Southwest Fisheries Center Administrative Report H-80-2

ESTIMATING MONK SEAL POPULATIONS USING CHANGE-IN-RATIO AND LEAST SQUARES METHODS

Jerry A. Wetherall
Southwest Fisheries Center Honolulu Laboratory
National Marine Fisheries Service, NOAA
Honolulu, Hawaii 96812

DRAFT

February 1980

SUMMARY

This report describes some statistical methods for estimating the size of monk seal populations using observations of hauled out seals during the molting season. Two techniques are examined: (1) a change-in ratio (CIR) technique and (2) a least squares (LS) method. The first approach requires counts of the number of seals molting on each of a series of beach surveys and counts of pre-molt and post-molt seals on the first and last surveys of the series. The paper critically evaluates past efforts at applying the CIR method to monk seal populations and develops new estimators grounded in a stochastic model of molting dynamics.

The general theoretical model also gives rise to a set of LS estimators. They are based on a sequence of counts of molting seals. Both the CIR and LS methods rely on auxiliary information, namely haul out probabilities for different classes of seals and the probability distributions for time spent in the pre-molt state and in the molt state.

NOTATION

The following symbols and definitions will be adopted in describing the CIR and LS estimators and their variances. Throughout the paper, a hat (^) over a symbol signifies that the quantity is estimated. Some symbols of lesser importance are not listed here but are defined sufficiently in the text.

N = Total number of seals in the population, assumed constant during the molting season.

n = Number of beach surveys conducted. On each survey all hauled out seals are enumerated and classified as pre-molt, molting, or post-molt.

t_i = The instant in time at which the i^{th} survey is assumed to be conducted, measured from the beginning of the molting season, $i = 1, 2, \ldots, n$.

m_i = Number of molting seals counted on i^{th} survey, $i = 1, 2, \ldots, n$.

r_i = Number of pre-molt seals counted on the i^{th} survey, $i = 1, 2, \ldots, n$.

h_i = Total number of seals hauled out on the i^{th} survey, $i = 1, 2, \ldots, n$.

s_i = Number of molting seals in the population at time t_i.

p_i = Proportion of non-molting seals at time t_i which are in the pre-molt state, $i = 1, 2, \ldots, n$.

μ_i = Probability that a molting seal is hauled out at time t_i, $i = 1, 2, \ldots, n$.

b_i = Number of seals beginning molt in the time interval (t_{i-1}, t_i), $i = 2, 3, \ldots, n$.

c_i = Number of seals completing molt in the interval (t_{i-1}, t_i), $i = 2, 3, \ldots, n$.

ω_i = Number of seals which began molting on the i^{th} day of the molting season, in a sample of $\omega_{\bullet}$ seals.

π_i = Probability that a seal is molting at time t_i, $i = 1, 2, \ldots, n$.

ϕ_i = Probability that a seal begins molting in (t_{i-1}, t_i), $i = 2, 3, \ldots, n$.

L = Index of the last survey in a series of surveys contributing to a CIR estimate.

y = Amount of time a seal spends in the pre-molt state, measured from the beginning of the molting season (a random variable).

x = Amount of time a seal spends in the molting state (a random variable).

λ = Maximum value of y; i.e., maximum pre-molt time.

τ = Maximum value of x; i.e., maximum molt duration.

$\lambda + \tau$ = Length of molting season.

$f(x)$ = Probability density function for x.

$g(y)$ = Probability density function for y.

$F(x)$ = Cumulative distribution function for x.

f_i = Probability that a seal completes its molt exactly i days after beginning it, $i = 1, 2, \ldots, \tau$.

g_i = Probability that a seal begins molting on the i^{th} day of the molting season (at the beginning of the day). $i = 1, 2, \ldots, \lambda$.

$\sigma^2(\cdot)$ = Variance of a random variable or statistic.

$V(\cdot)$ = Estimator of $\sigma^2(\cdot)$.

w_i = Statistical weight.

