on HYPOTHESIS TESTING (data).
HYPOP5.WK1	Lotus 123™ spreadsheet for the fifth problem in the laboratory on HYPOTHESIS TESTING (data).
KS2.EXE	Interactive program to calculate statistics for the Kolmogorov-Smirnov Two-Sample Test (requires one file for input) (software).
KS2M.EXE	Interactive program to calculate statistics for the Kolmogorov-Smirnov Two-Sample Test (requires one file for input). Requires a math-coprocessor (software).
MREXP.DOC	This document (Microsoft Word™) (text).
MREXPROB.DOC	Laboratories (Microsoft Word™) (text).
NUMBER1.WK1	Lotus 123™ spreadsheet for the first problem in the laboratory on PETERSEN'S MODEL: ANALYSIS (data).
NUMBER2.WK1	Lotus 123™ spreadsheet for the second problem in the laboratory on PETERSEN'S MODEL (Analysis) (data).
XSQ.EXE	Interactive program to calculate statistics for the contingency table analysis for an RxC table (maximum 5x5) (software).
XSQM.EXE	Interactive program to calculate statistics for the contingency table analysis for an RxC table (maximum 5x5). Requires a math-coprocessor (software).

HOW TO PLOT CUMULATIVE DISTRIBUTION FUNCTIONS

Plotting cumulative distribution functions is part of determining whether two samples have similar length distributions. Plots of cumulative distribution functions augment hypothesis tests based on empirical distribution functions, such as the Kolmogorov-Smirnov Two-Sample Test. Most standard statistical packages for personal computers have this test and will plot the cumulative distribution functions for you. However, programs KS2.EXE and KS2M.EXE do not have such a feature, so these plots have been provided for you whenever possible in these pages.

If you at some time decide to use KS2.EXE or KS2M.EXE instead of the standard packages, you can use LOTUS 123™ to plot the cumulative distribution functions. Below is one of several methods that can be used to make these plots.

1. Import two lists of numbers into separate columns in the spreadsheet. The first list is comprised of measured lengths from the first population; the second list is comprised of measured lengths from the second population.

2. Find all the unique lengths in both lists (distributions):

   a. Copy the first list to a new column. At the bottom of the this new column, copy the second list to create one long, combined list with <u>all</u> the measurements in it. Keep the two original lists!

   b. Add a field name to the top of the combined list.

   c. Copy the field name to two other empty columns in the spreadsheet - one column for the criterion range and one for the output range of a data query (make sure there's nothing below the output range).

   d. Now do the Data Query. First specify your input, criterion, and output ranges. The input range is the combined list of measurements, including the field name. The criterion range is one of the copies of the field name and one blank cell beneath it. The output range is just the other copy of the field name.

   e. Once you've specified all the ranges, type U for Unique. You get as output a sorted list of all the unique lengths from both distributions, and can now delete or erase the combined list.

3. Now do Data Distribution using the column of unique lengths as your bin range and the measurements from the first, original list as the values range. Move the resulting distribution to a column away from the bin range. Now do the same thing for the second original list.

4. Calculate the relative frequency and cumulative relative frequency (cumulative distribution function) of measurements in the first original list for each length in the bin range. Do the same for the second original list.

5. Now do the graph:

   Type is XY
   X is the column of unique lengths (bin range)
   A is the column of cumulative relative frequencies for the first list (population).
   B is the column of cumulative relative frequencies for the second list (population).
   Set the Y axis manually to 0 minimum and 1 maximum.
   Make the format of A and B as line only.

To get different line formats for A and B and to label the two lines, you would need to save the graph and import it into Freelance™ or a similar graphics package.

CUMULATIVE DISTRIBUTION OF CHI-SQUARE: χ^2 STATISTICS FOR HYPOTHESIS TESTING

	Probability of a Greater Value								
df	0.99	*0.95*	0.90	*0.75*	0.50	*0.25*	0.10	*0.05*	0.01
1	-	-	0.02	*0.10*	0.45	*1.32*	2.71	*3.84*	6.63
2	0.02	*0.10*	0.21	*0.58*	1.39	*2.77*	4.61	*5.99*	9.21
3	0.11	*0.35*	0.58	*1.21*	2.37	*4.11*	6.25	*7.81*	11.34
4	0.30	*0.71*	1.06	*1.92*	3.36	*5.39*	7.78	*9.49*	13.28
5	0.55	*1.15*	1.61	*2.67*	4.35	*6.63*	9.24	*11.07*	15.09
6	0.87	*1.64*	2.20	*3.45*	5.35	*7.84*	10.64	*12.59*	16.81
7	1.24	*2.17*	2.83	*4.25*	6.35	*9.04*	12.02	*14.07*	18.48
8	1.65	*2.73*	3.49	*5.07*	7.34	*10.22*	13.36	*15.51*	20.09
9	2.09	*3.33*	4.17	*5.90*	8.34	*11.39*	14.68	*16.92*	21.67
10	2.56	*3.94*	4.87	*6.74*	9.34	*12.55*	15.99	*18.31*	23.21

HYPOTHESIS TESTING AND ITS ROLE IN MARK-RECAPTURE EXPERIMENTS

The success of any mark-recapture experiment depends on meeting several necessary conditions. Most of the time, the experiment can be corrected when violation of one or more of these conditions has been found. Hypothesis tests are the means of this detection.

Traditional hypothesis tests begin with a null hypothesis and an alternative. A difference between two or more statistics is compared to sampling-induced variation in the statistics. If the comparison indicates that observing this difference is an improbable event, the null hypothesis is rejected in favor of the alternative. Before any data are collected, however, decisions are made as to the desired power of the test to detect a meaningful difference between statistics. These decisions influence the required sample size.

Hypothesis tests in mark-recapture experiments still involve differences in statistics, but power and the "meaningful difference" are defined *after* data have been collected, not before. The null hypothesis is an expression of one or more of the necessary conditions in the experiment. The power of the test will be determined not by consideration of the test, but by the desire to obtain a precise estimate of abundance and the sample size as determined on that basis. The "meaningful" difference is not the difference between sets of statistics in the test, but between biased and unbiased estimates of abundance.

Investigation of bias in a mark-recapture experiment begins by estimating bias with the unaltered model. Then one by one, null hypotheses for each of the necessary conditions in the experiment are tested, first graphically and then mathematically. Rejection of a null hypothesis means that the estimate of abundance from the unaltered model is biased. The model is then altered to produce an unbiased estimate of abundance, usually with a larger variance. If the difference between these two estimates of abundance is "meaningful" (the bias is large), the estimate from the altered model is chosen as the true estimate. If the difference is not "meaningful" (the bias is negligible), the estimate with the smaller variance is chosen for publication. If the null hypothesis is not rejected, abundance calculated from the unaltered model is the *de facto* estimate. This procedure continues until each necessary condition has been investigated.

NECESSARY CONDITIONS FOR ACCURATE USE OF PETERSEN'S MODEL

The requirements for unbiased estimates from two-event mark-recapture experiments on closed populations follow:

- The "AND" assumptions --- all must be fulfilled:
  - Marking does not effect the catchability of a fish
    - There is no handling-induced "trap happiness".
    - There is no handling-induced "trap shyness".
    - There is no handling-induced mortality.
  - Fish do not lose marks between sampling events,
  - Recruitment and death of fish can not occur between sampling events
- The "OR" assumptions --- at least one must be fulfilled:
  - Every fish has an equal probability of being marked and released alive during the first sampling event.
  - Every fish has an equal probability of being captured during the second sampling event.
  - Marked fish mix completely with unmarked fish between sampling events.

SAMPLE SIZE CALCULATIONS USING PETERSEN'S MODEL FOR MARK-RECAPTURE EXPERIMENTS

Below is an extension of Table 2 from Robson and Regier (1964) where $1-\alpha$ is the probability that the estimate of abundance will be within $\pm$ (A x 100) % of the true abundance and D is a function of α and A:

$1 - \alpha$	A	D
0.75	0.50	4.75
0.90	0.50	14.8
0.90	0.25	45.5
0.90	0.10	272
0.95	0.50	24.4
0.95	0.25	69.9
0.95	0.10	392

$1 - \alpha$	A	D
0.95	0.05	1,544
0.95	0.04	2,408
0.95	0.03	4,276
0.95	0.02	9,611
0.95	0.01	38,423
0.99	0.10	695
0.99	0.01	66,200

Prior to the start of the experiment:

$$M = C = \frac{N\,X}{(1 + X)} \qquad X = \left\{\frac{D}{N - 1}\right\}^{1/2}$$

where N = preseason "guesstimate" of abundance, M = the proposed number to release with marks during the first sampling event, C = the number captured during the second sampling event, and X = a collected constant. When the first sampling event of the experiment has been completed and M is known:

$$C = \frac{N\,X}{(1 + X)} \qquad X = \left\{\frac{D(N - M)}{(N - 1)M}\right\}$$

If for some reason C is fixed, the above equation can be used to calculate M simply by switching these two variables.

PLANNING MARK-RECAPTURE EXPERIMENTS BASED ON PETERSEN'S MODEL

Below is a check list for the planning of two-event mark-recapture experiments based on Petersen's model and its variations. Some of the items are just common sense, while others are not. Some items are relevant to every situation, others are not. Some items are mutually exclusive, and some of the items require an understanding of the behavior of fish that may or may not be available. Mark-recapture experiments can be so designed that if this critical information is not now available, it soon will be.

- To avoid and/or detect handling-induced effects on marked fish:
  - Use sampling gear and handling methods that minimally stress fish.
  - Sample when fish are less prone to injury.
  - Note and individually mark stressed fish that were released with marks, and do not mark severely stressed fish.
  - Use active sampling gear to avoid "trap-induced" behavior.
  - If passive sampling gear is used, conduct a separate, two-year experiment to detect "trap-induced" behavior.
  - Lengthen the hiatus until "trap-induced" behavior is "forgotten".
- To avoid and/or detect loss of marks between sampling events:
  - Double mark each fish during the first sampling event.
  - Check each fish captured during the second event for both marks.
- To avoid and/or detect growth recruitment between sampling events:
  - Sample during the first event so that all fish regardless of size have an equal probability of being captured.
  - Measure every fish in the sample from the second sampling event.
  - Determine the age of every fish in the mark-recapture experiment.
  - Keep the hiatus between the two sampling events short.
  - Keep the length of sampling events short.
- To avoid and/or detect unequal probabilities of capture for fish of different sizes:
  - Spread sampling across all areas of the stream or lake.
  - Sample at times when all sizes of fish are equally susceptible to capture.
  - Use sampling gear that is not size-selective.
  - Measure the length of every captured fish in both sampling events.

- To avoid and/or detect unequal probabilities of capture for fish in different parts of the lake or stream:
  - Mark fish such that each can be distinguished by where (when) it was released and by where (when) it was recaptured.
  - Spread sampling effort during one or both of the sampling events evenly over area (time).
- To avoid and/or detect partial or no mixing of marked and unmarked fish between sampling events:
  - Mark fish such that each can be distinguished by where (when) it was released and by where (when) it was recaptured.
  - Spread sampling effort during one or both of the sampling events evenly over area (time).
  - Lengthen the hiatus between the two sampling events.
  - Sample when fish are more prone to movement within the study area.

OPTIONS WHEN NECESSARY CONDITIONS FOR PETERSEN'S MODEL HAVE NOT BEEN MET

Below is a general list of ways to detect violation of conditions for a successful two-event mark-recapture experiment on a closed population. Included are ways to change analyses to correct for these violations. Unless specifically noted, estimated abundance is germane to the time of the FIRST EVENT.

- Handling-induced effects on marked fish:
  - If recapture rates of marked fish released while notably stressed are significantly lower than other marked fish and:
    - if "stressed" and "unstressed" fish have similar size distributions when released, estimate abundance without "stressed" fish, then add the number of "stressed" fish released with marks to the estimate.
    - if "stressed" and "unstressed" fish have dissimilar size distributions when released, stratify the population into two or three groups based on size, and repeat the procedure above for each group.
  - If "trap-induced" behavior has been found in a separate, two-year experiment, use the fraction of marked fish in the population from that experiment to estimate abundance.
- Loss of marks between sampling events:
  - Use the mark that was not lost to identify recaptured fish.
  - Estimate the true fraction of recaptured fish in the sample from separate rates of loss for both marks (Seber 1982).
- Recruitment between sampling events:
  - If both GROWTH RECRUITMENT AND MORTALITY (or emigration) occurred between sampling events, cull growth recruitment with methods of Robson and Flick (1965) or by dividing the population by age (p. 19-25).
  - If RECRUITMENT (or immigration) of fish of all sizes and no or little mortality (or emigration) occurred between sampling events, no correction is needed, but note that the abundance estimated will be for the SECOND SAMPLING EVENT, not the first as is the usual case.
- Recruitment between sampling events (continued):
  - If RECRUITMENT (or immigration) of fish of all sizes AND MORTALITY (or emigration) occurred between sampling events, no correction is possible, and note that the estimated abundance from this experiment will be TOO LARGE.

- Unequal probabilities of capture for fish of different sizes:
  - If length distributions of MARKED fish and RECAPTURED fish are similar, probabilities of capture were equal for fish of all sizes during the SECOND EVENT, and no correction is needed.

  - If length distributions of MARKED fish and RECAPTURED fish are dissimilar, probabilities of capture were not equal for fish of all sizes during the SECOND SAMPLING EVENT. Stratify the population into two or three groups based on size to produce two or three mark-recapture experiments to estimate abundance for each group.

  - If a large proportion of UNMARKED fish during the SECOND EVENT are larger or smaller than RECAPTURED fish, probability of capture during the FIRST EVENT for these larger or smaller fish is near zero. Ignore these larger or smaller fish in the SECOND EVENT and estimate the abundance just for the fish of moderate size.

- Unequal probabilities of capture for fish in different parts of the lake or stream or failure of marked fish to mix completely with unmarked fish:
  - If the fraction of the population comprised of marked fish is similar in all areas of the lake or stream in the SECOND EVENT, either every fish had an equal chance of being caught during the FIRST EVENT or marked fish mixed completely with unmarked fish BETWEEN EVENTS. In either case, "OR" conditions have met and no correction is needed (if fractions are dissimilar, see next page, same hierarchy).

    - If inspection of data shows marked fish DID NOT MOVE from area to area between sampling events, pool the data and calculate one estimate of abundance (see next page if marked fish did move).

      - If sampling during the FIRST EVENT was spread evenly across the lake or up the stream, this estimate is germane to the entire lake or stream.

      - If sampling during the FIRST EVENT was constrained leaving large expanses of lake or stream with no sampling effort, the abundance estimate is a minimal estimate germane to parts of the stream or lake in the immediate vicinity of sampling.

- Unequal probabilities of capture . . or failure to mix . . (continued)
  - If the fraction . . . is similar in all areas . . . (continued)

    - If inspection of data shows marked fish DID MOVE from area to area between sampling events, pool the data and calculate one estimate of abundance. The estimate will be germane to the entire lake or stream regardless of how sampling was or was not constrained during the FIRST EVENT.

  - If the fraction in the population comprised of marked fish is dissimilar among areas of the lake or stream, there is no evidence that "OR" conditions have been met:

- If inspection of data shows marked fish DID NOT MOVE from area to area between sampling events, estimate abundance for each area of the stream or lake.

  - If sampling during the FIRST EVENT was spread evenly across the lake or up the stream, the sum of the estimates across areas is an unbiased estimate of abundance for the entire lake or stream.

  - If sampling during the FIRST EVENT was constrained leaving large expanses of lake or stream with no sampling effort, the sum of the estimates across areas is a MINIMAL ESTIMATE germane to parts of the stream or lake in the immediate vicinity of sampling.

- If inspection of data shows marked fish DID MOVE from area to area, estimate abundance with the method of Darroch (1961) for the entire population (see p. 27).

