

January 25, 1996
Gary C. White
Department of Fishery and Wildlife
Colorado State University
Fort Collins, CO 80523
(970)491-6678
gwhite@cnr.colostate.edu

RH: NOREMARK Software Reference Manual ! White

Program NOREMARK Software Reference Manual

Gary C. White, Department of Fishery and Wildlife, Colorado State University, Fort Collins, CO 80523

Key words: closure, Lincoln-Petersen, mark-resight, maximum likelihood, population estimation, sighting surveys.

Scientific and Technical Information

Estimation of population size of a geographically and demographically closed but free-ranging population is a common problem encountered by wildlife biologists. The earliest approaches to this problem were developed by Petersen (1896) and later Lincoln (1930), where capture-recapture techniques were applied. Extensions to the simple 2 occasion Lincoln-Petersen estimator were developed for multiple occasions (Schnabel 1938, Darroch 1958), for removal experiments (Zippin 1956, 1958), and for heterogeneity of individual animals (Burnham and Overton 1978, 1979, Chao 1988). For the capture-recapture technique, Otis et al. (1978) and White et al. (1982) provided a summary of the available methods, and others (White et al. 1978, Rexstad and Burnham 1991) describe Program CAPTURE for computing these estimators of population size.

More technologically advanced approaches to the problem of abundance estimation have incorporated animals marked with radio transmitters. The initial sample of animals is captured and marked with radios, but recaptures of these animals are obtained by only observing them, not actually recapturing them. The limitation of this procedure is that unmarked animals are not marked on subsequent occasions. The advantage of this procedure is that resighting occasions are generally much cheaper to acquire than when the animals must be physically captured and handled. The mark-resight procedure has been tested with known populations of mule deer (Bartmann et al. 1987), and used with white-tailed deer (Rice and Harder 1977), mountain sheep (Furrow et al. 1981, Neal et al. 1993), black and grizzly bears (Miller et al. 1987), and coyotes (Hein 1992). Eberhardt (1990) further investigated the Petersen estimator with the Chapman correction (Chapman 1951) for small population sizes where animals could immigrate/emigrate from the study area.

Here, I present a user's manual for Program NOREMARK, a program to compute mark-resight estimators of population abundance. Four estimators of abundance are provided: the joint hypergeometric maximum likelihood estimator (JHE) (Bartmann et al. 1987, White and Garrott

1990, Neal et al. 1993), the joint hypergeometric maximum likelihood estimator extended to incorporate animals moving on and off the study area (Neal et al. 1993, White 1993), the Minta-Mangel estimator (Minta and Mangel 1989), and Bowden's estimator (Bowden 1993). The latter two estimators do not require the assumption that each animal in the population has the same probability of resighting on a particular occasion, as the first 2 estimators require.

Notation

- T_i Number of marked (telemetered) animals in the population at the time of the i^{th} survey, $i=1,\dots,k$. When the number of marked animals is assumed constant across surveys, the value is denoted as T .
- M_i Number of marked animals in the population that are on the area surveyed at the time of the i^{th} sighting survey. For all M_i constant, define $M \equiv M_i$.
- n_i Number of animals seen during the i^{th} sighting survey, consisting of m_i marked animals and u_i unmarked animals, so that $n_i = m_i + u_i$.
- f_i Number of times marked animal i was observed during the k surveys (sighting frequencies), $i = 1, \dots, T$. . Note that this is not the same use of f_j as in Otis et al. (1978) or White et al. (1982).
- m . Total number of sightings of marked animals, so that $m = \sum m_i = \sum f_i$.
- u . Total number of sightings of unmarked animals, so that $u = \sum u_i$, where $i = 1, \dots, N - T$.
- \bar{f} Mean capture frequency of marked animals, m/T .
- s_f^2 Variance of the sighting frequencies of the marked animals, $s_f^2 = \frac{\sum_{i=1}^T (f_i - \bar{f})^2}{T}$.