ε_i = Random error term (residual) for the i^{th} data point in LS model.

e = Base of natural logarithm

CHANGE-IN-RATIO ESTIMATES

Eberhardt's Model

Basic CIR methodology is discussed thoroughly in Seber (1973). The approach allows one to estimate the size of a population by knowing the proportion of the population in a specified class at the beginning and end of a time interval and the number of members added or removed from that class and the other classes during the interval.

The application of CIR estimators to monk seals was suggested by Footnote 1 Eberhardt.[1] Assuming seals hauled out and observed during a survey were identifiable as pre-molt, molting or post-molt, he developed an estimator based on the proportions of non-molting seals in the post-molt class during two beach surveys and the number of seals observed in the molting class on all visits made in the intervening period . His approach assumes that visits are timed in such a way that all seals molting during the interval are observed and that none are counted more than once, or alternatively, that during the interval the multiple counts are exactly balanced by the number of molting seals not seen. He further assumes that all molting seals are hauled out at the time of surveys and that all non-molting seals haul out with equal probability when the first and last surveys are conducted. Under these conditions his estimator is

$$\hat{N}_E = \frac{\sum_{i=2}^{L} m_i + p_L(m_1 - m_L)}{p_1 - p_L} + m_1$$

where the symbols are defined in the previous section.

DeMaster's Model

Because such assumptions are likely to be violated in practice,

Footnote 2 DeMaster[2] suggested a procedure to adjust the counts of molting seals by dividing each count value by the expected number of times a molting seal will be seen during the surveys and by μ, the average haul out probability for molting seals. DeMaster calls the first of these constant divisors the "bias." In essence, his adjusted CIR estimator is

$$\hat{N}_D = \frac{\sum_{i=2}^{L}\left(\frac{m_i}{\text{bias} \cdot \mu}\right) + p_L\left(\frac{m_1 - m_L}{\mu}\right)}{p_1 - p_L} + \left(\frac{m_1}{\mu}\right) .$$

DeMaster computed the bias under the following restrictive set of assumptions:

(1) molt duration, x, ranges from 1 to 2I days, where I is a fixed interval between successive surveys,

(2) x is "normally distributed,"

(3) pre-molt time, y, is uniformly distributed [DeMaster made this assumption tacitly, but claims the results are not dependent on the distribution of y].

Under these conditions it is easy to show that his results reduce to bias = $E(x)/I$ where $E(\cdot)$ is the expectation operator. This result requires that y is uniform and, in his case, that surveys are conducted at a constant interval, I. The distribution and range of x do not matter. DeMaster seems unaware of these conditions, and neglects other difficulties as well. For example, the bias factor is correct only for counts made at times t_i such that $\tau \leq t_i < \lambda$, where λ is the maximum time in the pre-molt state and τ is the maximum molt duration (in DeMaster's case $\tau = 2I$). Hence, if the first (benchmark) survey is made at time t_1, and the k^{th} survey is the first survey such that $t_k > \tau$, then $m_2, m_3, \ldots, m_{k-1}$ should be excluded from the summation and the bias factor for m_k would be $E(x)/(t_k - t_1)$. Counts on subsequent surveys would be corrected by the usual bias factor, $E(x)/I$, subject to $t_i < \lambda$. For $\lambda \leq t_i < \lambda + \tau$ the bias term does not reduce to such simple expressions.

DeMaster suggested that an optimal survey design would result from choosing $I = E(x)$, but it is not clear why this would be so since, except for the difficulties mentioned here, it is a simple matter to compute and apply the bias factor. In any event, DeMaster's assumptions on the distribution of x actually result in bias = 1 regardless of I, so his particular model is inapplicable except in the moot case when $I = E(x)$.

A General Model

In most situations DeMaster's simple approach will not be satisfactory because y will not be uniformly distributed and it may not be possible or convenient to survey at regular intervals. What is needed is a more general procedure that will permit greater realism and more flexibility in survey design and analysis.

A general model can be constructed to accommodate any assumed distributions for x and y and any pattern of sampling. Let $f(x)$ be the probability density function for x and $g(y)$ be the density function for y. The molting season begins at time zero with all N seals in a pre-molt state and progresses until all seals have completed molting and entered the post-molt state. Assuming x and y are independent, the probability that a seal is molting at time t_i is

$$\pi_i = \int_0^{t_i} g(u) \Big(1 - F(t_i - u)\Big) du$$

where $F(t_i - u) = \int_0^{t_i - u} f(x)dx$.