SIZE-SELECTIVE SAMPLING

Detection

Results of Hypothesis Tests (K-S and χ^2) on Lengths of Fish MARKED during the First Event and RECAPTURED during the Second Event	Results of Hypothesis Tests (K-S) on Lengths of Fish CAPTURED during the First Event and CAPTURED during the Second Event
Case I:	
Accept H_o	Accept H_o
There is no size-selectivity during either sampling event.	
Case II:	
Accept H_o	Reject H_o
There is no size-selectivity during the second sampling event but there is during the first.	
Case III:	
Reject H_o	Accept H_o
There is size-selectivity during both sampling events.	
Case IV:	
Reject H_o	Reject H_o
There is size-selectivity during the second sampling event; the status of size-selectivity during the first event is unknown.	

Case I: Calculate one unstratified abundance estimate, and pool lengths, sexes, and ages from both sampling events to improve precision of proportions in estimates of composition.

Case II: Calculate one unstratified abundance estimate, and only use lengths, sexes, and ages from the second sampling event to estimate proportions in compositions.

Case III: Completely stratify both sampling events, and estimate abundance for each stratum. Add abundance estimates across strata to get a single estimate for the population. Pool lengths, ages, and sexes from both sampling events to improve precision of proportions in estimates of composition, and apply formulae to correct for size bias to the pooled data (p. 18).

Case IV: Completely stratify both sampling events and estimate abundance for each stratum. Add abundance estimates across strata to get a single estimate for the population. Use lengths, ages, and sexes from only the second sampling event to estimate proportions in compositions, and apply formulae to correct for size bias to the data from the second event.

Whenever the results of the hypothesis tests indicate that there has been size-selective sampling (Case III or IV), there is still a chance that the bias in estimates of abundance from this phenomenon is negligible. Produce a second estimate of abundance by not stratifying the data as recommended above. If the two estimates (stratified and unbiased vs. biased and unstratified) are dissimilar, the bias is meaningful, the stratified estimate should be used, and data on compositions should be analyzed as described above for Cases III or IV. However, if the two estimates of abundance are similar, the bias is negligible in the UNSTRATIFIED estimate, and analysis can proceed as if there were no size-selective sampling during the second event (Cases I or II).

Correction

Begin by estimating the conditional fractions from your samples:

$$p_{ij} = n_{ij}/n_i$$

where

n_i = the number sampled from stratum i in the mark-recapture experiment
n_{ij} = the number sampled from stratum i that belong to group j
p_{ij} = the estimated fraction of the fish in group j in stratum i

Note that $\sum_j p_{ij} = 1$. The variance for p_{ij} is:

$$V[p_{ij}] = \frac{p_{ij}(1 - p_{ij})}{n_i - 1}$$

The estimated abundance of group j in the population (N_j) is:

$$N_j = \sum_i p_{ij} N_i$$

where N_i = the estimated abundance in stratum i of the mark-recapture experiment. The variance for N_j is a sum of the exact variance of a product from Goodman (1960):

$$V[N_j] = \sum_i (V[p_{ij}]N_i^2 + V[N_i]p_{ij}^2 - V[p_{ij}]V[N_i])$$

The estimated fraction of the population that belongs to group j (p_j) is:

$$p_j = N_j/N$$

where $N = \sum N_i$. The variance of the estimated fraction can be approximated with the delta method (Seber 1982):

$$V[p_j] \approx \sum_i V[p_{ij}] \left\{ \frac{N_i}{N} \right\}^2 + \frac{\sum_i \{V[N_i]\ (p_{ij} - p_j)^2\}}{N^2}$$

CULLING GROWTH RECRUITMENT FROM "CLOSED" TWO-SAMPLE MARK-RECAPTURE EXPERIMENTS

The technique of isolation by age

Whenever age of a fish can be determined without resorting to its death, information on age can be used to remove effects of growth recruitment from a mark-recapture experiment. Those age groups that are fully-recruited to sampling during the first event are redefined as the population of interest. For instance, if all age 5 fish and older are fully-recruited to sampling, abundance of fish age 5 and older during the first sampling event is the statistic that will be calculated. In this specific case, only those marked fish age 5 and older during the first event and only those fish age (5+k) and older captured during the second event will be included in the experiment. The increment *k* is the number of years between sampling events (k=0,1,2,etc.). Growth recruitment to this redefined population is impossible. So long as mortality rates of marked and unmarked fish of the same age are the similar, an unbiased estimate of abundance results from the experiment.

Petersen's model (Chapman's or Bailey's) is still used to estimate abundance and its variance, only the data are coded in the following manner:

$$M' = \sum_{j=r}^{T} M_{ij} \qquad C' = \sum_{j=r+k}^{T} C_{(i+k)j} \qquad R' = \sum_{j=r+k}^{T} R_{(i+k)j}$$

where i is the year of the first sampling event (and the year for which abundance is estimated), T is the age of the oldest individual in either sample, r is the age that fish in the studied population are fully recruited to sampling, M is the number of fish marked and released alive during the first sampling event, C is the number captured during the second event, and R is the number recaptured. The numbers R', C', and M' are substitutions for R, C, and M in the traditional model. From this point on, calculations are the same and hypothesis tests are the same. If the population must be stratified by size, that stratification can be done on age. In this situation, size-selectivity in sampling during the second event may be due to age-specific differences in mortality rates if the hiatus between events is a year or longer.

For this procedure to work, age must be determined for every fish in the experiment. That is M' + C' - R' fish if marked fish are individually identifiable or M' + C' fish if they are not. Therefore, each fish must be sampled at least once (and maybe some twice) in such a way that its age can be determined.

The technique of Robson and Flick (1965)

Instructions:

I. Sort R measurements of length from RECAPTURED fish in the SECOND SAMPLING EVENT in ascending order and remove all redundant lengths to make of list of length R′.

II. Treat the list as boundaries of cells in a length frequency histogram and assign measurements of UNMARKED fish captured during the SECOND SAMPLING EVENT to these cells, thereby building a length frequency distribution for unmarked fish with R′ or R′+ 1 cells.

III. Calculate "running averages" for these length frequencies by "leaving one out". The running average for the first cell would be the average of all the frequencies. The running average for the second cell would be the average of all the frequencies excluding the frequency for the first cell. The running average for the third cell would be the average of all the frequencies excluding the frequencies for the first and second cells. Etc.

VI. Plot the running averages against the upper bound of each cell. From the origin outwards the plot should decline, then flatten out, and on the far right "go crazy" (see p. 21 for examples). *Where the plot flattens out is where growth recruitment is no longer significant.* The running average of the cell at the left-most edge of this plateau is an unbiased estimate of the number of unmarked fish for every marked fish in the population ($\bar{u}_{r+1}$). This cell before it is the last cell to be significantly influenced by recruitment and is called the rth cell.

V. Run a series of hypothesis tests (R′-1 tests) using FLICK.WK1. The rth cell will be the one with an unusual number of rejections below it (shorter lengths) and an expected number of rejections above it (longer lengths).

VI. Calculate abundance:

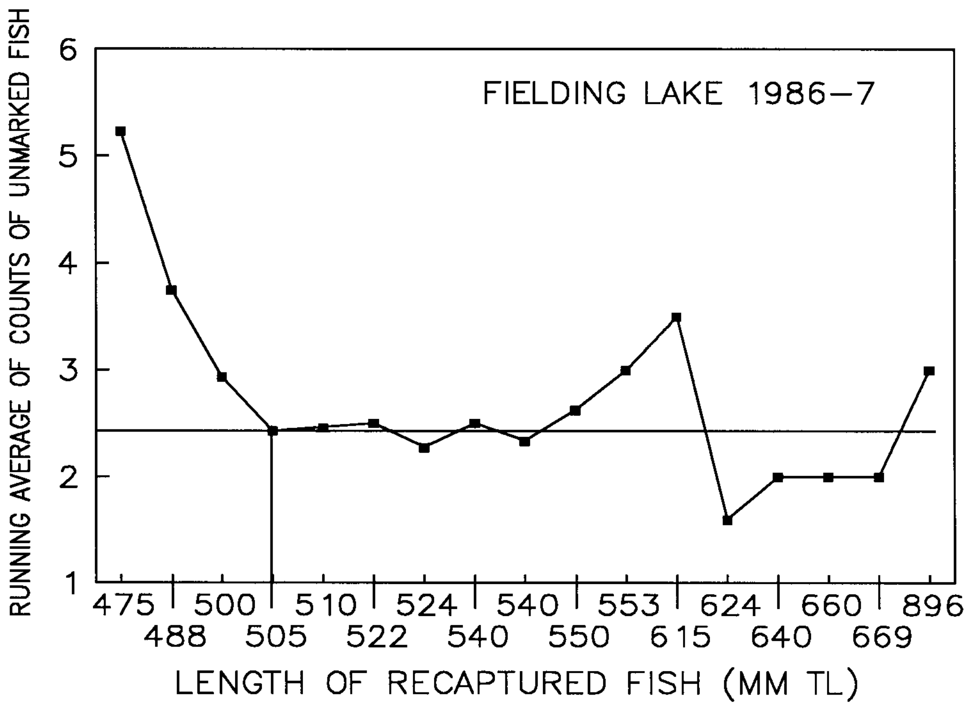
$$\hat{N} = (M + 1)(\bar{u}_{r+1} + 1) - 1$$

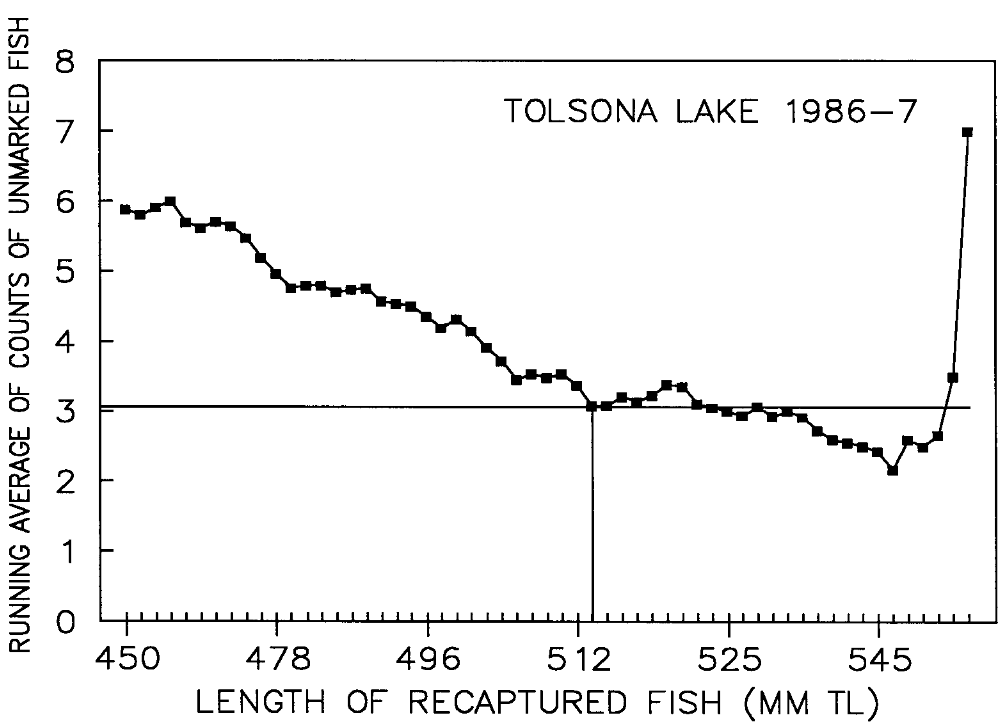
and its variance:

$$V[\hat{N}] \approx \frac{(M + 1)^2}{(R' + 1 - r)(R' - r)} \sum_{i=1}^{R'+1-r} (u_{r+i} - \bar{u}_{r+1})^2$$

R′	=	number recaptured second event with unique lengths
u	=	number of unmarked fish in a cell
M	=	number released with marks during first event
r	=	last cell influenced by recruitment

The estimate of abundance will be for the time of the FIRST EVENT.


FIELDING LAKE 1986-7
RUNNING AVERAGE OF COUNTS OF UNMARKED FISH
6
5
4
3
2
1
475 500 510 524 540 553 624 660 896
488 505 522 540 550 615 640 669
LENGTH OF RECAPTURED FISH (MM TL)


TOLSONA LAKE 1986-7
RUNNING AVERAGE OF COUNTS OF UNMARKED FISH
8
7
6
5
4
3
2
1
0
450 478 496 512 525 545
LENGTH OF RECAPTURED FISH (MM TL)

WARNING: **Fish released with marks during the first sampling event must be representative of the population - SIZE-SELECTIVITY in the sampling during the FIRST EVENT will bias estimates. Size-selectivity in the sampling during the second event is of no consequence.**

Example I - Culling growth recruitment during winter 1986-7 in the burbot population in Fielding Lake:

The table below is an example of the output from FLICK.WK1 for a population of burbot sampled in the summers of 1986 and 1987. Graphical representation of these data are in the upper plot on p. 21. Note that burbot grow and recruit to their populations during the winter, not the summer. Also note that all burbot ≥450 mm TL have equal probability of being caught with the sampling gear and procedures used in this study (Bernard et al. 1991).

Application of the nonparametric test described by Robson and Flick (1965) for growth recruitment. L(i) is the ith length of R', ordered, unique lengths of fish RECAPTURED during the SECOND SAMPLING EVENT, u(i) is the number of UNMARKED fish captured in the SECOND SAMPLING EVENT with lengths that fall between L(i) and L(i-1), U is the catch of unmarked fish during the second sampling event, and P[u>u(i)] is the Probability of a Type I error given $\bar{u}$(i+1--->R'). TO USE: 1) import unique lengths of recaptured fish in ascending order with cursor on A14, 2) import counts of unmarked fish with cursor at cell B14, 3) type {ALT}G, 4) check COL L for significant P[u>u(i)].

Lengths Recaptured Fish L(i)	Count Unmarked Fish u(i)	Q(i) = R'-i+1	U	U+Q(i)-u(i)	U-u(i)	U+Q(i)	Logarithms of Factorials of the Expressions Below: U	U+Q(i)-u(i)	U-u(i)	U+Q(i)	P[u>u(i)]
475	29	17	89	77	60	106	313.65	260.56	188.63	391.58	0.00 *
488	16	16	60	60	44	76	188.63	188.63	125.32	256.22	0.01 *
500	10	15	44	49	34	59	125.32	144.56	88.58	184.53	0.04 *
->505	2	14	34	46	32	48	88.58	132.95	81.56	140.67	0.50<-
510	2	13	32	43	30	45	81.56	121.53	74.66	129.12	0.50
522	5	12	30	37	25	42	74.66	99.33	58.00	117.77	0.17
524	0	11	25	36	25	36	58.00	95.72	58.00	95.72	1.00
540	4	10	25	31	21	35	58.00	78.09	45.38	92.13	0.24
550	0	9	21	30	21	30	45.38	74.66	45.38	74.66	1.00
553	0	8	21	29	21	29	45.38	71.25	45.38	71.25	1.00
615	13	7	21	15	8	28	45.38	27.89	10.59	67.89	0.01 *
624	0	6	8	14	8	14	10.59	25.19	10.59	25.19	1.00
640	2	5	8	11	6	13	10.59	17.49	6.57	22.55	0.36
660	2	4	6	8	4	10	6.57	10.59	3.16	15.10	0.33
669	1	3	4	6	3	7	3.16	6.57	1.76	8.51	0.57
896	3	2	3	2	0	5	1.76	0.65	ERR	4.77	ERR[a]
$	0	1	0	1	0	1	ERR	-0.08	ERR	-0.08	ERR

[a] Because there is no information beyond the last cell with which to compare the count from that cell, the ERR flags and negative numbers on the last line are to be expected.

The conclusion is to "break" the data at 500 mm TL with $r = 3$ and $\bar{u}_{r+1} = 2.43$. Calculation of estimated abundance and its variance are as follows:

i	Upper Bounds	Counts	Running Average	$(u_{r+i} - \bar{u}_{r+1})^2$
1	475	29	5.24	
2	488	16	3.75	
3	500	10	2.93	
4	505	2	2.43	0.18
5	510	2	2.46	0.18
6	522	5	2.50	6.61
7	524	0	2.27	5.90
8	540	0	2.33	5.90
9	550	0	2.63	5.90
10	553	0	3.00	5.90
11	615	13	3.50	111.76
12	624	0	1.60	5.90
13	640	2	2.00	0.18
14	660	2	2.00	0.18
15	669	1	2.00	2.04
16	896	3	3.00	0.33
17	∞	0	0.00	5.90
				159.70

Now with $r = 3$, $r + 1 = 4$, $\bar{u}_{r+1} = 2.43$, $R = 16$, and $M = 59$ (59 fish were released with marks in 1986):

$$\hat{N} = (59 + 1)(2.43 + 1) - 1 = 205$$

$$V[\hat{N}] \approx \frac{(59 + 1)^2}{(16 + 1 - 3)(16 - 3)} (159.70) = 3{,}159 \text{ (SE = 56)}$$

For the sake of comparison, the estimated abundance of burbot ≥450 mm TL in Fielding Lake in 1986 from a separate mark-recapture experiment is 213 (SE = 41). The estimate in this example is 205 (SE = 56).