Estimators

Four estimators of population size for marking and sighting surveys are provided in Program NOREMARK. First is the joint hypergeometric maximum likelihood estimator (JHE) (Bartmann et al. 1987, White and Garrott 1990, Neal 1990, Neal et al. 1993). JHE is the value of N which maximizes the following likelihood (\hat{N}):

$$\mathcal{L}(N \mid M, n_i, m_i) = \prod_{i=1}^k \frac{\binom{M}{m_i} \binom{N-M}{n_i-m_i}}{\binom{N}{n_i}} \quad (1)$$

and the terms are defined for all $i = 1$ to k sighting occasions. The estimate \hat{N} can be found by iterative numerical methods. Confidence intervals are determined with the profile likelihood method (Hudson 1971, Venzon and Moolgavkar 1988). This estimator assumes that all the marked animals are on the area surveyed for each survey, i.e., that the population is geographically and demographically closed, and thus N is the same for each survey. The number of marked animals (M) is the same for each survey in the above equation, although the probability of sighting animals is not assumed to be the same for each survey. An extension allowed in Program NOREMARK is to allow additional marked animals to be added to the population between sighting occasions. Thus, M_i replaces M in the above equation, but the value of N is still assumed constant across occasions.

Second, the JHE has been extended to accommodate immigration and emigration (Neal et al. 1993) through a binomial process. This estimator is labeled IEJHE, and does not assume that the population is geographically closed (but the population is still assumed to be demographically closed). Assume that the total population with any chance of being observed on the study area is N^* , and that at the time of the i^{th} sighting survey, N_i animals occur on the study area. I am interested in estimating the mean number of animals on the study area, \bar{N} , and possibly N^* . At the time of the i^{th} sighting occasion, a known number of the marked animals (M_i) are on the study area of the possible T_i animals with transmitters. The probability that an individual is on the study area on the i^{th} occasion can be estimated as M_i/T_i , or in terms of the parameters of interest as N_i/N^* . Then the likelihood function for the model that includes temporary immigration and emigration from the study area is a product of the binomial distribution for the probability that a marked animal is on the study area times the joint hypergeometric likelihood of Eq. (1):

$$\mathcal{L}(N^*, N_i \mid T_i, M_i, m_i, n_i) = \prod_{i=1}^k \binom{T_i}{M_i} \left(\frac{N_i}{N^*}\right)^{M_i} \left(1 - \frac{N_i}{N^*}\right)^{T_i - M_i} \frac{\binom{M_i}{m_i} \binom{N_i - M_i}{n_i - m_i}}{\binom{N_i}{n_i}} \quad (2)$$

The parameters N^* and N_i for $i=1$ to k can be estimated by numerical iteration to maximize this likelihood, with the constraints that $N_i > (M_i + u_i)$ and $N^* > N_i$ for $i=1$ to k . Profile confidence intervals can be obtained for the $k+1$ parameters. I was not interested in the k population estimates for each sighting occasion, but rather desired the mean of the N_i estimates. Therefore, I re-parameterized the likelihood to estimate the total population and mean population size on the

study area directly, and their profile likelihood confidence intervals. In the re-parameterized likelihood, I used $N_i = \bar{N} + \alpha_i$, where $\sum \alpha_i = 0$.

Third, Minta and Mangel (1989) suggested a bootstrap estimator (MM) of population size based on the sighting frequencies of the marked animals, f_i . For unmarked animals, sighting frequencies are drawn at random from the observed sighting frequencies of the marked animals until the total number of captures equals u . The number of animals sampled is then an estimate of the number of unmarked animals in the population, so that M plus the number sampled is an estimator for N . Only bootstrap samples where the number of sightings was exactly equal to u were used, i.e., cases where the cumulative sightings exceeded u were rejected. Minta and Mangel (1989) accepted the first value where the cumulative sightings equalled or exceeded u . The stopping rule I used results in less bias than the rule used by Minta and Mangel (1989). Minta and Mangel (1989) suggested the mode of the bootstrap replicates as the population estimate. Confidence intervals were computed as probability intervals with the 2.5th and 97.5th percentiles from the bootstrapped sample of estimates. White (1993) demonstrated that the MM estimator is basically unbiased, but that the confidence interval coverage was not at the expected 95% for $\alpha=0.05$. He suggested a modified procedure, but coverage still was not satisfactory.