The expected number of seals observed molting at the time of the i^{th} survey is just

$$E(m_i) = \mu_i \pi_i N \qquad (1)$$

where μ_i is the probability a molting seal is hauled out at t_i. Further, the expected number of seals beginning the molt between the previous survey at t_{i-1} and the present visit is

$$E(b_i) = N \int_{t_{i-1}}^{t_i} g(u)du = N\phi_i$$

and the expected number of seals completing the molt in (t_{i-1}, t_i) is

$$E(c_i) = \frac{E(m_{i-1})}{\mu_{i-1}} - \frac{E(m_i)}{\mu_i} + E(b_i) \qquad i = 2, 3, 4, \ldots, n$$

Given the sequence of counts of molting seals, m_1, m_2, m_3, ..., and the auxiliary information on $g(\cdot)$, $f(\cdot)$, and $\mu(\cdot)$, the number of seals beginning the molt and the number completing the molt in any survey interval can easily be estimated:

$$\hat{b}_i = m_i\left(\frac{E(b_i)}{E(m_i)}\right) \qquad E(m_i) > 0 \tag{2}$$

and

$$\hat{c}_i = \frac{m_{i-1}}{\hat{\mu}_{i-1}} - \frac{m_i}{\hat{\mu}_i} + \hat{b}_i . \tag{3}$$

A pair of CIR estimators can now be derived:

$$\hat{N} = \frac{\sum_{i=2}^{L} \hat{b}_i + \hat{p}_L\left(\frac{m_1}{\hat{\mu}_1} - \frac{m_L}{\hat{\mu}_L}\right)}{\hat{p}_1 - \hat{p}_L} + \frac{m_1}{\hat{\mu}_1} \tag{4}$$

and

$$\hat{N}' = \frac{\sum_{i=2}^{L} \hat{c}_i - (1 - \hat{p}_L)\left(\frac{m_1}{\hat{\mu}_1} - \frac{m_L}{\hat{\mu}_L}\right)}{\hat{p}_1 - \hat{p}_L} + \frac{m_1}{\hat{\mu}_1}$$

where p_i is the proportion of non-molting seals at time t_i which are pre-molts. Note that $\hat{N} = \hat{N}'$. In the special case where y is uniformly distributed,

$$\pi_i \sim E(x) \qquad \text{for} \qquad \tau \leq t_i < \lambda$$

and

$$\hat{b}_i = \frac{m_i}{\mu_i}\left((t_i - t_{i-1})/E(x)\right) .$$

This removes the equal interval restriction of the DeMaster model, but not its other shortcomings.

Another special case of the CIR arises when the n surveys span the entire molting season, i.e., the first survey is at time zero and the n^{th} (L^{th}) survey is at time $t_n \geq \lambda + \tau$. In this event

$$\sum_{i=2}^{n} E(b_i) = \sum_{i=2}^{n} E(c_i) = N .$$

Further, $p_1 = 1$, $p_L = 0$ and the CIR estimator at (4) reduces to

$$\hat{N} = \sum_{i=2}^{L} \hat{b}_i$$

This is akin to the Johnson's "molt estimate," where the population size is estimated by summing the unadjusted m_i over the entire season. (This simple procedure has apparently worked satisfactorily with the proper choice of sampling interval, so that unobserved molting seals and multiple counts of molting seals balance out over the season.)

For ease in computing the estimates of b_i, it is convenient to approximate the continuous functions of $E(b_i)$ and $E(m_i)$ by discrete probability functions and replace the integrals by summations. Thus

$$\hat{b}_i = \frac{m_i \sum_{j=t_{i-1}}^{t_i} g_j}{\mu_i \left\{ \sum_{j=1}^{t_i} g_j - \sum_{j=1}^{t_i} g_j \sum_{k=1}^{t_i-j+1} f_k \right\}}$$

where g_j = probability that a seal begins its molt on the j^{th} day of the molting season (at the beginning of the day), $j = 1, 2, \ldots, \lambda$.

f_k = probability that a seal completes its molt exactly k days after beginning it, $k = 1, 2, \ldots, \tau$.