Example II - Culling growth recruitment during winter 1986-7 in the burbot population in Tolsona Lake:

The table on the next page is an example of the output from FLICK.WK1 for a population of burbot sampled just before their lake froze in 1986 and just after it thawed in 1987. Graphical representation of these data are in the lower plot on p. 21. Note that burbot grow and recruit to their populations during the winter, not the summer. Also note that all burbot ≥450 mm TL have equal probability of being caught with the sampling gear and procedures used in this study (Bernard et al. 1991).

The conclusion is to "break" the data at 512 TL with $r = 31$ and $\bar{u}_{r+1} = 3.08$. With this break, $R = 56$ and $M = 517$ (517 fish were released with marks in 1986):

$$\hat{N} \quad = \quad (517 + 1)(3.08 + 1) - 1 = \quad 2{,}113$$

Since the sums of squares of the deviations from the mean are 109.85:

$$V[\hat{N}] \approx \frac{(517 + 1)^2}{(56 + 1 - 31)(56 - 31)} (109.85) \quad = \quad 45{,}347 \quad (SE = 212)$$

For the sake of comparison, the estimated abundance of burbot ≥450 mm TL in Tolsona Lake in 1986 from a separate mark-recapture experiment is 1,901 (SE = 120). The estimate in this example is 2,113 (SE = 212).

Lengths Recaptured Fish L(i)	Count Unmarked Fish u(i)	Running Average	P[u>u(i)]		Lengths Recaptured Fish L(i)	Count Unmarked Fish u(i)	Running Average	P[u>u(i)]	
450	10	5.88	0.20	x	513	3	3.08	0.43	
451	0	5.80	1.00		514	0	3.08	1.00	
452	1	5.91	0.86		515	5	3.21	0.25	
458	22	6.00	0.03	* x	516	1	3.13	0.76	
460	10	5.70	0.19	x	517	0	3.23	1.00	
461	1	5.62	0.85		518	4	3.38	0.35	
463	9	5.71	0.23		520	8	3.35	0.11	x
465	14	5.64	0.10	* x	523	4	3.11	0.32	
470	19	5.47	0.04	* x	524	4	3.06	0.31	
475	16	5.19	0.05	* x	525	4	3.00	0.31	
478	14	4.96	0.07	* x	526	1	2.94	0.75	
479	3	4.76	0.56		528	5	3.07	0.23	
482	5	4.80	0.39		529	2	2.93	0.55	
483	9	4.80	0.18	x	530	4	3.00	0.30	
484	3	4.70	0.56		534	5	2.92	0.21	
485	4	4.74	0.46		535	4	2.73	0.27	
489	12	4.76	0.10	* x	537	3	2.60	0.36	
490	6	4.57	0.30		539	3	2.56	0.36	
492	6	4.54	0.30		540	3	2.50	0.35	
494	10	4.50	0.13	x	545	4	2.43	0.22	
496	10	4.35	0.12	x	550	0	2.17	1.00	
497	0	4.19	1.00		555	3	2.60	0.35	
499	10	4.31	0.12	x	558	2	2.50	0.50	
500	12	4.15	0.07	* x	560	1	2.67	0.73	
503	10	3.91	0.10	* x	568	0	3.50	1.00	
505	12	3.72	0.05	* x	∞	7	7.00	ERR	
506	1	3.45	0.78						
508	5	3.53	0.28						
509	2	3.48	0.60						
510	8	3.54	0.13	x					
512	11	3.37	0.05	* x					

$\alpha =$ 0.10 (*)

10 of 31 significant (3-4 expected) | 0 of 25 significant (2-3 expected)

$\alpha =$ 0.20 (x)

17 of 31 significant (6-7 expected) | 1 of 25 significant (5 expected)

Other uses:

The following is an example of using techniques from Robson and Flick (1965) other than to estimate abundance (from Bernard et al. 1991).

> "The mark-recapture experiments described above on burbot populations in Tolsona and Fielding Lakes were also used to detect temporary aversion or inclination by burbot to be recaptured in hoop traps. The hypothesis of no trap-induced behavior was tested by comparing the fraction of the population with marks a few weeks after the release of marked fish to the fraction in a sample drawn from the population much later. Similarity between these two fractions would be due either to the subsidence of trap-induced behavior before the first opportunity for recapture or its continuance beyond the last opportunity. Prior to comparing these fractions, the nonparametric technique of Robson and Flick (1965) was used to cull burbot that had grown into the sampled population between sampling events. Immigration between sampling events was unlikely because Tolsona Lake is landlocked in the winter and Fielding Lake is relatively isolated from other lakes with burbot or from any large rivers."
>
> "In late August 1986, a sample of 49 burbot $\geq$450 mm TL (burbot fully recruited to the gear) from Fielding Lake contained 13 burbot that had been released from the sampling event three weeks earlier in late July. In late July 1987 some 51 weeks after marked burbot had been released into Fielding Lake, 107 fully recruited burbot were captured, 17 of which had been released with marks in July, 1986. Comparison of the lengths of marked and unmarked burbot in the sample in 1987 showed that there was a higher relative frequency of smaller fish with no marks than smaller fish with marks (Kolmogorov-Smirnov Two-Sample Test Statistic $D = 0.45$, $P = 0.003$) indicating that recruitment to the population through growth had occurred during the previous year. Growth recruitment above 505 mm TL was undetectable ($\alpha = 0.10$), and there were 2.44 unmarked burbot for every marked burbot above this length. Since this ratio is indicative of the unmarked to marked ratio for all the population (Robson and Flick 1965), the fraction of marked burbot in the sample drawn in 1987 adjusted for growth recruitment was 29%. The same fraction from the sample drawn in 1986 was 27% (=13/49)."

PARTIAL STRATIFICATION - DARROCH'S METHOD

Equations

There are situations in which none of the "OR" conditions have been met. Sampling has been locally (or temporally) concentrated and hypothesis tests have shown that marked fish had not completely mixed with unmarked fish between the two sampling events. If there has been no mixing of marked and unmarked fish among sampling strata between sampling events, several estimates based on some form of Petersen's model, one for each stratum, are added to produce a minimum estimate of abundance. The estimate is a minimum because it is only relevant to the areas (or times) sampled; because fish did not move among strata, nothing is known about the areas (or times) not sampled. When mixing of marked and unmarked fish has been partial, abundance of the entire population can be estimated with the methods of Darroch (1961).

The formulation of Darroch's method is a multidimensional expansion of Petersen's model. When the number of strata in the first sampling event(s) equals the number in the second event (t):

$$\hat{U} = \mathbf{u}' \mathbf{M}^{-1} \mathbf{a}$$

where $\hat{U}$ is the estimated abundance of unmarked fish in the population during the second sampling event, **u** and **a** are vectors and **M** is a matrix:

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ u_j \\ \cdot \\ u_t \end{bmatrix} \qquad \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \cdot \\ a_i \\ \cdot \\ a_s \end{bmatrix}$$

$$\mathbf{M} = \begin{bmatrix} m_{11} & m_{12} & \cdot\ \cdot\ \cdot & m_{1t} \\ m_{21} & m_{22} & \cdot\ \cdot\ \cdot & m_{2t} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & m_{ij} & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ m_{s1} & m_{s2} & \cdot\ \cdot\ \cdot & m_{st} \end{bmatrix}$$

where u_j is the number of unmarked fish in jth stratum during the SECOND sampling event, a_i is the number of marked fish released during the ith stratum during the FIRST sampling event, and the m_{ij} is the number of marked fish released into the ith stratum during the first event that were recaptured in the jth stratum during the second sampling event. The vectors **a** and **u** correspond to the scalar values of M and C and the matrix **M** to R in Petersen's model. Since the matrix **M** must be inverted ($\mathbf{M}^{-1}$ is the inverse of **M**), it must be of full rank and non-singular. Once abundance of unmarked fish has been estimated, the estimate of the entire abundance is:

$$\hat{N} = \hat{U} + \sum_{i=1}^{s} a_i$$

This method has all the same conditions as does Petersen's model, except of course, the "OR" conditions are not a concern here.

Because methods to calculate variance from Darroch's method are approximate and because statistical bias of Darroch's method has not been investigated (Seber 1982), bootstrap methods of Efron (1982) are used to both estimate variance and statistical bias for specific experiments. Each individual fish in an experiment has a capture history. For instance, in an experiment with two strata during each sampling event (s=t=2), there are eight possible capture histories:

Caught in stratum A during the 1st event
    and recaptured in stratum A during the 2nd event

Caught in stratum A during the 1st event
    and recaptured in stratum B during the 2nd event

Caught in stratum A during the 1st event
    and not recaptured during the 2nd event

Caught in stratum B during the 1st event
    and recaptured in stratum A during the 2nd event

Caught in stratum B during the 1st event
    and recaptured in stratum B during the 2nd event

Caught in stratum B during the 1st event
    and not recaptured during the 2nd event

Not caught in the 1st event
    but captured during the 2nd event in stratum A

Not caught in the 1st event
    but captured during the 2nd event in stratum B

Each bootstrap estimate is based on n fish sampled randomly with replacement from the tallied capture histories built on the original data. The capture histories of the n resampled fish are accumulated to produce a new set of vectors $\mathbf{a}^*$ and $\mathbf{u}^*$ and the matrix $\mathbf{M}^*$ that are used with the equations above to produce a single bootstrap estimate. The process is repeated until 100 to 1000 such bootstrap estimates have been calculated; the lower number is strictly for estimating statistics while the higher is used to build confidence intervals. The overall estimate of abundance and its variance are:

$$\bar{N}^* = \frac{\sum_{k=1}^{B} N_k^*}{B} \qquad V[\bar{N}^*] = \frac{\sum_{k=1}^{B} (N_k^* - \bar{N}^*)^2}{B - 1}$$

where B is the number of bootstrap estimates made (sets of n resampled fish) and N_k^* is the estimated abundance from the kth bootstrap sample. The final statistics are:

Abundance:	$\hat{N}$
Variance:	$V[\bar{N}^*]$
Statistical Bias:	$\lvert \hat{N} - \bar{N}^* \rvert$

As to be expected, the statistical bias is large when sample sizes are small.

Another set of useful statistics from Darroch's method are the probabilities of capture in the strata during the second sampling event (p_j):

$$\hat{\mathbf{p}} = \mathbf{M}^{-1}\,\mathbf{a}$$

$$\mathbf{p} = \begin{bmatrix} 1/p_1 \\ 1/p_2 \\ . \\ 1/p_j \\ . \\ 1/p_t \end{bmatrix}$$

All probabilities must meet the condition $0 < p_j \leq 1$ for all strata (probability of capture can not be zero, negative, or greater than one). If this condition is not met in estimating abundance from the original data, the overall estimate is seriously flawed. Conversely, the more bootstrap estimates that have impossible probabilities of capture, the more statistical bias is inherent in the overall estimate.

Abundance can also be estimated by stratum during the second sampling event by using this formulation of the equations:

$$\vec{\mathbf{U}} = \mathbf{D_u}\,\mathbf{M}^{-1}\,\mathbf{a}$$

$$\vec{\mathbf{U}} = \begin{bmatrix} U_1 \\ U_2 \\ . \\ U_j \\ . \\ U_t \end{bmatrix} \qquad \mathbf{D_u} = \begin{bmatrix} u_1 & 0 & 0 & . & . & 0 \\ 0 & u_2 & 0 & . & . & 0 \\ . & . & u_3 & & & 0 \\ . & . & & u_j & & . \\ . & . & & . & . & . \\ 0 & 0 & 0 & . & . & u_t \end{bmatrix}$$

The U_j are the number of unmarked fish in the population extant in the jth stratum during the second sampling event. Estimated abundance for each stratum must meet the condition $0 < U_j$ for all strata. If this constraint is violated in estimating abundance from the original data, then the estimate of overall abundance is seriously flawed. Conversely, the more bootstrap estimates that have "negative abundance" in one or more strata, the more statistical bias is inherent in the overall estimate.

Example: chinook salmon returning to the Kenai River

During 1988, a mark-recapture experiment was used to estimate the abundance of chinook salmon entering the Kenai River. Fish were caught just above tidewater in drift gill nets, measured to the nearest mm, marked with a numbered spaghetti tag and by removal of their adipose fin, and released. Because of the large mesh used in the gill nets, only chinook salmon age 1.3 and older were fully recruited to the sampling gear. To reduce handling stress, captured fish were immediately cut from the gill net and were held in a submerged, padded enclosure. The first sampling event began 20 May and ended 28 July. Sampling for the second sampling event was conducted in conjunction with a creel survey in the sport fishery upstream. Creels were inspected for marked and unmarked chinook salmon from 20 May through 31 July; 1,858 fish were inspected of which 61 were marked. Loss of tags was negligible. Because there was no significant differences between the length distributions of marked and recaptured fish (Kolmogorov-Smirnov Two-Sample Test, $D = 0.05$; $n = 2365,61$; $P = 0.99$), sampling during the second event was considered not to be size-selective.

Although sampling on this migratory population occurred at restricted locations, sampling effort was spread equally throughout the season in hopes of meeting the "OR" conditions and of using Petersen's model. To test the hypothesis that every fish had an equal chance of being captured during the first event (complete mixing of marked and unmarked fish is impossible here), data from the second event were arbitrarily broken into two-week periods as follows:

	Number w/ Marks	Number w/o Marks	Fraction w/ Marks
20-31 May	11	175	.06
1-15 June	18	364	.05
16-30 June	17	355	.05
1-15 July	6	306	.02
16-31 July	9	597	.01

Unfortunately, these fractions are significantly different ($\chi^2 = 16.44$, $df = 4$, $P < 0.01$) which means that Petersen's method should not be used to estimate abundance of the entire population. Therefore, all data were stratified into the following periods with like fractions of marked fish: 20-31 May, 1-30 June, and 1-28 July. These data are:

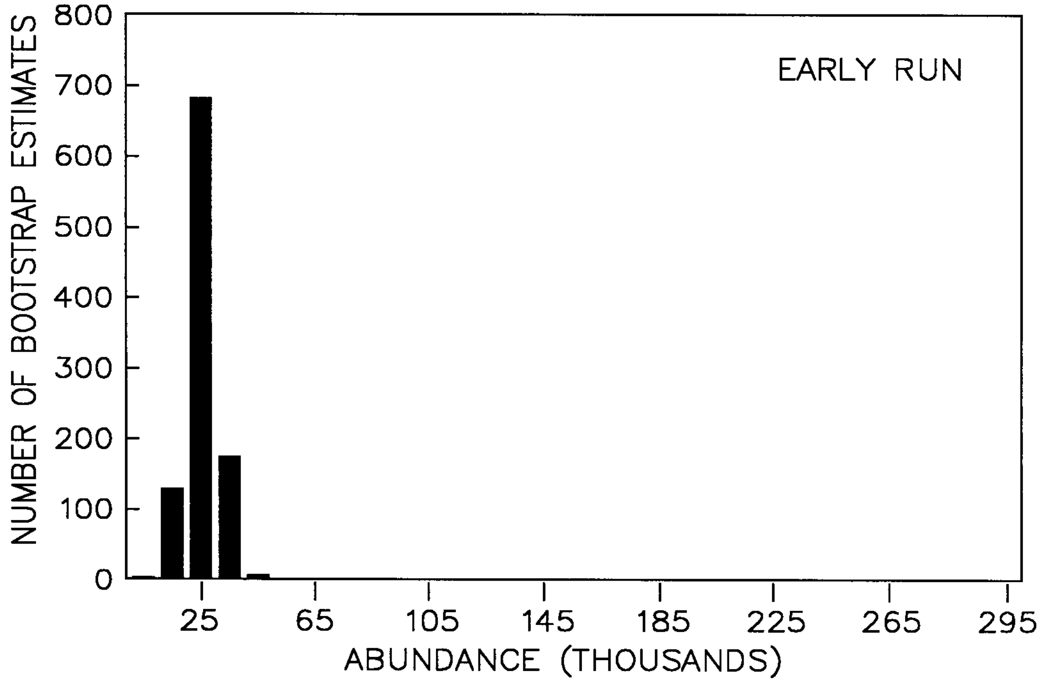
	Stratum	Recaptured in Second Event 1	2	3	Not Recaptured	% Not Recaptured
Release in	1	11	3	0	239	5.5
First Event	2	0	32	2	1,095	3.0
	3	0	0	13	1,240	1.0
Number Unmarked		175	719	903		
% Unmarked		5.9	4.6	1.6		