Fourth, Bowden (1993) suggested an estimator for the Minta-Mangel model where the confidence intervals on the estimate were computed based on the variance of the resighting frequencies of the marked animals. He approached the problem from a sampling framework, where each animal in the population has the attribute f_i of the number of times it was resighted. The values of f_i are known for the marked animals, and the sum of the f_i 's ($= u$) are known for the unmarked animals. Then, an unbiased estimator of the population size is

$$\hat{N} = \frac{\left(\frac{(u. + m.)}{\bar{f}} + \frac{s_f^2}{\bar{f}^2} \right)}{\left(1 + \frac{s_f^2}{T\bar{f}^2} \right)}$$

with variance

$$\text{Vâr}(\hat{N}) = \hat{N}^2 \frac{\left(\frac{1}{T} - \frac{1}{\hat{N}} \right) \frac{s_f^2}{\bar{f}^2}}{\left(1 + \frac{s_f^2}{T\bar{f}^2} \right)^2} .$$

Confidence intervals for N are computed from a log-transformation as

$$\hat{N} / \exp\left(t_{1 - \frac{\alpha}{2}, T-1} \hat{C}\hat{V}(\hat{N}) \right) \quad \text{and} \quad \hat{N} \times \exp\left(t_{1 - \frac{\alpha}{2}, T-1} \hat{C}\hat{V}(\hat{N}) \right)$$

the backspace key erases the previous character, and the delete key deletes the current character. The Insert key can be used to insert text into previously entered text. Once the title is correct, you move the cursor to the next data entry field by hitting the Enter key, which in this case moves you to the alpha level field. The default value is 0.05, which you probably do not want to change. However, should you desire a 90% confidence interval instead of a 95% interval, enter the value 0.10 in this field. After the correct value is entered, move to the number of sighting occasions field with the Enter key.

In this field, you enter the number of sighting occasions. For the JHE estimator, a sighting occasion is an attempt to view animals in the population, keeping track of the number of marked and unmarked animals observed. For the example here, I'll assume that 14 sighting occasions were conducted. After the value 14 has been entered, the user could proceed to the next data entry screen by hitting the Enter key. However, you might have noted a mistake in one of the previous fields. This mistake could be corrected by using the up arrow to move the cursor to the field with the error, and making the change. To exit the data screen, you could hit the Enter key until you reach the last field, or you can use the short cut key PgDn, which takes you immediately to the next data entry screen. The following shows the screen just prior to hitting return to specify 14 resighting occasions for the data set described in Neal et al. (1993). These data are not the best example to illustrate this estimator with because the population was not closed, as is assumed by this estimator. However, I believe it is informative to see the difference in the results from this estimator and the immigration-emigration estimator, which is appropriate for these data.

The user has specified a title to identify the data. In addition, an α level for computing confidence intervals has been specified as 0.05, giving a 95% confidence interval. Finally, 14 recapture occasions have been specified.

Note that additional information is available on any of these input requests by hitting the F1 (Help) key. A description of the input being requested will be popped on the screen for the user to read. To return to entering data, just hit Esc to exit the help screen.

the same sighting probability on a particular occasion and is sampled without replacement. For both these estimators, it is inappropriate to apply them to data collected by sampling with replacement, i.e., animals are seen more than once during a particular occasion.

Bowden's estimator allows each animal's sighting probability to differ from the others, and sampling can be with or without replacement. Sighting heterogeneity allows to a certain degree that the study area does not have to be geographically closed, in that some animals can be off the study area for a particular occasion(s), and hence have a zero sighting probability. The resulting estimate is the total population using the study area, which is not the same as the average density of animals on the study area. Thus, the estimate resulting from Bowden's estimator may not be exactly what the researcher had in mind if the study area is not closed.