Variance Estimates and Confidence Intervals

We turn now to the precision of the CIR estimates, as measured by $\sigma^2_{\hat{N}}$, the variance of $\hat{N}$. The estimator $\hat{N}$ at (4) is a function of several random variables, each with its own variance. The delta method and other results may be used to combine these in an approximation of $\sigma^2_{\hat{N}}$. Ignoring all covariances, and replacing expected values by observations or estimates, an approximation to $\sigma^2_{\hat{N}}$ is

$$V(\hat{N}) = A_1 \sum_{i=2}^{L} V(\hat{b}_i) + A_2 V(m_1) + A_3 V(m_L) + A_4 V(\hat{\mu}_1) + A_5 V(\hat{\mu}_L) + A_6 V(\hat{p}_L) + A_7 V(\hat{p}_1)$$

where $A_1 = \left(\hat{p}_1 - \hat{p}_L\right)^{-2}$

$$A_2 = \left(\frac{\hat{p}_L^2 + V(\hat{p}_L)}{\left(\hat{p}_1 - \hat{p}_L\right)^2} + 1\right)\left(\frac{1}{\hat{\mu}_1^2}\right)$$

$$A_3 = \frac{\hat{p}_L^2 + V(\hat{p}_L)}{\left(\hat{p}_1 - \hat{p}_L\right)^2 \hat{\mu}_L^2}$$

$$A_4 = \left(\frac{\hat{p}_L^2 + V(\hat{p}_L)}{\left(\hat{p}_1 - \hat{p}_L\right)^2} + 1\right)\left(\frac{m_1^2}{\hat{\mu}_1^4}\right)$$

$$A_5 = \frac{\hat{p}_L^2 + V(\hat{p}_L)\, m_L^2}{\left(\hat{p}_1 - \hat{p}_L\right)\hat{\mu}_L^4}$$

$$A_6 = \left(\frac{1}{\hat{p}_1 - \hat{p}_L}\right)^2 \left(\frac{m_1}{\hat{\mu}_1} - \frac{m_L}{\hat{\mu}_L}\right)^2 + \left(\hat{N} - \frac{m_1}{\hat{\mu}_1}\right)^2 \left(\frac{1}{\hat{p}_1 - \hat{p}_L}\right)^2$$

$$A_7 = \left(\hat{N} - \frac{m_1}{\hat{\mu}_1}\right)^2 \left(\frac{1}{\hat{p}_1 - \hat{p}_L}\right)^2$$

Here we use $V(\cdot)$ to denote an estimate of a sampling variance, $\sigma^2(\cdot)$. Of the component variance terms, $V(\hat{b}_i)$ requires further elaboration. Using the same approximating techniques as above, $V(\hat{b}_i)$ may be expressed as a function of $V(\hat{g}_i)$, $V(\hat{f}_i)$, $V(m_i)$, and $V(\hat{\mu}_i)$ as follows:

$$V(\hat{b}_i) = (\hat{\mu}_i\hat{\pi}_i)^{-2}\{m_i^2 V(\hat{\phi}_i) + \hat{\phi}_i^2 V(m_i) + V(m_i)V(\hat{\phi}_i)\}$$
$$+ \left(\frac{m_i\hat{\phi}_i}{\left(\hat{\mu}_i\hat{\pi}_i\right)^2}\right)^2 \{\hat{\mu}_i^2 V(\hat{\pi}_i) + \hat{\pi}_i^2 V(\hat{\mu}_i) + V(\hat{\pi}_i)V(\hat{\mu}_i)\}$$

where
$$V(\hat{\phi}_i) = \sum_{j=t_{i-1}}^{t_i} V(\hat{g}_j)$$

and
$$V(\hat{\pi}_i) = \sum_{j=1}^{t_i} V(\hat{g}_j) + \sum_{j=1}^{t_i} \left\{ \hat{g}_j^2 \sum_{k=1}^{t_i-j+1} V(\hat{f}_k) + \left(\sum_{k=1}^{t_i-j+1} \hat{f}_k \right)^2 V(\hat{g}_j) + V(\hat{g}_j) \sum_{k=1}^{t_i-j+1} V(\hat{f}_k) \right\}.$$