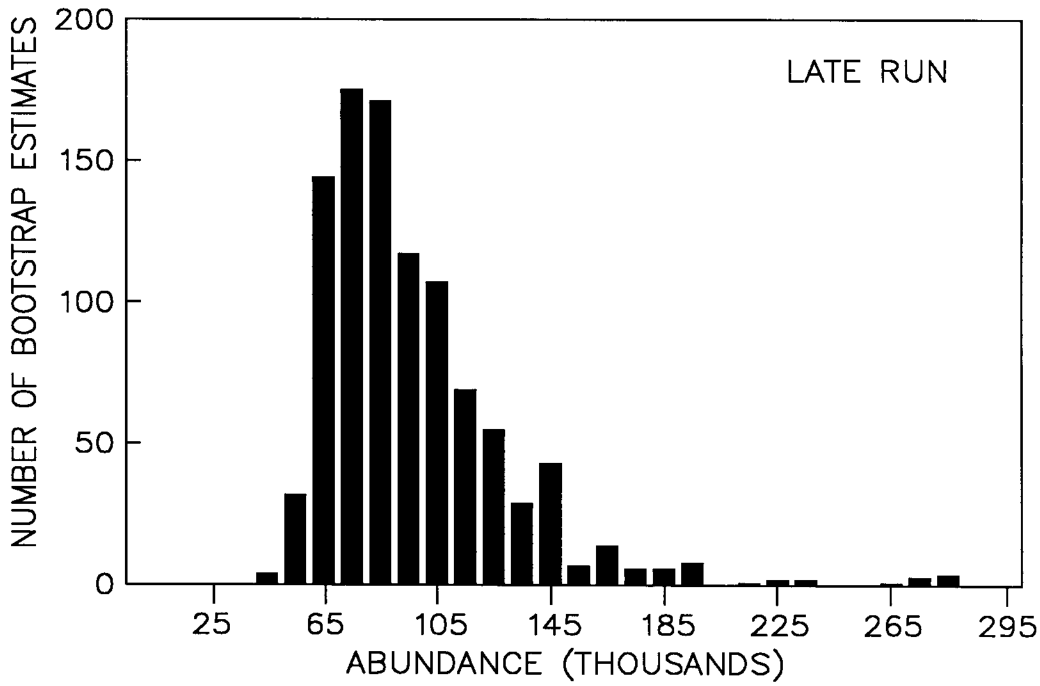
Since there was some mixing of marked and unmarked fish among strata, Darroch's method with bootstrapping was used to estimate abundance:

	Strata 1	2	1+2	3
From original data:				
U_j	2,629	21,036	23,665	87,035
N_j	2,882	22,165	25,047	88,288
From resampled data:				
U_j	2,806	21,276	24,082	95,379
N_j	3,059	22,405	25,464	96,632
$SE[N_j]$			5,496	33,436
Number:				
$p_j \leq 0$	18	3		0
$0 < p_j \leq 1$	978	997		1000
$p_j < 1$	4	0		0

The estimates from the first two strata were combined because this combined estimate is germane to the early run to the Kenai River; the third stratum corresponds to the late run which is a separate stock.

Estimated abundance for the early run is a solid estimate. Few bootstrap samples produced unrealistic probabilities of capture, and the estimated statistical bias is low (417 fish or about 2%). However, the precision of the estimate is poor (CV = 22%). Bootstrap estimates were almost symmetrically distributed around the mean:


EARLY RUN
NUMBER OF BOOTSTRAP ESTIMATES
800
700
600
500
400
300
200
100
0
25
65
105
145
185
225
265
295
ABUNDANCE (THOUSANDS)


LATE RUN
NUMBER OF BOOTSTRAP ESTIMATES
200
150
100
50
0
25
65
105
145
185
225
265
295
ABUNDANCE (THOUSANDS)

In contrast, estimated abundance for the late run is not as satisfying. Although no bootstrap samples produced unrealistic probabilities of capture, statistical bias is high (8,344 fish or about 9%), and precision in the estimate poor (CV = 38%). Distribution of bootstrap estimates is highly skewed toward greater abundance.

Hypotheses on why the experiment worked so poorly for the late run can be developed from rates of capture in the two sampling events, closure of the recreational fishery, and the presence of commercial fisheries. Inspection of data from the second event showed that a much lower fraction of the late run was marked than was of the early run. Since the sampling crews daily spent the same amount of time fishing during both runs, these fractions should have been similar. However, sampling effort was actually lower for the larger, late run. Effective sampling effort is the time gill nets are fished. In this experiment, gill nets were pulled when a fish was caught. For a large run, more fish are caught while the nets are fishing, but the nets are fished less. The result is an upper limit on the number of fish that can be marked. In this situation, the larger the run, the smaller the fraction of it can be marked, which makes the numbers of recaptured fish lower, variances higher, and biases greater. The second hypothesis is that fish marked late in first event were not fully available to the fishery by the time it closed 1 August.

The third hypothesis concerns interception of marked fish from the late run in commercial fisheries. Radiotelemetry has shown that some chinook salmon in the Kenai River back down the river before moving upstream. Some actually move back into Cook Inlet. Data from other studies around Alaska are consistent with this behavior for marked fish. If handling causes this behavior in chinook salmon, a higher fraction of the marked population than of the unmarked population will move back out to sea where they will be exposed to commercial drift and set net gill net fisheries for sockeye salmon. Since these commercial fisheries do not begin until late June, the "backing out" of marked fish does not affect the estimated abundance of the early run.

JOLLY-SEBER'S MODEL

Abundance, survival rates, and recruitment can be estimated with the techniques of Jolly (1965) and Seber (1965). All of the requirements of the previous models apply here with two important exceptions: recruitment and mortality can occur between sampling events without danger of biasing the abundance estimates. In an experiment based on these methods, recruitment can be due to either growth or immigration. With this relaxation of the assumptions behind closure of the population, sampling events can be spaced far apart so as to promote the complete mixing of marked and unmarked fish. Therefore, the short-term availability of fish to the sampling gear is not as critical with the Jolly-Seber models as with models used in experiments on closed populations. However, selectivity for fish of different sizes by the sampling gear is still a problem with this methodology.

Estimating numbers of marked fish

The Jolly-Seber models are based on estimation of the number of marked fish in the population just prior to each sampling event through the comparison of two marked populations. For example, the day before the start of any sampling event, a population subjected to a multi-year mark-recapture experiment can be divided into two components:

1) Marked fish and

2) Unmarked fish.

During a sampling event, fish are captured, inspected for marks, and unmarked fish are marked; all live fish are released. On the day after a sampling event, the population can be divided into three components:

1) Marked fish that had not been caught during the sampling event;

2) Marked fish that had been caught during the sampling event; and

3) Unmarked fish.

If during the next sampling events all marked fish have equal probability of being captured, then the rates of recapture from these two marked groups of fish will be the same:

$$\frac{\text{Number in Group (1) Recaptured in Future Sampling}}{\text{Number in Group (1)}} = \frac{\text{Number in Group(2) Recaptured in Future Sampling}}{\text{Number in Group (2)}}$$

In the not so friendly, but more convenient notation for the ith sampling event:

$$\frac{z_i}{M_i} = \frac{r_i}{R_i}$$

Since fish recaptured during a sampling event are considered part of Group (2) and no longer part of Group (1), the number in Group (1) is adjusted accordingly:

$$\frac{z_i}{M_i - m_i} = \frac{r_i}{R_i}$$

in which m is the number of fish from Group (1) recaptured. Note that the m_i becomes part of the R_i. In the above equation, R_i, m_i, and eventually z_i and r_i are known; M_i is only unknown. Solving this equation gives:

$$\hat{M}_i = \frac{z_i R_i}{r_i} + m_i$$

Estimating abundance

Note that $\hat{M}_i$ is the estimated number of marked fish in the population *just before* the start of the ith sampling event. Assuming that marked and unmarked fish have the same probability of being captured during the ith sampling event, abundance of the population *just before* that sampling event can be estimated as:

$$\hat{N}_i = \hat{M}_i \frac{(m_i + u_i)}{m_i}$$

where u_i are the unmarked fish in the sample. Note that this is the core of Petersen's model, only M is estimated, not known.

There will be K-2 estimates of abundance in a mark-recapture experiment with K sampling events. Because z_K and r_K are needed to estimate M_K, yet can not be estimated until there is at least one more sampling event beyond K, no estimate of N_K is possible. Neither is an estimate of N_1 possible since there is no marked fish extant just before the start of the first sampling event, and $m_1 = 0$. Obviously K must be ≥ 3 to produce a single estimate of abundance.

Estimating survival rate

Remember $\hat{M}_i$ is the estimated number of marked fish in the population just before the start of the ith sampling event. The number of marked fish in the population just after the ith event is $M_i - m_i + R_i$. The survivors of these "$M_i - m_i + R_i$" fish will comprise the marked population just before the (i+1)th sampling event; their number will be M_{i+1}. Since the equations above can be used to estimate abundance of marked fish extant just before any sampling event in a series but the first, the survival rate of marked fish between the ith and (i+1)th sampling events can be estimated:

$$\hat{S}_{i,i+1} = \frac{\hat{M}_{i+1}}{\hat{M}_i - m_i + R_i}$$

If a mark-recapture experiment with K sampling events and K-1 periods, survival rates for only K-2 periods can be estimated because the z_K and r_K are needed to estimate M_K, yet are not available until there is at least one more sampling event beyond K. Obviously K must be $\geq$ 3 to produce a single estimated survival rate. The survival rate of marked fish between the first and second sampling events is estimated as:

$$\hat{S}_{1,2} = \frac{\hat{M}_2}{R_1}$$

because the number of marked fish in the population just after the first sampling event is known, not estimated.

Estimating surviving recruitment

Recruitment between sampling events is the last major statistic that can be gleaned from a mark-recapture experiment based on the Jolly-Seber method. Surviving recruitment to the population from any source (growth or immigration) between the ith and (i+1)th sampling events *that is still alive just before the (i+1)th event* is calculated as:

$$\hat{A}_{i,i+1} = \hat{N}_{i+1} - \hat{S}_{i,i+1} \{\hat{N}_i - m_i - u_i + R_i\}$$

The term $(m_i + u_i - R_i)$ decrements the population for any fish that are "killed" during a sampling event. Note that this is a minimum estimate of actual recruitment.

There will be K-3 estimates of surviving recruitment in a mark-recapture experiment with K sampling events. Since there is no estimate of abundance for the first and last sampling events, surviving recruitment can not be estimated for the first or last periods in the experiment. Obviously, K must be $\geq$ 4 to obtain estimates of recruitment.

Planning

As yet there is no means of determining sample sizes to meet objective criteria for statistics estimated with Jolly-Seber methods short of simulation. Sample sizes needed to estimate abundance from Jolly-Seber methods are larger than those needed to estimate abundance from Petersen's method. Since there are no real corrections in the Jolly-Seber methods for statistical biases from small sample sizes, statistics from this method are prone to this kind of bias more so than estimates from Petersen's method with the same sample sizes. Although resampling techniques (i.e., bootstrapping) can not be used to correct this bias from small sample sizes, it can be used to estimate it.

Sampling events in an experiment based on Jolly-Seber methods should be spaced far enough apart so that significant mortality occurs between events, but close enough together that this mortality does not reduce the marked population to the point where few fish can be recaptured. If sampling events are so far apart that most of the marked fish die between them, then the few recaptures will produce considerable statistical bias in the estimates. If two sampling events are so close that little recruitment or little mortality occurs between them, these events should be collapsed into a single event (see section below on Meshing Experiments). If they are not collapsed, two sets of imprecise statistics will be generated as estimates for what is in reality the same parameters. These statistics will be imprecise because the sample sizes upon which they are based will be smaller than need be.

The conditions for accurately using Jolly-Seber methods are the same as those for Petersen's method except that recruitment and mortality can occur between sampling events. Every fish must have an equal probability of being captured during each sampling event or all marked fish must mix completely with unmarked fish between sampling events. Fish must retain their marks, and capture must not affect behavior of fish. And finally, the longer the hiatus between two sampling events, the better the chance that marked fish will spatially mix with unmarked fish between events.

<u>Software and inputs</u>

There are several programs for the personal computer that can be used to produce statistics for the Jolly-Seber method. The programs available from personnel in the Division of Sport Fish are: JOLLY, JOLLYAGE, POPAN-2, POPAN-3, and RECAP. Data are input into programs such as these either as a "B-table" or as capture histories. An example of the B-table is:

	Statistics by Event							
	1	2	3	4	5	6	7	8
Recaptured for the FIRST TIME from Event 1	0	123	35	14	5	3	5	9
Recaptured for the FIRST TIME from Event 2		0	79	32	33	18	11	5
Recaptured for the FIRST TIME from Event 3			0	51	36	13	11	8
Recaptured for the FIRST TIME from Event 4				0	45	13	4	5
Recaptured for the FIRST TIME from Event 5					0	63	14	8
Recaptured for the FIRST TIME from Event 6						0	22	9
Recaptured for the FIRST TIME from Event 7							0	21
Captured	531	502	349	206	349	239	249	195
Released with Tags	531	497	349	206	348	239	249	195

Each column of the B-table is a list of statistics for a particular sampling event. If there are K sampling events, there will be K columns. The first

K-1 rows in the B-table are numbers of fish recaptured for the FIRST TIME during a sampling event. For instance, the "32" in the second row, fourth column means that 32 fish in the experiment had been captured during the second event, not seen during the third event, and had been recaptured during the fourth event. The "9" in the first row, eighth column means that 9 fish had been captured during the first event, but had not been caught again until being recaptured during the last sampling event. Once a fish has been recaptured during a sampling event, that fish is considered a "freshly" released, marked fish from that sampling event. The Kth row is the number of fish captured during a sampling event, and the (K+1)th row is the number of fish released with marks during that sampling event.

The B-table is tricky to build, is easy to read, and is easy to "store" on a single sheet of paper. The B-table has the disadvantage of being difficult to resample (i.e., use in bootstrapping).

Capture histories are a matrix with K+1 columns and 2^K-1 rows. Each row corresponds to a different capture history, the first K columns correspond to the K sampling events, and the (K+1)th column corresponds to the number of fish in that experiment with that capture history:

1	2	3	4	5	6	7	8	n
0	0	0	0	0	0	0	1	21
0	0	0	0	0	0	1	0	20
0	0	0	0	0	0	1	1	2
0	0	0	0	0	1	0	0	60
0	0	0	0	0	1	0	1	1
0	0	0	0	0	1	1	0	2
0	0	0	0	0	1	1	1	0
0	0	0	0	1	0	0	0	40
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
1	1	1	1	1	1	1	0	0
1	1	1	1	1	1	1	1	0

For instance, the history "0 0 0 0 0 1 0 1" is the set of fish that were not captured until the 6th sampling event, not caught on the 7th event, but were recaptured on the 8th event. There was one fish with this capture history. The history "0 0 0 0 1 0 0 0" is shared by 40 fish; these were fish that were captured for the first time during the 5th sampling event and have yet to be recaptured in the experiment.

Capture histories are difficult to store, difficult to read, but are relatively easy to build, especially with a data-base manager. Capture histories are also easy to use in resampling.