The Minta-Mangel estimator requires the same assumptions as Bowden's estimator, although this estimator is derived under the assumption of sampling with replacement. However, using this estimator for situations where sampling without replacement is not particularly inappropriate, because the estimator would not be particularly biased and confidence interval length would be only slightly larger than the estimator derived under the assumption of sampling without replacement. I do not recommend the Minta-Mangel estimator because of the poor performance of the confidence interval coverage, and suggest that users of NOREMARK use Bowden's estimator if heterogeneity of sighting probabilities is serious or sampling is performed with replacement.

If there is little or no heterogeneity of sighting probabilities, the joint hypergeometric estimator should generate slightly smaller confidence interval lengths than Bowden's estimator because stronger assumptions are made. Neal et al. (1993) showed that confidence interval coverage drops to 80% for an expected 95% interval when reasonable sighting heterogeneity is simulated. The choice of Bowden's estimator over the joint hypergeometric estimator will depend on the degree of heterogeneity in sighting probabilities, and whether sampling is performed with or without replacement.

One additional criterion is useful in deciding between the joint hypergeometric estimator and Bowden's estimator. Often, conduct of the survey does not fall into logical occasions, i.e., animals may be counted multiple times during one "occasion". Examples are photographing bears at bait sites with motion-sensing cameras, where the same bear may visit the same bait site multiple times in a few hours time. Besides the fact that this type of survey probably has heterogeneity of sighting probabilities, structure of the survey basically precludes use of the joint hypergeometric estimator. Such surveys should be considered as sampling with replacement.

Program Testing

User's can verify their copy of the program by duplicating results in the previous sections. Input files supplied with the program can be retrieved with the F2 function key, and the output verified against the output shown in the above screen reproductions. The input file ANDREA.HYP provides input for the JHE and immigration-emigration estimator. ANDREA.MM provides input for the Minta-Mangel and Bowden's estimators.

Verification of the code has been performed by checking the numerical optimization results against independent codes. However, the strongest evidence that the code is correct is

that simulation results for data simulated under the correct model generates correct results. That is, estimators are unbiased, and confidence interval coverage is the expected $1 - \alpha$ level for reasonable inputs.

Programming Logic

NOREMARK user interface is implemented using the CA-Clipper compiler. This language is a superset of the dBase III language and provides the capabilities to select choices from a list and input fields interactively.

The design capability of the program is implemented via a dBase III database named INTERP.DBF that the code creates if the database is not present. The context-sensitive help function is also implemented via a dBase III database (HELP.DBF) which includes a memo field existing in HELP.DBT. If these files are not present in the local directory, the program creates them.

Numerical optimization is performed with FORTRAN codes for the joint hypergeometric and the immigration-emigration estimators. For the joint hypergeometric estimator, the FORTRAN objects are linked directly into the NOREMARK.EXE file. For the immigration-emigration estimator, the NOREMARK.EXE code creates an input file (IE_NRM.INP) to the IE_NRM.EXE file. This file is then executed, reads IE_NRM.INP, computes the estimator, creates the output file IE_NRM.OUT, and exits. NOREMARK.EXE then reads the estimation results from IE_NRM.OUT.

A golden section search in one dimension (Press et al. 1986) is used to compute the maximum of the joint hypergeometric estimator. The maximum is bracketed with routine MNBRAK, then the golden section search initiated. NPSOL (Gill et al. 1986) is used to compute the optimum of the immigration-emigration estimator. NPSOL is a set of FORTRAN subroutines designed to minimize a smooth function subject to constraints, which may include simple bounds on the variables, linear constraints and smooth nonlinear constraints. NPSOL uses a sequential quadratic programming algorithm, in which the search direction is the solution of a quadratic programming subproblem. The algorithm treats bounds, linear constraints, and nonlinear constraints separately.