To complete the estimation of $V(\hat{N})$, some assumptions have to be made about the distributions of the component random variables. A reasonable set of assumptions may be the following:

(1) $V(m_i)$: The m_i could be considered Poisson, hence $V(m_i) = m_i$. Or, given the total count of seals on the i^{th} visit, h_i, m_i could be taken as binomial random variable in which case $V(m_i) = m_i\left(1 - \frac{m_i}{h_i}\right)$

(2) $V(\hat{p}_i)$: Both $\hat{p}_1$ and $\hat{p}_L$ are binomial given the total counts of non-molting seals on the first and L^{th} surveys. Thus

$$V(\hat{p}_i) = \frac{r_i}{\left(h_i - m_i\right)^2} \left(1 - \frac{r_i}{h_i - m_i}\right)$$

where r_i = number of pre-molt seals counted on i^{th} survey.
Here we assume pre-molt and post-molt haul out probabilities are the same on the first survey, and also on the L^{th} survey.

(3) $V(\hat{\mu}_i)$: Presumably the haul out probabilities for molting seals are estimated from sampling binomial distributions also. Thus

$$V(\hat{\mu}_i) = \left(\frac{m_i}{s_i^2}\right)\left(1 - \frac{m_i}{s_i}\right)$$

where s_i = total population of molting seals at time t_i

(4) $V(\hat{g}_i)$ and $V(\hat{f}_i)$: Elements of the probability distributions for pre-molt time and duration of molt are presumably estimated by sampling multinomial populations. Thus the i^{th} element of the pre-molt time distribution is estimated as

$$\hat{g}_i = \frac{\omega_i}{\sum_{i=1}^{\lambda} \omega_i} = \frac{\omega_i}{\omega_{\bullet}}$$

with estimated variance $V(\hat{g}_i) = \left(\frac{\omega_i}{\omega_{\bullet}^2}\right)\left(1 - \frac{\omega_i}{\omega_{\bullet}}\right)$, $\quad i = 1, 2, \ldots, \lambda$

where ω_i = number of seals beginning their molt on the i^{th} day of the season. Similar equations would apply to $\hat{f}_i$ and $V(\hat{f}_i)$, $i = 1, 2, \ldots, \tau$.

Approximate 95% confidence limits on N may be computed in the usual manner, assuming normality of $\hat{N}$, as

$$\hat{N} \pm 2V(\hat{N})^{1/2} .$$

LEAST SQUARES ESTIMATES

The general model of molting dynamics leads to other estimators besides the CIR. In particular, Equation (1) immediately suggests a set of moment estimators for N, namely

$$\hat{N}_i = \frac{m_i}{\hat{\mu}_i \hat{\pi}_i} \qquad i = 1, 2, \ldots, n$$

However, since each survey produces a separate estimate the question now is how best to combine the n estimates. An intuitively obvious procedure would be to simply take the arithmetic mean of the $\hat{N}_i$, i.e.,

$$\hat{N} = \sum_{i=1}^{n} \hat{N}_i / n .$$

Another approach leading to the same result is to consider a simple least squares criterion, i.e., choose $\hat{N}$ as the value of N which minimizes

$$S = \sum_{i=1}^{n} w_i \left(m_i - E(m_i) \right)^2 \qquad (5)$$

where w_i is a statistical weight to be specified. This model assumes additive error, i.e., $m_i = \mu_i \pi_i N + \varepsilon_i$ where the ε_i are independent random errors with zero mean and covariance matrix $W^{-1}\sigma_\varepsilon^2$ and W is a diagonal matrix of the weights. The general solution to (5) is

$$\hat{N} = \frac{\sum_{i=1}^{n} w_i \pi_i \mu_i m_i}{\sum_{i=1}^{n} w_i (\mu_i \pi_i)^2} .$$

Some special cases of interest may be obtained directly using the results in Draper and Smith (1966, p. 80-81) or other standard statistical texts. Among these is the case where w_i is assumed proportional to $(\mu_i \pi_i)^{-2}$, and this is

$$\hat{N} = \frac{\sum_{i=1}^{n} \frac{m_i}{\mu_i \pi_i}}{n} = \frac{\sum_{i=1}^{n} \hat{N}_i}{n}$$

as in the intuitive estimate above.