Changing estimates with time

In an ongoing mark-recapture experiment, estimates of abundance, survival rate, and recruitment from the Jolly-Seber method change from year to year. For instance, abundance of fish in 1989 was estimated at 1,237 in 1990 for a hypothetical population. In 1991, abundance of fish in 1989 in the same population was estimated at 1,589. In 1992, abundance of fish in 1989 was estimated as 1,345. Eventually as the years progress beyond 1989, estimated abundance for that year will no longer "skip about". The same phenomenon is observed for estimates of survival rates and recruitment.

This temporary "instability" in estimates is from the ongoing recapture of marked fish. Remember back to statistics z_i and r_i. These statistics are defined as numbers recaptured in *future* sampling. This means *recapture in more than one sampling event beyond the ith event.* In the example in the previous paragraph, estimates of abundance for 1989 made in 1990 are based on fish recaptured during only one sampling event (1990). In 1991, the estimate of abundance is based on fish recaptured in two events: 1991 and 1990. In 1992, abundance in 1989 is estimated based on recaptured fish from three events: 1992, 1991, and 1990. Obviously, the estimate of abundance based on the most sampling events is the best. Eventually, all marked fish associated with the ith sampling event die, and estimates remain constant for that event.

Meshing experiments

If several two-event experiments with models based on closed populations have been conducted over a series of years on the same stock, experiments based on open and on closed populations can be combined [see Pollock (1982)]. The advantage in pooling data from two-event experiments into one event is that 1) sample sizes are increased and 2) statistics for the first year in the multi-year experiment can be calculated.

Data from two-event mark-recapture experiments on closed populations can be pooled because of the condition that abundance between sampling events does not change (the population is closed). While mortality and recruitment will change abundance of any population between any two sampling events, recruitment and mortality is negligible during the short hiatus in many mark-recapture experiments based on Petersen's model. If so, data from two parent events can be pooled to produce a single event by simply ignoring all those fish released with marks during the first of the parent events and recaptured during the second. For instance, $R_i = M + C - R$ and $(u_i + m_i) = C$ if you prefer the notation of Ricker (1975).

Caveats

Mark-recapture experiments based on Jolly-Seber methods are difficult to implement and "tricky" to analyze. Presence of size-selective sampling reduces chances for successfully using this approach. Splitting populations in two or three parts to accommodate size-selective sampling can reduce sample sizes for each subpopulation to the point where statistical bias is fatal. Also, difficulties arise when the survival rates of fish in all but the subpopulation with the largest fish reflect both mortality and growth recruitment out of the subpopulation.

If estimates of abundance, survival rates, or surviving recruitment "do not make sense", often some mistake has been made in coding or managing the data prior to final analysis. Setting up tagging histories and B-tables are difficult, and until the basics are well understood, mistakes are easily made. If abundance drops by 50% while CPUE in sampling gear doubled from the previous year, look for an error in your manipulation of the data. If an estimated survival rate is significantly greater than unity, look for an error. If surviving recruitment is significantly lower than zero, look for an error. The statistics obtained from the mark-recapture experiment must be consistent with other information. If recruitment is exceptionally large one year, then size distributions should be skewed to smaller fish. However, there are occasions when there is an inconsistency and no error has been made. These cases are rare and are exiting. These situations mean that we are about to learn something new about the population.

Great care must be taken in the field. Use only one kind of secondary mark each year, and change that mark every year. Repeat these secondary marks only after enough years have past that only an insignificant few fish with that mark remain from previous years. Be meticulous in the recording of the data, especially secondary marks on fish that have lost their primary mark. Failure to do so will not only compromise the data collected that year, but also the statistics for several sampling events before and after.

Example: burbot in Tolsona Lake

Since September, 1986, Jolly-Seber methods have been used to estimate abundance, survival rates, and surviving recruitment for the burbot population in Tolsona Lake. Tolsona Lake has 120 ha of surface, a maximum depth of 5 m, and is landlocked. Sixty to 120 baited hoop traps have been set for two days during each sampling event, and all burbot captured were measured to the nearest mm TL and marked with an individually numbered tag and with removal of a fin. The same fin was removed from all fish captured during a year, and each year a different fin was selected for removal. Since sampling burbot to estimate their age is lethal, no data on age were collected. Data collected early in the program showed that burbot $\geq$450 mm TL are fully recruited to the sampling gear. Other studies in Alaska and elsewhere have shown that this is the size at which burbot mature, switch to a fish diet, and enter hook-and-line fisheries. Analysis for purposes of fisheries management was therefore restricted to this group of fully recruited burbot. Subsequent analysis of data collected in this stock assessment program showed that these fully-recruited burbot exhibited no "trap happiness" or "trap shyness" and that marked, fully-recruited burbot mixed completely with their unmarked brethren. Because the lake is shallow, there were no problems with decompression in sampled burbot. Measured rate of tag loss was less than a few percent per year; measured rate of fin regeneration was 0%. Estimates of survival rates, recruitment, and abundance from Jolly-Seber and other methods for burbot $\geq$450 mm TL in Tolsona Lake are (SEs are in parentheses):

Midpoint of Sample Dates	Days in Hiatus	Abundance	Survival Rate	Surviving Recruitment
9/26/86		1,901 (120)		
	237		0.60 (0.05)	159 (170)
6/3/87		1,300 (121)		
	336		0.73 (0.07)	599 (133)
5/26/88		1,545 (162)		
	96		0.77 (0.09)	22 (118)
9/01/88		1,214 (148)		
	267		0.79 (0.14)	629 (139)
5/24/89		1,590 (191)		
	112		0.91 (0.17)	85 (146)
9/14/89		1,535 (276)		
	241		0.66 (0.18)	1,067 (323)
5/25/90		2,085 (512)		

Mark-recapture experiments in 1986 and 1987 were designed as two-event experiments based on Petersen's model; sampling events were but a couple of weeks apart. Experiments in 1988, 1989, and 1990 also had two sampling events each, however, these sampling events were several months apart so survival rates and surviving recruitment could be estimated for winter and for summer. Events from experiments in 1986 and 1987 were "collapsed" into a single event for each year making 8 sampling events in this experiment.

Results from the experiment showed the effects of the fishery on the population dynamics of the stock:

Years/Months	Fishery Status	Annual Survival Rate	Surviving Recruitment Annual		Summer		Winter
1986-7 Oct-May	Fishery Closed January	0.46	159	=			159
1987-8 Jun-May	Fishery Opened November	0.71	599				
1988-9 Jun-May	Fishery Open	0.61	651	=	22	+	629
1989-0 Jun-May	Fishery Open	0.59	1,152	=	85	+	1,067

The fishery for burbot in Tolsona Lake is (was) a set-line fishery through the ice from November through April. Little interest was shown in the fishery in 1986 until results of our stock assessment program reached the public after October of that year. Within the first 30 days of the fishery, tags from over 10% of the fish we marked had been voluntarily returned. The fishery was closed in January by emergency order. The mark-recapture experiment showed a dramatic decline in abundance, survival rate, and surviving recruitment for that winter even with the early closure. The fishery was reopened in November, 1987 with stricter regulation of set lines and reduced interest by anglers. The better survival rate for 1987-8 (0.71) and the lower rates in later years (0.61 and 0.59) is consistent with the fishery being turned off, then on with tighter regulation and reduced fishing effort. The increase of surviving recruitment in 1989-90 is born out by length-frequency diagrams of burbot captured during the later years of the program. This shot of recruitment is consistent with faster growth and better survival of young burbot that can be expected when a relatively unexploited population is "pulse fished."

Besides the obvious boost to fisheries management, the Jolly-Seber method has increased our knowledge on how mark-recapture experiments can be designed for lacustrine populations of burbot. Estimates of surviving recruitment for winter (629, 1067) have far out paced estimates for summer (22, 85). In fact, summer estimates are not significantly different than zero, indicating that whatever recruitment occurs is negligible even though some growth does occur. This phenomenon has been observed in populations in two other lakes. The implication is that with negligible recruitment during the summer, a two-event experiment with the events placed anywhere within the summer months is essentially a "closed" experiment. This information adds flexibility to designing mark-recapture experiments for stocks of this species in Alaska's lakes.

BOOKKEEPING

The following text is taken from the Results Section of Clark et al. (1988).

> "During the first sampling event (2 - 9 June 1987), 2,199 northern pike were caught (2,061 fish were tagged and released, 117 fish were released without being tagged, and 21 fish died during sampling). Thirteen of the 2,061 fish caught, marked, and released during the first sampling event were tagged and released in 1986. In 1986, 279 northern pike were tagged and released (65% caught in seines, 25% caught in fyke nets, and 10% caught in gill nets).
>
> During the second sampling event (23 - 29 June 1987), 793 northern pike were caught (87 were recaptured fish, 228 fish were tagged and released for the first time, 471 fish were released without being tagged, and 8 fish died during sampling including 1 recaptured fish from the first sampling event). Of the 87 recaptured northern pike caught during the second sampling event, 4 were fish released during 1986 (3 tagged fish and 1 fish that had lost its tag) and 83 were fish marked and released during the first sampling event (78 tagged fish including 1 mortality and 5 fish that had lost their tags). Fish that had lost their tags were retagged during the second sampling event. Tag loss for fish tagged during 1986 was estimated to be 5.9% (1 tag loss of 17 recaptured northern pike). Tag loss for fish tagged during the first sampling event in 1987 was estimated to be 6.0% (5 tag losses of 83 recaptured northern pike).
>
> <u>Abundance</u>
> Estimated abundance of northern pike over 299 mm FL in George Lake in June 1987 was 17,662 (standard error = 2,105 fish). In 1987, 1,051 fish were released into "Area "A" and 1,010 fish were released into "Area B" of George Lake (Figure 2) during the first sampling event. During the second sampling event, 744 northern pike over 299 mm FL, including 83 recaptured fish, were caught. One hundred two northern pike were captured in "Area A" during the second event, including 17 recaptured fish (14 of these fish had been released in "Area A", 2 in "Area B", and 1 had lost its tag). Six hundred forty-one northern pike were captured in "Area B" during the second event, including 62 recaptured fish (19 of these fish were released in "Area A", 43 in "Area B" and 4 had lost their tags). . . . . . ."

The following appendices are taken from the Results Section of Clark et al. (1988).

Appendix Table 1. Numbers of northern pike tagged in each of two areas in George Lake 2 - 9 June and their location 23 - 29 June 1987.

	Area of Recapture		
Area of Release	A	B	Not Recaptured
Released in A	14	19	1,018
Released in B	2	43	965
	$\chi^2 = 18.90$[1]	$P < 0.005$	DF = 2

[1] The χ^2 value is the test statistic for the hypothesis of equal probability of recapturing fish in either half of George Lake (Seber 1982).

Appendix Table 2. Numbers of tagged and untagged northern pike captured by area in George Lake 23 - 29 June 1987.

	Area	
Category	A	B
Recaptured Fish	16 (17)[1]	62 (66)
New Fish	86 (85)	579 (575)

$\chi^2 = 3.39$[2]	$0.05 < P < 0.010$[3]	DF = 1
($\chi^2 = 3.60$	$0.05 < P < 0.10$	DF = 1)[4]

[1] Marked fish that had lost their tags were considered unmarked fish in the reported analysis. The correct figures are in parentheses.
[2] The χ^2 value is the test statistic for the hypothesis of equal probability of capturing tagged fish in either half of George Lake (Seber 1982).
[3] The probability statement should read "$0.05 < P < 0.10$"
[4] The results of the hypothesis test based on the adjusted contingency table.

Appendix Table 3. Numbers of northern pike tagged by length class in George Lake 2 - 9 June then sampled 23 - 29 June 1987.

Length Class (mm FL)	Recaptured		Not Recaptured
300 - 349	6		127
350 - 399	9		249
400 - 449	12		329
450 - 499	19		361
500 - 549	19		362
550 - 599	12		251
Over 600	6		299
	$\chi^2 = 5.88$[1]	$0.25 < P < 0.5$	DF = 6

[1] The χ^2 value is the test statistic for the hypothesis of equal probability of capturing fish across the seven length classes.

LAB NO. 1: HYPOTHESIS TESTING

Problem 1

During the third week in September *(1986)*, 683 fish *(burbot)* were released into the lake *(Tolsona Lake)*; 149 of these fish were recaptured during sampling two weeks later. The population of released fish were divided into the following groups according to their total length:

	Number Released	Number Recaptured	p	
<450 mm	163	9	*.06*	*Group I*
450-499 mm	292	73	*.25*	*Group II*
500-549 mm	203	61	*.30*	*Group III*
≥550 mm	25	6	*.24*	*Group IV*

IvsIIvsIIIvsIV - $v2 = 35.24$, $df = 3$ $P < 0.01$
IIvsIIIvsIV - $v2 = 1.67$, $df = 2$ $0.25 < P < 0.50$
Ivs(II+III+IV) - $v2 = 33.33$, $df = 1$, $P < 0.01$

1) What are the rates of recapture for the four groups?

2) Are these rates significantly different? What is the risk you run of being wrong by concluding the rates are different?

3) Write out the hypothesis that you just tested. Now write out the alternative hypothesis.

   H_o: $p_l = p_o$ *for all lengths l*
   H_a: $p_l \neq p_o$ *for all at least one length*

4) If you had to split the population into two groups based on recapture rates, at what length would you make the split? (HINT: This activity requires several hypothesis tests).

   450 mm TL

Problem 2

During a week's sampling during the autumn 1986, several hundred fish *(burbot)* were captured, measured to the nearest mm TL, marked, and released into Wilson's *(Tolsona)* Lake. Two weeks later, over 100 of these fish were recaptured. The same sampling gear was used in the same way for both sampling events. Measurements by sampling event are in the file HYPOP2.WK1.

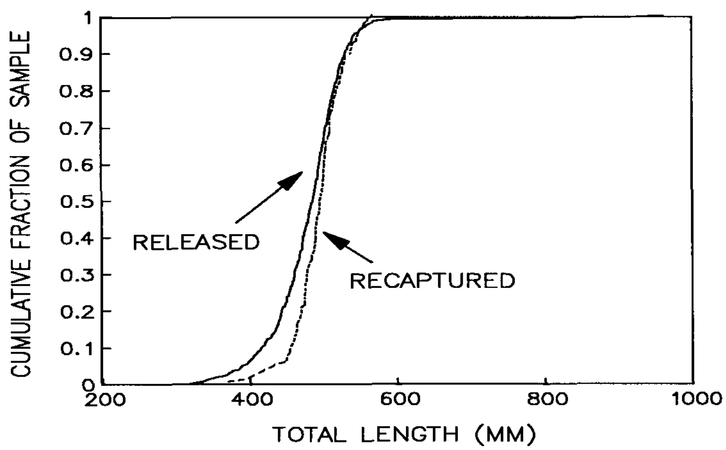
1) Are the length distributions for *released* and *recaptured* fish significantly different? What's the test statistic and the chance that you're wrong in this judgment?

   Kolmogorov-Smirnov Two-sample Test, $D_{max} = .21;\ n = 683{,}149;\ P < 0.01$

2) What's the hypothesis that you just tested. What's the alternative hypothesis.

   $H_o: F(l_{marked}) = G(l_{recaptured})$ $\qquad$ $H_a: F(l_{marked}) \neq G(l_{recaptured})$

   Plots of the cumulative distribution function for both samples are as follows (the dashed line corresponds to recaptured fish):

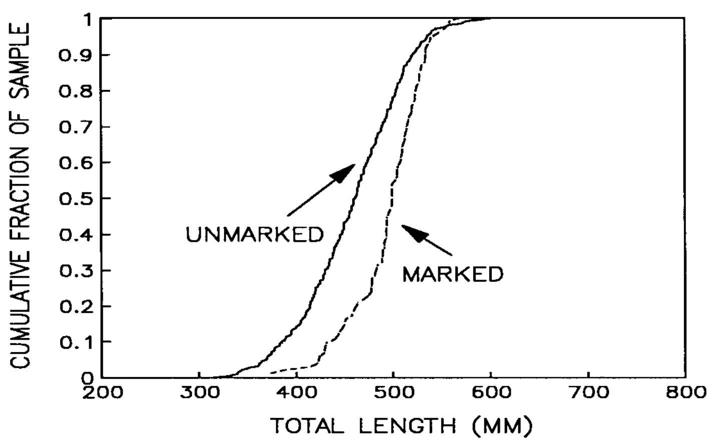


3) In what way do these two plots differ? What is the meaning of this difference?