The NOREMARK code is not portable to other operating systems at this time because the CA-Clipper compiler is available only for DOS.

Program Files

The following is a list of files distributed with the program as an archive created with PKZIP. The archive can be unzipped with PKUNZIP.

Directory of C:\NOREMARK

README	1ST	262	01-05-94	10:29a
ANDREA	HYP	369	06-20-92	8:31a
ANDREA	MM	250	06-20-92	9:54a
ANDREA	OUT	3,692	04-27-94	4:07p
NOREMARK	EXE	611,328	02-25-95	9:30a

INTERP	DBF	29,186	06-20-92	8:14a
HELP	DBF	1,116	02-02-94	6:38a
HELP	DBT	13,989	02-02-94	6:38a
IE_NRM	EXE	350,720	04-07-94	6:49p
DOSXMSF	EXE	374,950	01-30-93	12:00a
BADGER	MM	169	06-22-92	10:01a
BISON76	MM	170	06-22-92	10:33a
BISON77	MM	160	06-22-92	12:47p
PORCUPIN	MM	109	06-22-92	1:41p
KUFELD	MM	250	11-15-93	8:43a
NOODLES	HYP	224	06-17-93	12:37p
NOODLES	MM	371	06-17-93	12:36p
USERMAN	WP	291,342	02-25-95	9:46a

The file ANDREA.OUT is output for checking the program using as input the files ANDREA.HYP (JHE and immigration-emigration estimators) and ANDREA.MM (Minta-Mangel and Bowden's estimators). The files NOREMARK.EXE and IE_NRM.EXE are executable files for the program. DOSXMSF.EXE is the DOS memory extender file for IE_NRM.EXE. INTERP.DBF is the dBase database for interpolations with the design procedure. HELP.DBF and HELP.DBT are the help documents in dBase III format. The next 7 files are example input from Minta and Mangel (1989), Bowden (1993), or just a test case based on a simple experiment with marked macaroni. The last file USERMAN.WP is this document in WordPerfect 6.0 format.

Literature Cited

- Arnason, A. N., Schwarz, C. J., and Gerrard, J. M. 1991. Estimating closed population size and number of marked animals from sighting data. *J. Wildl. Manage.* 55:716-730.
- Bartmann, R. M., G. C. White, L. H. Carpenter, and R. A. Garrott. 1987. Aerial mark-recapture estimates of confined mule deer in pinyon-juniper woodland. *J. Wildl. Manage.* 51:41-46.
- Bowden, D. C. 1993. A simple technique for estimating population size. Dept. of Statistics, Colorado State Univ., Fort Collins, Colo. 17pp.
- Buckland, S. T., and P. H. Garthwaite. 1990. Algorithm AS 259 -- Estimation confidence intervals by the Robbins-Monro search process. *Applied Statistics* 39:413-424.
- Burnham, K. P., and W. S. Overton. 1978. Estimation of the size of a closed population when capture probabilities vary among animals. *Biometrika* 65:625-633.
- Burnham, K. P., and W. S. Overton. 1979. Robust estimation of population size when capture probabilities vary among animals. *Ecology* 60:927-936.
- Chao, A. 1988. Estimating animal abundance with capture frequency data. *J. Wildl. Manage.* 52:295-300.
- Chapman, D. G. 1951. Some properties of the hypergeometric distribution with applications to zoological sample censuses. *University of California Publication in Statistics* 1:131-160.
- Darroch, J. N. 1958. The multiple recapture census: I. Estimation of a closed population. *Biometrika* 45:343-359.
- Eberhardt, L. L. 1990. Using radio-telemetry for mark-recapture studies with edge effects. *J. Applied Ecol.* 27:259-271.
- Furlow, R. C., M. Haderlie, and R. Van den Berge. 1981. Estimating a bighorn sheep population by mark-recapture. *Desert Bighorn Council Transactions* 1981:31-33.
- Gill, P. E., W. Murray, M. A. Saunders, and M. H. Wright. 1986. User's guide for NPSOL (Version 4.0): a FORTRAN package for nonlinear programming. Tech. Rep. SOL 86-2, Systems Optimization Laboratory, Dept. of Operations Research, Stanford Univ., Stanford, Calif. 54pp.
- Hein, E. W. 1992. Evaluations of coyote attractants and a density estimate on the Rocky Mountain Arsenal. M. S. Thesis, Colorado State Univ., Fort Collins. 58pp.