The variance of $\hat{N}$ may be estimated by

$$V(\hat{N}) = \frac{\sum_{i=1}^{n} (\hat{N}_i - \hat{N})^2}{n(n-1)} .$$

Another set of least squares estimators is appropriate when the error is multiplicative, i.e., when the model is $m_i = \mu_i \pi_i N e^{\varepsilon_i'}$. In this event

$$\ln m_i = \ln N + \ln(\mu_i \pi_i) + \varepsilon_i'$$

where ε_i' is independently distributed with mean zero and covariance matrix $W'^{-1}\sigma^2_{\varepsilon'}$. Now the least squares estimate of $\ln N$ is found by minimizing

$$S' = \sum_{i=1}^{n} w_i'(\ln m_i - (\ln N + \ln(\mu_i \pi_i)))^2 .$$

The general result is

$$\widehat{\ln(N)} = \frac{\sum_{i=1}^{n} w_i' \ln\left(\frac{m_i}{\mu_i \pi_i}\right)}{\sum_{i=1}^{n} w_i'} .$$

When the variance of m_i is proportional to $(\mu_i \pi_i)^2$ as in the special case discussed above, it is well-known that a logarithmic transformation will stabilize the variance so equal weighting will be appropriate in the logarithmic model. Hence in this special case,

$$\widehat{\ln N} = \frac{\sum_{i=1}^{n} \ln\left(\frac{m_i}{\mu_i \pi_i}\right)}{n} = \frac{\sum_{i=1}^{n} \ln \hat{N}_i}{n}$$

with estimated variance

$$V(\widehat{\ln N}) = \frac{\sum_{i=1}^{n}\left(\ln \hat{N}_i - \widehat{\ln N}\right)^2}{n(n-1)} .$$

The corresponding estimate of N is

$$N^* = e^{\widehat{\ln N}} = \left\{ \prod_{i=1}^{n} \hat{N}_i \right\}^{1/n}$$

i.e., the geometric mean of the individual $\hat{N}_i$.

When the additive model is employed, confidence limits on N may be estimated assuming normality of $\hat{N}$. Approximate 95% confidence limits then are $\hat{N} \pm 2V(\hat{N})^{1/2}$. When the multiplicative model is used, confidence limits for N may be constructed by transforming limits for $\ln N$. Thus if $\widehat{\ln N}$ is normally distributed, approximate 95% confidence limits on N are $N^* \, e^{\pm 2V(\widehat{\ln N})^{1/2}}$.

NUMERICAL EXAMPLES

To demonstrate the use of the estimation techniques, consider the following hypothetical set of survey observations:

i	t_i	h_i	r_i	m_i
1	0	56	56	0
2	8	49	48	1
3	16	47	46	1
4	24	58	57	0
5	32	61	57	3
6	40	47	43	2
7	48	51	40	7
8	56	61	45	5
9	64	70	49	11
10	72	74	41	16
11	80	77	30	18
- - -	- - -	- - -	- - -	- - -
12	88	75	28	9
13	96	81	25	11
14	104	62	16	7
15	112	69	15	12
16	120	67	7	9

In addition to this set of survey observations it is assumed from other studies that the probability functions for pre-molt time and molt duration are as follows:

$$g_i = \begin{cases} 0.0005 & i = 1, 2, \ldots, 15 \\ 0.001 & i = 16, 17, \ldots, 30 \\ 0.0034 & i = 31, 32, \ldots, 45 \\ 0.0078 & i = 46, 47, \ldots, 60 \\ 0.0137 & i = 61, 62, \ldots, 75 \\ 0.0127 & i = 76, 77, \ldots, 90 \\ 0.0074 & i = 91, 92, \ldots, 105 \\ 0.0083 & i = 106, 107, \ldots, 120 \\ 0 & i = 121, 122, \ldots \end{cases}$$

$$f_i = \begin{cases} 0 & i = 1, 2, 3 \\ 0.0139 & i = 4 \\ 0.0208 & i = 5 \\ 0.903 & i = 6 \\ 0.0972 & i = 7 \\ 0.2222 & i = 8 \\ 0.1875 & i = 9 \\ 0.1806 & i = 10 \\ 0.0972 & i = 11 \\ 0.0764 & i = 12 \\ 0.139 & i = 13 \\ 0 & i = 14, 15, \ldots \end{cases}$$