   The lower tail for the recaptured fish is shifted towards larger fish which means that smaller fish were less likely to be recaptured than larger fish.

Problem 3

In the spring 1987 you return to Wilson's *(Tolsona)* Lake and capture 620 fish *(burbot)*, 106 of which had been marked during last year's sampling. The same sampling gear was used in the same way to capture fish in 1987 as was used in 1986. All fish were measured to the nearest mm TL. Cumulative distribution functions of lengths for MARKED (recaptured fish) and UNMARKED fish (captured for the first time in 1987) are plotted and compared:



The dashed line corresponds to recaptured (marked) fish. Data on marked and unmarked fish are in the file HYPOP3.WK1.

1) Are the length distributions for unmarked and marked fish significantly different? What's the test statistic and the chance that you're wrong in this judgment?

   Kolmogorov-Smirnov Two-sample Test, $D_{max} = .39$; $n = 514,106$; $P < 0.01$

2) Relationship between the two plots for this problem is graphically similar to the relationship between the two plots in Problem 2. However, these two relationships indicate different phenomena at work, both of which are important to successfully conducting mark-recapture experiments? What are these two phenomena?

   The phenomenon in Problem 2 is size-selectivity in the sampling in October; in this problem, its presence of growth recruitment.

Problem 4

One spring during spawning season *(in 1987)*, northern pike were captured at two locations (spawning grounds) at the far ends of a long lake *(George Lake)*. These fish were marked according to which end of the lake they had been captured and were released. A week later these spawning grounds were again visited, and more fish were captured. Fish in the second sample were of two types: marked or unmarked. Data collected during this second trip are:

	Number w/ Marks	Number w/o Marks	p	
North End	16	86	*.16*	*Group I*
South End	62	579	*.10*	*Group II*

IvsII - $\chi^2 = 3.87$, $df = 1$, $0.01 < P < 0.05$

1) What fraction of the population carries marks in each section of the lake?

2) Are these fractions significantly different? What is the risk you run of being wrong by concluding the fractions are different?

3) What's the hypothesis that you just tested. What's the alternative hypothesis.

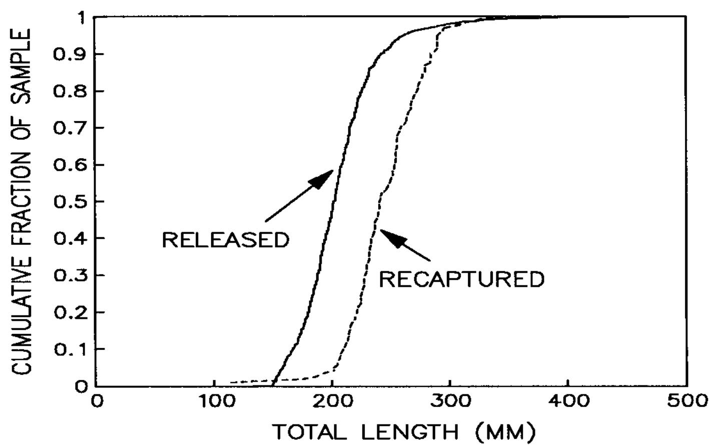
   $H_o: p_{North} = p_{South}$ $H_a: p_{North} \neq p_{South}$

4) If no marked fish had been released in the north end of the lake, what would these data suggest about behavior of these fish between samples?

   These fish move quickly across the lake at this time of the year.

Problem 5

Fyke nets were used to capture and mark 1,457 *(rainbow)* trout from a small lake *(Honeybee Lake)* during the spring *(1990)* of which 127 were recaptured during sampling in the autumn. All fish were measured to the nearest mm FL. Cumulative distribution functions of lengths for trout *marked* in the spring and trout *recaptured* in the fall were plotted and compared:



A list of measurements for trout released in the spring and of trout recaptured in the autumn are in the file HYPOP5.WK1.

1) Are the length distributions for released and recaptured fish significantly different? What's the test statistic and the chance that you're wrong in this judgment?

   Yes.
   Kolmogorov-Smirnov Two-sample Test, $D_{max} = .57$; $n = 1,457,127$; $P < 0.01$

2) This situation is analogous to the one in Problem 2 in that the same kinds of data are involved. However, plots of the cumulative distribution functions are radically different in this problem compared to those in Problem 2. Size-selectivity of sampling is the phenomenon that shaped the plots for Problem 2. A different phenomenon entirely molded the plots in this problem. What is this phenomenon? (HINT: Sampling events in Problem 2 were two weeks apart; in this problem, sampling events were several months apart. What can a fish do a lot of in several months that it can only do a little of in two weeks?)

   The phenomenon in Problem 2 is size-selectivity in the sampling in October; in this problem, its presence of growth.

Problem 6

A sport fishery *(on the Kenai River)* for adult *(chinook)* salmon was sampled from mid May to the end of July with a creel survey *(in 1988)*. The fishery is on the lower reaches of a large river below the spawning grounds. In samples from the survey, 1,858 fish were inspected; some of these fish had been marked while others had not. The arbitrary breakdown of samples into two-week periods is as follows:

	Number w/ Marks	Number w/o Marks	p	
20-31 May	11	175	*.06*	*Group I*
1-15 June	18	364	*.05*	*Group II*
16-30 June	17	355	*.05*	*Group III*
1-15 July	6	306	*.02*	*Group IV*
16-31 July	9	597	*.01*	*Group V*

IvsIIvsIIIvsIVvsV - $\chi^2 = 16.44$, $df = 4$, $P < 0.01$
IvsIIvsIII - $\chi^2 = 0.53$, $df = 2$, $0.90 < P < 0.95$
IVvsV - $\chi^2 = 0.25$, $df = 1$, $0.50 < P < 0.75$
(I+II+III)vs(IV+V) - $\chi^2 = 15.54$, $df = 1$, $P < 0.01$

1) What are the fractions of samples with marks for each two-week periods?

2) Are these fractions significantly different? What risk are you taking by concluding that they are different?

3) Write out the hypothesis that you just tested. Now write out the alternative hypothesis.

   H_o: $p_i = p_o$ *where i is a two-week period and p is fraction of population with tags.*

   H_a: $p_i \neq p_o$ *for at least one i*

4) Division of data into 5 two-week periods was arbitrary. Based on significant differences (if any) in the fractions of marked fish in the samples, what is a better division of the data? (HINT: This activity requires several hypothesis tests).

   20 May - 30 June and July

Problem 7

Fish *(burbot)* were captured at all depths in a deep lake *(Paxson Lake)* with gear resting on the bottom. Biologists noted that those fish captured at depths greater than 15 m were distressed, so much so that any remaining on the surface for a moderate amount of time died. Therefore, fish brought up from these depths were hurriedly marked and returned to the water. From 1986 through 1988, 2,442 fish were captured and 118 subsequently recaptured at least once:

Depth of First Capture	Number Recaptured	Number NOT Recaptured	p	
< 15 m	41	914	*.04*	*Group I*
≥ 15 m	77	1,410	*.05*	*Group II*

IvsII - $\chi^2 = 0.99$, $df = 1$, $0.25 < P < 0.50$

1) What are the rates of recapture for the two groups?

2) Are these rates significantly different? What is the risk you run of being wrong by concluding the rates are different?

3) Write out the hypothesis that you just tested. Now write out the alternative hypothesis.

   H_o: $p_{<15} = p_{\geq 15}$ H_a: $p_{<15} \neq p_{\geq 15}$

4) What do you conclude from this exercise about the survivability of fish initially captured below 15 m?

   Survival rates are the same as those for fish initially captured in shallower water.

LAB NO. 2: PETERSEN'S MODEL (PLANNING)

Problem 1

If you expect that the abundance of the population is 5,000 fish of which 300 have been marked during the first sampling event, how many must be inspected for marks during the second sampling event so that there is a 90% chance [$=(1 - \alpha)100$] that the estimate is within 25% (d) of the true value?

$C = 624$

Problem 2

Some of the chinook salmon returning to the Taku River are captured in the fishwheels at Canyon Island, marked, and released, and are sampled in four tributaries in Canada. Spawning occurs in tributaries other than those that are visited. The fishwheels are located in a canyon through which the entire river flows. What you want is a simple, mark-recapture experiment based on Petersen's model to estimate abundance of migrating chinook salmon.

1) Since not every fish has an equal probability of being inspected for marks, how must the fishwheels be operated at Canyon Island to increase the chances that every chinook salmon has an equal probability of being marked regardless of when it enters the river?

   The fishwheels must be operated 24 h a day, 7 d a week, on both banks simultaneously, and all captured fish must be marked.

2) Since bigger chinook salmon tend to migrate farther off shore, what additional measurements must be made to insure that estimated abundance will be unbiased? Do you expect this phenomenon to affect the analysis in the mark-recapture experiment? How so?

   Length of individual fish. Yes. The experiment will be completely stratified by length of fish.

3) Stream flow in the Taku River has traditionally been several hundred cfs in May when chinook salmon first enter the river; by late June when the last chinook salmon pass Canyon Island the stream flow is several thousand cfs. Since faster water tends to force salmon to migrate more along the shore, what affect will this phenomenon have at the fishwheels? Will this kind of sampling meet the condition that probabilities of capture be the same for all fish in the first sampling event? What action can you take to insure that the estimate of abundance will not be biased from this phenomenon?

   Probability of capture will be low in May and high in June. No. Fish must be marked so as to identify when they were marked at the fishwheels.

4) Last year the estimated abundance of chinook salmon age 1.3 and older in the Taku River was 6,543. One hundred ninety eight of these older fish were marked in the four fishwheels operated last year and 456 were inspected for marks in four tributaries with weirs and foot surveys. Based on the unusual number of chinook salmon age 1.2 that were captured in fishwheels last year and throughout southeast Alaska, you expect that the escapement of age 1.3 and older chinook salmon to the Taku River to be about double (13,000) last year's abundance. If you want an estimate of abundance that has a 90% chance of being within 25% (d) of the actual value, how many fishwheels should you operate at Canyon Island? (you can not talk the Canadians into expanding their sampling on the tributaries).

Six fishwheels. If the abundance doubles, then the number inspected in the tributaries and the number captured at Canyon Island double as well: 456 to about 900 and 198 to 400, respectively. If 900 are inspected on the spawning grounds, then about 600 need to be tagged to meet objective criteria:

$$585 = \frac{13{,}000\ (0.0471)}{1 + 0.0471} \qquad 0.0471 = \frac{45.5(13{,}000 - 900)}{(13{,}000 - 1)\ 900}$$

If 4 fishwheels are expected to catch about 400 older chinook salmon, then 6 fishwheels should catch 600.

5) How many fishwheels would you operate on each bank at Canyon Island?

Three on each bank.

6) If you fish the 4 wheels that you have next year, what would you expect your relative precision ($d = Z_{\alpha/2} \cdot SE[N]/N$) to be for your estimate of older chinook salmon? If you fished 8 wheels? Only 2 wheels? ($Z_{\alpha/2}$ = 1.645 when α = 0.10)

Four Wheels: If 4 wheels are fished at a catch rate of 100 per wheel, 400 of 13,000 fish will be caught. The expected number of recaptured animals in the tributaries would be (400/13,000)(900) = 28.

$$SE[\hat{N}] = \left\{ \frac{(13{,}000)\ (400-28)\ (900-28)}{(28 + 1)\ (28 + 2)} \right\}^{1/2} = 2{,}202$$

$$d = \frac{2{,}202\ (1.645)}{13{,}000} = 0.28$$

Eight Wheels: If 8 wheels are fished at a catch rate of 100 per wheel, 800 of 13,000 fish will be caught. The expected number of recaptured animals in the tributaries would be (800/13,000)(900) = 56.

$$\widehat{SE[N]} = \left\{ \frac{(13{,}000)\ (800-56)\ (900-56)}{(56 + 1)\ (56 + 2)} \right\}^{1/2} = 1{,}572$$

$$d = \frac{1{,}572\ (1.645)}{13{,}000} = 0.20$$

Two Wheels: If 2 wheels are fished at a catch rate of 100 per wheel, 200 of 13,000 fish will be caught. The expected number of recaptured animals in the tributaries would be (200/13,000)(900) = 14.

$$\widehat{SE[N]} = \left\{ \frac{(13{,}000)\ (200-14)\ (900-14)}{(14 + 1)\ (14 + 2)} \right\}^{1/2} = 2{,}988$$

$$d = \frac{2{,}988\ (1.645)}{13{,}000} = 0.38$$

7) The expected accuracy next year as a function of sampling effort with the fishwheels is as follows:

No. Fishwheels	d (relative precision)
2	0.38
4	0.28
6	0.23
8	0.20
10	0.18

If you desire a relative precision of 15% instead of 25%, what is your most cost effective means of meeting that objective?

Get the Canadians to sample more chinook salmon on the spawning grounds.

Problem 3

1) Trout are caught in four fyke nets set at equal distance around the shoreline in a round, 100 ha lake. Under what conditions will all trout have an equal probability of being captured?

Trout move so rapidly ALONG THE SHORE that they encounter at least one fyke net during a sampling event; AND trout move so rapidly ACROSS THE LAKE that they encounter at least one fyke net during a sampling event.

2) If fyke nets are the only gear available to you, how can you improve the chances that each trout has the same probability of capture regardless of its location in the lake?

Fish more fyke nets and/or fish them longer and spread them evenly along the shore.

3) Where can a traditional fyke net never be set? What potential problem will this cause in estimating abundance?

   The middle of the lake. If fish reside in the middle of the lake and movement across the lake is slow, estimated abundance will be biased because none of the "OR" conditions will be met.

4) During the first sampling event, you planned to fish four fyke nets spaced evenly around the lake until 500 fish have been captured. During the second night of sampling, someone stole one of the nets. Should you obtain a replacement or continue fishing with only three nets? Why or why not?

   Obtain a replacement that day. Without that net fishing, trout in that quadrant of the lake will have lower probability of being captured.

5) If you have reason to believe that 75% of the trout are along only half the shoreline, should fyke nets be set at a higher density along this shore to increase sample size? Why or why not? What if the species was not trout, but a species noted for its constant movement; would that change your mind?

   Aggregating sampling effort on concentrations of fish will increase sample sizes, but those fish in the area in which sampling is concentrated will have a higher probability of being captured than will fish outside the area. Under this situation, an unbiased estimate of abundance can be expected only if marked fish mix completely with unmarked fish between sampling events. This is unlikely with a more sedentary species like trout. If the species moved more (say northern pike), chances of complete mixing between sampling events is much improved.

Problem 4

The only way that large numbers of northern pike can be caught in George Lake (1,823 ha) is by seining the shallow spawning grounds during the three week period just after the lake opens in the spring. There are only four such spawning grounds that are shallow enough for a beach seine in George Lake; other spawning grounds are not seinable. If you plan to estimate abundance of spawning northern pike in George Lake with a mark-recapture experiment in which fish are marked with individually numbered tags:

1) What is the likelihood that every fish in the lake will have the same chance of being caught (good, fair, poor)?

   Poor

2) If marked northern pike mix COMPLETELY with unmarked fish between sampling events, what model do you expect to use to estimate abundance?

   Petersen's model

3) If marked northern pike mix PARTIALLY with unmarked fish between sampling events, what model do you expect to use to estimate abundance?

Darroch's model

4) If marked northern pike do NOT MIX at all with unmarked fish between sampling events, what model do you expect to use to estimate abundance?

Petersen's model for each sampling site to obtain a minimum estimate of abundance.

5) Of these three scenarios, which do you think is the most likely? If this is the expected scenario, what extra information must you tell your field crew to collect about each northern pike?

Partial mixing and Darroch's model. Field crews should record the location of capture for each fish each time its captured.