- Hudson, D. J. 1971. Interval estimation from the likelihood function. *J. Royal Stat. Soc. Series B* 33:256-262.
- Leslie, D. M., Jr. and C. L. Douglas. 1979. Desert bighorn sheep of the River Mountains, Nevada. *Wildl. Monogr.* 66:1-56.
- Leslie, D. M., Jr. and C. L. Douglas. 1986. Modeling demographics of bighorn sheep: current abilities and missing links. *North American Wildlife and Natural Resources Conference Transactions* 51:62-73.
- Miller, S. D., E. F. Becker, and W. H. Ballard. 1987. Black and brown bear density estimates using modified capture-recapture techniques in Alaska. *International Conference on Bear Research and Management* 7: 23-35.
- Minta, S. and M. Mangel. 1989. A simple population estimate based on simulation for capture-recapture and capture-resight data. *Ecology* 70:1738-1751.
- Neal, A. K. 1990. Evaluation of mark-resight population estimates using simulations and field data from mountain sheep. M. S. Thesis, Colorado State Univ., Fort Collins. 198pp.
- Neal, A. K., G. C. White, R. B. Gill, D. F. Reed, and J. H. Olterman. 1993. Evaluation of mark-resight model assumptions for estimating mountain sheep numbers. *J. Wildl. Manage.* 57:436-450.
- Otis, D. L., K. P. Burnham, G. C. White, and D. R. Anderson. 1978. Statistical inference from capture data on closed animal populations. *Wildl. Monogr.* 62:1-135.
- Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. 1986. *Numerical recipes: the art of scientific computing.* Cambridge Univ. Press, Cambridge, U.K. 818pp.
- Schnabel, Z. E. 1938. estimation of the size of animal populations by marking experiments. *U. S. Fish and Wildlife service Fisheries Bull.* 69:191-203.
- Rexstad, E., and K. Burnham. 1991. Users' guide for interactive program CAPTURE. Colo. Coop. Fish and Wildl. Res. Unit, Colo. State Univ., Fort Collins. 29pp.
- Rice, W. R. and J. D. Harder. 1977. Application of multiple aerial sampling to a mark-recapture census of white-tailed deer. *J. Wildl. Manage.* 41:197-206.
- Venzon, D. J. and Moolgavkar, S. H. 1988. A method for computing profile-likelihood based confidence intervals. *Appl. Stat.* 37:87-94.

- White, G. C. 1993. Evaluation of radio tagging marking and sighting estimators of population size using Monte Carlo simulations. Pages 91-103 in J.-D. Lebreton and P. M. North (eds.). *Marked Individuals in the Study of Bird Population*, Birkhäuser Verlag, Basel, Switzerland.
- White, G. C., K. P. Burnham, D. L. Otis and D. R. Anderson. 1978. User's manual for program CAPTURE. Utah State Univ. Press, Logan, UT. 40pp.
- White, G. C., D. R. Anderson, K. P. Burnham, and D. L. Otis. 1982. Capture-recapture and removal methods for sampling closed populations. Los Alamos National Laboratory. LA-8787-NERP. Los Alamos, N. M. 235pp.
- White, G. C. and R. A. Garrott. 1990. *Analysis of wildlife radio-tracking data*. Academic Press, New York, N. Y. 383pp.
- Zippin, C. 1956. An evaluation of the removal method of estimating animal populations. *Biometrics* 12:163-169.
- Zippin, C. 1958. The removal method of population estimation. *J. Wildl. Manage.* 22:82-90.