It is further assumed that at times t_1 and t_{11} the pre-molt and post-molt seals are hauled out with equal probability. This gives

$$\hat{p}_1 = \frac{56}{56 - 0} = 1.0$$

and

$$\hat{p}_L = \hat{p}_{11} = \frac{30}{77 - 18} = 0.5085 .$$

In addition, it is known that $\mu_i = 0.91$ for all i. Under these conditions the following basic statistics are computed:

i	$\hat{\pi}_i$	$\hat{\phi}_i$	$\hat{N}_i$	$\hat{b}_i$	$\hat{c}_i$
2	0.00358	0.0040	307	1.2	0.1
3	0.00441	0.0045	249	1.1	1.1
4	0.00767	0.0080	--	5.4 (rows 4–5)	3.2 (rows 4–5)
5	0.01262	0.0128	261		
6	0.02634	0.0272	83	2.3	3.4
7	0.03979	0.0404	193	7.8	2.3
8	0.06093	0.0624	90	5.6	7.8
9	0.08451	0.0860	143	12.3	5.7
10	0.10713	0.1096	164	18.0	12.5
11	0.10218	0.1046	194	20.2	18.0
12	0.09931	0.1016	100	10.1	20.0
13	0.06843	0.0698	177	12.3	10.1
14	0.05787	0.0592	133	7.9	12.3
15	0.06381	0.0655	207	13.5	8.0
16	0.06490	0.0664	152	10.1	13.4

Change-in-Ratio Estimates

To compute the CIR estimates we use statistics for $i = 2, 3, \ldots, 11$. On the fourth survey $m_i = 0$, so we derive a combined estimate of the quantity $(b_4 + b_5)$ which is $(3)(0.0080 + 0.0128)/(0.01262)(0.91) = 5.4$. Also we estimate $(c_4 + c_5)$ by $(1/0.91) - (3/0.91) + 5.4 = 3.2$.

The CIR estimate of N based on the b_i is

$$\hat{N} = \frac{73.9 - (0.5085)\left(\frac{18}{0.91}\right)}{1.0 - 0.5085} + 0$$

$$= \frac{73.9 - 10.0582}{0.4915}$$

$$= 129.9 .$$

The corresponding estimate based on the c_i is

$$\hat{N}' = \frac{54.1 + (0.4915)\left(\frac{18}{0.91}\right)}{0.4915} + 0$$

$$= \frac{54.1 + 9.722}{0.4915}$$

$$= 129.8$$

which shows that $\hat{N}$ and $\hat{N}'$ are identical.

An updated version of this paper will include variance estimates and confidence limits based on the CIR methodology.

Least Square Estimates

Based on the 14 individual estimates of N, the additive error model yields

$$\hat{N} = \frac{\sum_{i=1}^{14} \hat{N}_i}{14} = 175.2$$

with variance estimate

$$V(\hat{N}) = 311.87 .$$

So 95% confidence limits on N are 139.9 and 210.5.

The multiplicative model gives an estimate of

$$N^* = \left(\prod_{i=1}^{14} \hat{N}_i \right)^{1/14} = 163.4$$

with 95% confidence limits on N of 132.3 and 201.9.

LITERATURE CITED

DRAPER, N., and H. SMITH.

1966. Applied regression analysis. John Wiley & Sons, 407 p.

SEBER, G. A. F.

1973. The estimation of animal abundance. Griffin, 506 p.

TEXT FOOTNOTES

1. Personal communication between Lee Eberhardt and Brian and Pattie Johnson, 3 September 1979.

2. Personal communication from Douglas DeMaster to Robert Hofman, Marine Mammal Commission.