LAB NO. 3: PETERSEN'S MODEL (ANALYSIS)

Problem 1

During the third week in September *(1986)*, 683 burbot were released into Tolsona Lake. Burbot were captured in 150 baited hoop traps set on the bottom of the 120-ha, land-locked lake at locations randomly selected from parallel transects on an overlay. The overlay completely covered the mapped surface of the lake. Each captured burbot was measured to the nearest mm TL and was marked with an individually numbered tag and by removal of a fin. No burbot < 300 mm TL were caught. No burbot died during handling.

During the first week in October in the same year, 587 burbot were caught in the same lake with 75 baited hoop traps set on the bottom at locations again randomly selected from parallel transects on an overlay. Any repetition of locations of sets with those in the first sampling event was coincidental. Only one burbot was recaptured without a tag, no burbot was recaptured with all its fins, and 148 burbot were recaptured with a tag and a missing fin. No physical damage was apparent from any recaptured burbot.

Answer the following questions by analyzing the description of the sampling design above and with the data in the file NUMBER1.WK1.

1) What steps were taken to insure that burbot did not loose their marks between sampling events? Were they successful?

   Double marking. Yes

2) What evidence is available to indicate that captured burbot did not become "trap shy"? Did not become "trap happy"? How would these behaviors affect estimates of abundance?

   There is no evidence. Estimates would be too high with "trap shyness" or too low with "trap happiness."

3) What steps were taken to insure that little recruitment or mortality occurred between sampling events?

   The lake was landlocked and the sampling events were only two weeks apart.

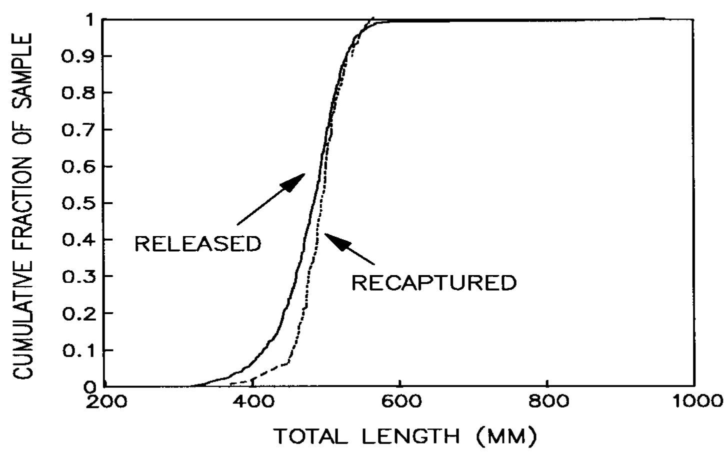
4) What arguments can you give to show that every fish had an equal probability of being captured during each sampling event regardless of its location in the lake?

   Sampling effort was most likely spread evenly across the lake by design.

5) What steps were taken to promote the mixing of marked burbot with unmarked burbot between sampling events?

   Sampling effort was most likely spread evenly across the lake which shortened the distance fish had to move to mix.

6) The plot below is a comparison of cumulative distribution functions of the lengths of burbot released with MARKS and of those subsequently RECAPTURED:



Did every burbot regardless of size have an equal chance of being caught? What is the risk you run of being wrong if you conclude the recapture rates are different? What's the hypothesis you're testing?

No. IvsIIvsIIIvsIV (see below) - $\chi^2 = 35.24$, $df = 3$ $P < 0.01$
H_o: $p_l = p_o$ *for all lengths l*

No. Kolmogorov-Smirnov Two-sample Test $D_{max} = 0.21, n = 683;149$, $P < 0.01$
H_o: $F(l_{marked}) = G(l_{recaptured})$

7) If your decision is that every burbot regardless of size DID have an equal probability of being sampled, estimate the abundance.

$$\hat{N} = \frac{(683+1)(587+1)}{(149+1)} - 1 = 2,681$$

8) If your decision is that every burbot DID NOT have equal probability of capture because of size-selectivity in sampling, divide the population into separate strata, and estimate the abundance for the population. For which event is sampling size-selective?

	Number Released	*Number Recaptured*	*p*	
>450 mm	*163*	*9*	*.06*	*Group I*
450-499 mm	*292*	*73*	*.25*	*Group II*
500-549 mm	*203*	*61*	*.30*	*Group III*
≥550 mm	*25*	*6*	*.24*	*Group IV*

IvsIIvsIIIvsIV - $\chi^2 = 35.24$, *df* = 3 $P < 0.01$
IIvsIIIvsIV - $\chi^2 = 1.67$, *df* = 2 $0.25 < P < 0.50$
Ivs(I+II+III+IV) - $\chi^2 = 33.33$, *df* = 1, $P < 0.01$

Divide data into two groups: < 450 mm and ≥ 450 mm TL

< 450 mm TL:

$$\hat{N} = \frac{(163+1)(113+1)}{(9+1)} - 1 = 1{,}869$$

Total Abundance = 3,624

≥ 450 mm TL:

$$\hat{N} = \frac{(520+1)(474+1)}{(140+1)} - 1 = 1{,}755$$

At least the second sampling event.

9) If the estimate of abundance from stratifying the data into two size groups is 35% higher than the estimate obtained without any stratification, which of these two estimates is the better estimate? Why?

The estimate calculated from the stratified data is the better. Hypothesis tests indicate there is bias from size-selective sampling in the estimate calculated without stratification. Bias from size-selective sampling is removed by stratifying the data into size groups, estimating abundance for each size group, and adding estimates across size groups. In this instance, the 35% difference in estimates is a measure of the bias in the estimate from unstratified data.

Problem 2

During the last week in May, 1989, 1,234 northern pike were released into George Lake. Northern pike were captured in beach seines along three widely separated beaches on the eastern, northwestern, and southwestern ends of the 1,823-ha lake. These beaches were spawning grounds and most of the fish caught were mature adults. Each captured northern pike was measured to the nearest mm FL and was marked with an individually numbered tag and by removal of a fin. No fish < 300 mm FL was caught. Since sampling was so concentrated, the area of the lake in which each northern pike was captured was noted. No northern pike died during handling.

Three days after completion of the first sampling event, 1,195 northern pike were caught in the same lake at the same sites with the same gear. Fifty-seven northern pike had been recaptured. Area of the lake in which each northern pike was captured was noted. No physical damage was apparent from any recaptured northern pike. Numbers of northern pike with and without marks captured during the second sampling event by area of the lake are:

	Number w/ Marks	Number w/o Marks	p
Area A	15	282	*.05*
Area B	15	336	*.04*
Area C	27	518	*.05*

Statistics on recovery during the second sampling event of marked northern pike released into different areas of the lake during the first sampling event are:

	Recaptured in Area A	Area B	Area C	Not Recaptured
Released in Area A	8	2	9	307
Released in Area B	3	9	2	332
Released in Area C	4	4	16	538

Answer the following questions by analyzing the description of the sampling design above and with the data in the file NUMBER2.WK1.

1) What steps were taken to insure that northern pike did not loose their marks between sampling events? Were they successful?

   Double marking. Yes. Fins can not grow back in three days.

2) What evidence is available to indicate that captured northern pike did not become "trap shy"? Did not become "trap happy"? How would these behaviors affect estimates of abundance?

   "Trap shyness" and "trap happiness" are unlikely behaviors when using beach seines. Estimates would be too high with "trap shyness" or too low with "trap happiness."

3) What steps were taken to insure that little recruitment or mortality occurred between sampling events?

Northern pike were spawning during sampling, and the sampling events were only three days apart.

4) What arguments can you give to show that every fish had an equal probability of being captured during each sampling event regardless of its location in the lake OR that marked fish mixed completely with unmarked fish across the lake? What is the risk you run of being wrong if you conclude the recapture rates are different? What's the hypothesis you're testing?

No arguments are available to show that northern pike at various locations had an equal probability of being captured during the SECOND sampling event.

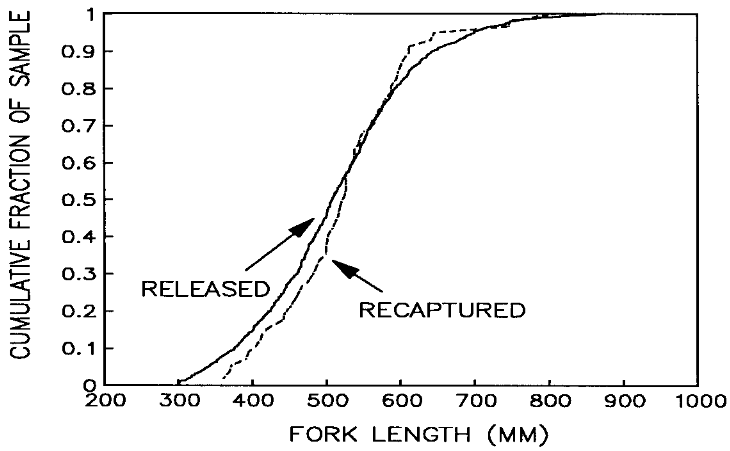
Isolation and concentration of sampling effort is an argument that fish across the lake WOULD NOT have an equal probability of being captured during EITHER sampling event.

Test of the hypothesis H_o: $p_a = p_b = p_c$ *was not rejected* ($\chi^2 = 0.28$, $df = 2$ $0.98 < P < 0.95$ *which indicates that either every fish had an equal probability of being marked during the FIRST sampling event or northern pike mixed completely across the lake BETWEEN sampling events.*

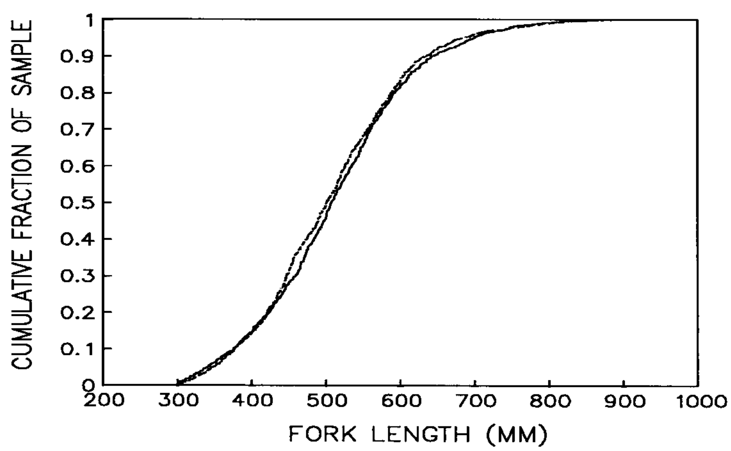
5) What argument do you have that an estimate of abundance from the data above is for all northern pike in George Lake and not just for northern pike in the vicinity of the three beaches?

Inspection of the second table in the description above showed that at least some mixing occurred between sampling events indicating that northern pike sampled around these beaches were not separate, isolated populations.

6) The plot below is a comparison of cumulative distribution functions of the lengths of northern pike released with MARKS and of those subsequently RECAPTURED:



The plot below is a comparison of cumulative distribution functions of the lengths of northern pike released with MARKS and of those CAPTURED several days later:



Did every northern pike regardless of size have an equal chance of being caught during at least one sampling event? What is the risk you run of being wrong if you conclude the recapture rates are different? What are the hypotheses you are testing?

Yes.
In comparing length distributions of marked and recaptured fish with the Kolmogorov-Smirnov Two-sample Test $D_{max} = 0.12$; $n = 1234,57$; $P = 0.41$
H_o: $F(l_{marked}) = G(l_{recaptured})$

In comparing length distributions of marked and recaptured fish with the χ^2 *test:*

	Number Not Recaptured	Number Recaptured	p	
300-399 mm	163	5	0.03	Group I
400-449 mm	191	6	0.03	Group II
450-499 mm	210	10	0.05	Group III
500-549 mm	206	17	0.08	Group IV
550-599 mm	172	10	0.05	Group V
≥600 mm	196	9	0.04	Group VI

IvsIIvsIIIvsIVvsVvsVI - $\chi^2 = 6.78$, $df = 5$, $0.10 < P < 0.25$
H_o: $p_I = p_{II} = p_{III} = p_{IV} = p_V = p_{VI}$

In comparing length distributions of marked and captured fish with the Kolmogorov-Smirnov Two-sample Test $D_{max} = 0.06$; $n = 1234,1195$; $P = 0.03$
H_o: $F(l_{marked}) = G(l_{captured})$

7) What is the abundance of northern pike (≥300 mm FL in George Lake in 1988?

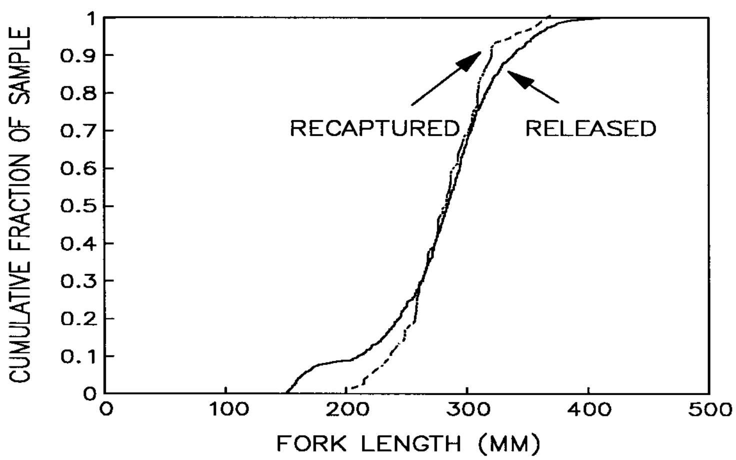
$$\hat{N} = \frac{(1,234+1)(1,195+1)}{(57+1)} - 1 = 25,466$$

LAB NO. 4: ADJUSTMENTS IN ESTIMATES OF COMPOSITION

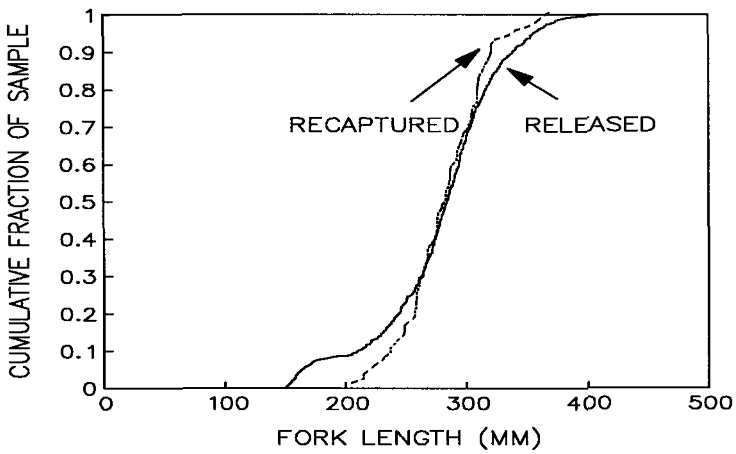
Problem 1

During the third week in July, 1989, 1,255 Arctic grayling were captured in the Upper Chena River with electrofishing gear, marked with an individually numbered tag and by removal of their adipose fin, and released into the river. A scale sample was taken from each fish, and its age was determined later. Sampling gear was moved along both banks of the entire river one mile at a time. No fish smaller than 149 mm FL was included in the experiment. During the first week in August, 952 Arctic grayling were caught in the Upper Chena River, 84 of which had been marked in July. Sampling methods were those that were used in July except no scales were taken during the second event. Twenty fish were inadvertently killed during handling in both sampling events.

The plot below is a comparison of cumulative distribution functions of Arctic grayling released with MARKS and of those subsequently RECAPTURED:



The plot below is a comparison of cumulative distribution functions of Arctic grayling released with MARKS and of those CAPTURED during the second sampling event:



Data for these comparisons and on the age of sampled fish can be found in file AGE1.WK1.

Since sampling to determine age composition of Arctic grayling was the same as that to estimate abundance in a mark-recapture experiment, the experiment can be used to determine if the samples taken to determine age composition are representative of the population, and if not, how to correct them.

1) Age composition determined from the scales taken in this study are relevant to which sampling event?

   The first event.

2) Was sampling size-selective during the SECOND sampling event? What is the risk you run of being wrong if you conclude that sampling was size-selective? What are the hypotheses you're testing?

   Yes.

   Comparing cumulative distribution functions of the lengths of marked and of recaptured fish with the Kolmogorov-Smirnov Two-sample Test D_{max} = *0.13; n = 1263,84; P = 0.11 which shows (along with the plot) that sampling was size-selective during the SECOND sampling event.*
   H_o: $F(l_{marked}) = G(l_{recaptured})$

   Comparing length distributions of marked and recaptured fish with the χ^2 *test also shows that sampling was size-selective during the SECOND sampling event:*

	Number Not Recaptured	*Number Recaptured*	*p*	
150-249 mm	*322*	*11*	*0.03*	*Group I*
250-299 mm	*512*	*45*	*0.08*	*Group II*
300-349 mm	*290*	*24*	*0.08*	*Group III*
≥350 mm	*55*	*4*	*0.07*	*Group IV*

IvsIIvsIIIvsIV - $\chi^2 = 8.34$, $df = 3$, $0.025 < P < 0.05$
H_o: $p_I = p_{II} = p_{III} = p_{IV}$

IIvsIIIvsIV - $\chi^2 = 0.15$, $df = 1$, $0.50 < P < 0.75$
H_o: $p_{II} = p_{III} = p_{IV}$

3) Was sampling size-selective during the FIRST sampling event? What is the risk you run of being wrong if you conclude that sampling was size-selective? What are the hypotheses you're testing?

Yes.
Comparing cumulative distribution functions of MARKED and RECAPTURED fish with the Kolmogorov-Smirnov Two-sample Test $D_{max} = 0.06$; $n = 1234,951$; $P = 0.04$ *showed that fish during the first sampling event were captured at different rates by size as were those caught in the second event, however, the plot indicates that this difference is functionally negligible.*
H_o: $F(l_{marked}) = G(l_{captured})$

4) If a) sampling was size-selective during the second sampling event, b) size distributions of fish caught during the first and second events functionally (meaningfully) different, and c) data to determine age are collected only during the first event, could the data on age composition be corrected for size-selective sampling? If not, how could you sample in the future to avoid this problem?

No. Take scales during the second sampling event.

5) If you were to stratify the population into groups based on length to remove the effects of size-selective sampling, how many groups would be the most efficient and where would the splits be?

Two. Near 250 mm FL.

6) Based on the analysis of size-selectivity of sampling, you divide the population into two groups: fish 150-259 mm FL and fish ≥ 260 mm FL. The statistics for these two subpopulations are:

	Marked	Captured	Recaptured
150-259 mm FL	395	266	17
≥ 260 mm FL	860	686	67

What model should be used to estimate abundance in this situation (Chapman's or Bailey's)? Why? What's the estimated abundance for each group?

Bailey. Sampling does not change the marked to unmarked ratio during the second event because sampling progresses through different fish as it moves along the river.

150-259 mm FL

$$\hat{N} = \frac{395\ (266+1)}{(17+1)} = 5{,}859$$

≥ 260 mm FL

$$\hat{N} = \frac{860\ (686+1)}{(67+1)} = 8{,}688$$

7) What's the *fraction* of each group comprised of two-year olds? Three-year olds? Fours? Fives? Six on up to ten-year olds?

	Fractions	
Age	*Small Fish*	*Large Fish*
2	*0.26*	*0.00*
3	*0.46*	*0.01*
4	*0.19*	*0.10*
5	*0.06*	*0.21*
6	*0.04*	*0.40*
7	*0.00*	*0.13*
8	*0.00*	*0.08*
9	*0.00*	*0.07*
10	*0.00*	*0.01*

8) What's the estimated *number* of Arctic grayling in each group comprised of two-year olds? Three-year olds? Fours? Fives? Six on up to ten-year olds?

	Abundance	
Age	*Small Fish*	*Large Fish*
2	*1,500*	*0*
3	*2,669*	*47*
4	*1,090*	*888*
5	*347*	*1,787*
6	*237*	*3,492*
7	*16*	*1,113*
8	*0*	*675*
9	*0*	*627*
10	*0*	*59*

9) What's the estimated *number* of Arctic grayling *in the entire population* comprised of two-year olds? Three-year olds? Fours? Fives? Six on up to ten-year olds? What's the fraction by age group *in the entire population*?

Age	*Abundance*	*Fraction*
2	1,500	0.103
3	2,716	0.187
4	1,977	0.136
5	2,135	0.147
6	3,729	0.256
7	1,128	0.078
8	675	0.046
9	627	0.043
10	59	0.004

10) What's the SE for the estimated abundance of five-year old Arctic grayling in the entire population?

$p_{s,5} = 22/371 = 0.06 \qquad p_{b,5} = 151/734 = 0.21$

where b denotes fish ≥ 260 mm FL
s denotes fish 150-259 mm FL

$$V[p_{s,5}] = \frac{0.06(1-0.06)}{371 - 1} = 0.000152$$

$$V[p_{b,5}] = \frac{0.21(1-0.21)}{734 - 1} = 0.000226$$

$$V[N_s] = \frac{5{,}859\ (395)(266-17)}{(17+1)(17+2)} = 1{,}684{,}976$$

$$V[N_b] = \frac{8{,}688\ (860)(686-67)}{(67+1)(67+2)} = 985{,}714$$

$$V[N_5] = (0.000152)(5{,}859)^2 + (1{,}684{,}976)(0.06)^2 - (1{,}684{,}976)(0.000152)$$
$$+ (0.000226)(8{,}688)^2 + (985{,}714)(0.21)^2 - (985{,}714)(0.000226)$$
$$= 71{,}334$$

$SE[N_5] = 267$

Problem 2

According to the latest data from the Statewide Harvest Survey (for year 1993), estimated harvest of northern pike in the sport fishery at Newmar Lake was 4,437. Maximum sustained yield from populations of northern pike in the vicinity of Newmar Lake has been estimated from other projects to be about 16% of abundance annually. In 1993, you visited the lake in early June, captured 833 northern pike in beach seines and released 798 with marks. A month later, you returned to the lake and caught four fish in beach seines and 15 in gill nets. None of these fish survived to be released. In spring, 1994 you returned to the lake with beach seines and caught 630 northern pike in beach seines; 33 of the fish captured had marks. All fish were measured to the nearest mm FL, and a sample of scales was taken from each fish.

1) Can an unbiased estimate of abundance of northern pike be obtained? If so, how? (HINT: Scales).

   Yes. By use of information on age composition.

2) If an unbiased estimate of abundance can not be obtained (the scale samples were lost in a bar), what is the biased estimate?

$$\hat{N} = \frac{(798+1)(630+1)}{(33+1)} - 1 = 14{,}828$$

3) Should the harvest from the population in Newmar Lake be restricted?

   Yes. $(14{,}828)(0.16) = 2{,}372 < 4{,}437$

LAB NO. 5: PARTIALLY STRATIFIED EXPERIMENTS (DARROCH'S METHOD)

Problem 1

During the first week in June, 1987, 2,061 northern pike were released into George Lake. Northern pike were captured in beach seines along widely separated beaches on the eastern and northwestern ends of the 1,823-ha lake. These beaches are spawning grounds, and most of the fish caught were mature adults. Each captured northern pike was measured to the nearest mm FL and was marked with an individually numbered tag and by removal of a fin. No fish < 300 mm FL were caught. Since sampling was so concentrated, the area of the lake in which each northern pike was captured was noted. Of the 2,199 northern pike caught during this sampling event, 117 died during handling.

Two weeks after completion of the first sampling event, 744 northern pike were caught in the same lake at the same sites with the same gear. Eighty-three northern pike were recaptured. Area of the lake in which each northern pike was captured was noted. Eight northern pike died during handling. Numbers of northern pike with and without marks captured during the second sampling event by area of the lake are:

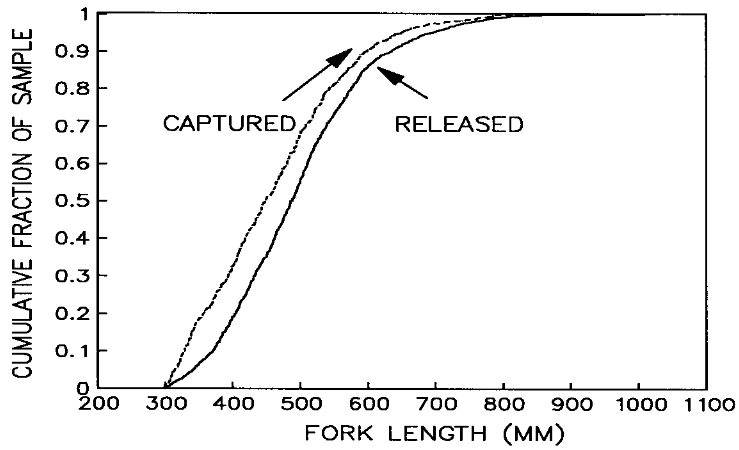
	Number w/ Marks	Number w/o Marks	*p*
Northwestern End	17	85	*.17*
Eastern End	66	575	*.10*

IvsII - $\chi^2 = 3.60$, $df = 1$, $0.05 < P < 0.10$

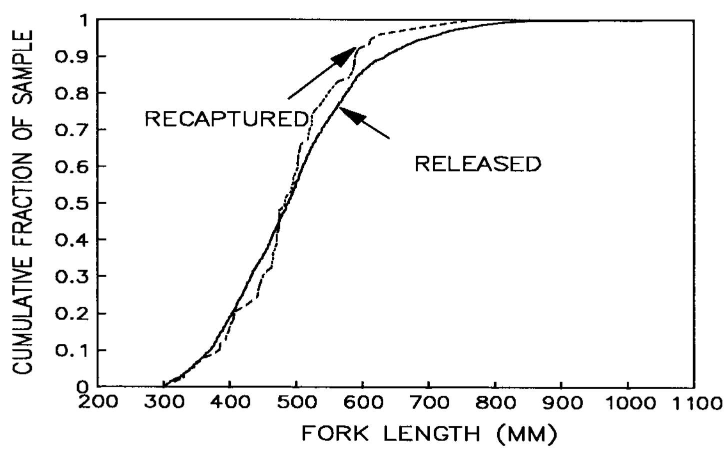
Statistics on recovery during the second sampling event of marked northern pike released into different areas of the lake during the first sampling event are:

	Recaptured in Northwestern End	Recaptured in Eastern End	Not Recaptured
Released Northwestern End	14	19	1,018
Released Eastern End	2	43	965

Plots of the cumulative distribution functions of lengths of northern pike captured during both sampling events are as follows (the dashed line corresponds to northern pike caught during the second sampling event):



Plots of the cumulative distribution functions of lengths of northern pike captured and released alive during the first sampling event and the lengths of northern pike recaptured during the second sampling event are as follows (the dashed line corresponds to marked northern pike recaptured during the second sampling event):



Answer the following questions by analyzing the description of the sampling design and the data listed above and by analyzing in the file DARROCH1.WK1.

1) What steps were taken to insure that northern pike did not loose their marks between sampling events?
Double marking.

2) What evidence is available to indicate that captured northern pike did not become "trap shy"? Did not become "trap happy"?

"Trap shyness" and "trap happiness" are unlikely behaviors when using beach seines.

3) What steps were taken to insure that little recruitment or mortality occurred between sampling events?

Northern pike were spawning during sampling, and the sampling events were only two weeks apart.

4) Did every fish have an equal probability of being caught, marked, and released alive during the FIRST SAMPLING EVENT regardless of where it was LOCATED in the lake? What is the risk you run of being wrong if you say NO? What's the hypothesis you're testing?

No. H_o: $p_{\mathrm{Eastern}} = p_{\mathrm{Northwestern}}$

5) Did every fish have an equal probability of being caught during the SECOND SAMPLING EVENT regardless of where it was LOCATED in the lake? What is the risk you run of being wrong if you say NO? What's the hypothesis you're testing? Can you even test this hypothesis? Is this test possible?

No test of this hypothesis is possible.

6) Did every fish have an equal probability of being caught during the SECOND SAMPLING EVENT regardless of its SIZE? What is the risk you run of being wrong if you say NO? What's the hypothesis you're testing?

Yes.

Kolmogorov-Smirnov Two-sample Test $D_{max} = 0.10$; $n = 2,062,77$; $P = 0.45$ *based on length distribution of fish captured in the first event against the distribution of those recaptured during the second event (Case II).*
H_o: $F(l_{\mathrm{marked}}) = G(l_{\mathrm{recaptured}})$

	Number Not Recaptured	*Number Recaptured*	*p*	
300-349 mm	*127*	*6*	*.045*	*Group I*
350-399 mm	*249*	*9*	*.035*	*Group II*
400-449 mm	*329*	*12*	*.035*	*Group III*
450-499 mm	*361*	*19*	*.050*	*Group IV*
500-549 mm	*362*	*19*	*.050*	*Group V*
550-599 mm	*251*	*12*	*.046*	*Group VI*
≥600 mm	*299*	*6*	*.020*	*Group VII*

IvsIIvsIIIvsIVvsVvsVIvsVII - $\chi^2 = 5.88$, $df = 6$, $0.25 < P < 0.50$
H_o: $p_{I} = p_{II} = p_{III} = p_{IV} = p_{V} = p_{VI} = p_{VII}$

7) Was mixing of marked fish with unmarked fish between sampling events COMPLETE? PARTIAL? Or did these fish mix at all? Why do you say so?

Mixing was partial because the off diagonal elements in the second table above are not zero indicating that some mixing did occur. The rejection of the hypothesis that the fraction of marked northern pike was the same in both areas of the lake indicates that mixing was not complete.

8) What model do you suggest be used to estimate abundance? On what size groups of northern pike?

Darroch's model. All northern pike $\geq$ 300 mm FL.

9) If you're answer is other than "Darroch's model", you're sadly mistaken. At this juncture, some bootstrapping would be in your future, however, the machines in the Training Room are too slow to realistically do the resampling. Therefore, we've don the simulations for you, and the results are in the file DARROCH2.WK1.

What is the estimated abundance of northern pike in George Lake in 1987? Its standard error? What is the potential bias in the estimate? What fractions of the bootstrap samples have unrealistic probabilities of capture? What's the estimate of abundance with Petersen's model? What it's standard error. Why are the abundance estimates from the two models so close?

$\hat{N} = 18{,}253$ $\quad SE[\hat{N}_{BOOT}] = 2{,}405$ $\quad BIAS = 18{,}253 - 18{,}753 = 500$

None.

$$\hat{N} = \frac{(2061+1)(744+1)}{(83+1)} - 1 = 18{,}287$$

$$V[N] = \frac{18{,}287(2061-83)(744-83)}{(83+1)(83+2)} = 3{,}348{,}667$$

$SE[N] = 1{,}830$

Although unequal probabilities of capture and incomplete mixing of marked and unmarked fish across the lake imparted some bias into the estimate obtained with Petersen's model, probabilities were enough alike and the extent of mixing was broad enough to render that bias negligible.

LITERATURE CITED

Bernard, D. R., G. A. Pearse, and R. H. Conrad. 1991. Hoop traps as a means to capture burbot. North American Journal of Fisheries Management 11:91-104.

Darroch, J. N. 1961. The two-sample capture-recapture census when tagging and sampling are stratified. Biometrika 48:241-60.

Efron, B. 1982. The jackknife, the bootstrap, and other resampling plans. Society Industrial Applied Mathematics Publication Number 38, Philadelphia.

Goodman, L. G. 1960. On the exact variance of a product. Journal of the American Statistical Association 66:708-713.

Jolly, G. M. 1965. Explicit estimates from capture-recapture data with both death and immigration - stochastic model. Biometrika 52:225-247.

Pollock, K. H. 1982. A capture-recapture design robust to unequal probability of capture. Journal of Wildlife Management 46:752-757.

Ricker, W. E. 1975. Computation and interpretation of biological statistics of fish populations. Fisheries Research Board of Canada, Bulletin Number 191.

Robson, D. S. and H. A. Regier. 1964. Sample size in Petersen mark-recapture experiments. Transactions of the American Fisheries Society 93:215-216.

Robson, D. S., and W. A. Flick. 1965. A nonparametric statistical method for culling recruits from a mark-recapture experiment. Biometrics 21:936-947.

Seber, G. A. F. 1965. A note on the multiple-recapture census. Biometrika 52:249-259.

Seber, G. A. F. 1982. The estimation of animal abundance and related-parameters, second edition. Charles Griffin and Company, Limited. London.