Tag Loss and the Petersen Mark-Recapture Experiment

G. A. F. Seber; R. Felton

Biometrika, Vol. 68, No. 1 (Apr., 1981), 211-219.

Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28198104%2968%3A1%3C211%3ATLATPM%3E2.0.CO%3B2-P

Biometrika is currently published by Biometrika Trust.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/bio.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

http://www.jstor.org/
Mon Jan 2 18:27:39 2006

Biometrika (1981), **68**, 1, *pp.* 211–19
Printed in Great Britain

Tag loss and the Petersen mark-recapture experiment

BY G. A. F. SEBER AND R. FELTON
Department of Mathematics, Auckland University

SUMMARY

A problem in the Petersen capture-mark-recapture experiment (Seber, 1973, Chapter 3) is the possible loss of tags. The effect of such a loss on the usual estimate of population size and the estimate of its variance is investigated for both single and double tagging. When there is double tagging, a correction for tag loss can be made based on the numbers recaptured that retain one or both tags, provided we can assume independence of the tags. Two cases are considered, distinguishable tags and nondistinguishable tags in which the latter have the same loss rates; new approximately unbiased estimates of population size and their variance estimates are introduced. It is shown that these estimates are moderately robust to dependence between the two tags and that there is little to be gained by using nondistinguishable tags.

Some key words: Capture-recapture; Double tagging; Mark-recapture; Tag loss; Population size.

1. INTRODUCTION

The so-called Petersen capture-mark-recapture experiment consists of taking a sample of individuals, called sample 1, from a closed population, marking them, and then releasing the marked sample into the population. After allowing for the marked and unmarked to mix, a second sample, sample 2, is taken and the number of marked individuals in the sample noted. Let N be the size of the population, n_i be the number caught in sample i $(i = 1, 2)$, m_2 be the number marked or tagged in sample 2, and $u_2 = n_2 - m_2$ be the number unmarked in sample 2. Here m_2 is the random variable and the n_i $(i = 1, 2)$ are assumed, for the moment, to be constants. Later, when n_2 is assumed random, we let $p_2 = 1 - q_2$ be the probability of capture in sample 2. It is usually assumed that:

(a) the population is closed, that is N is constant;
(b) all individuals have the same probability of being caught in sample 1;
(c) marking does not affect the catchability of an individual;
(d) the second sample is a simple random sample, i.e. each of the possible samples has an equal chance of being chosen;
(e) individuals do not lose their tags in the time between the two samples;
(f) all marks are reported on recovery in the second sample.

Given these assumptions, we see that m_2 has the hypergeometric distribution

$$f_1(m_2 \mid n_1, n_2) = \binom{n_1}{m_2}\binom{N-n_1}{n_2-m_2}\bigg/\binom{N}{n_2}, \tag{1}$$

for $\max\{0, n_1+n_2-N\} \leqslant m_2 \leqslant \min\{n_1, n_2\}$. Robson (1969) notes that the above model still holds even if only one of the samples is random because (1) is symmetric in n_1 and n_2.

It transpires that the Petersen estimate $\hat{N} = n_1 n_2/m_2$, suitably rounded to an integer, is the maximum likelihood estimate of N for (1). However a modification

$$N^* = (n_1+1)(n_2+1)/(m_2+1)-1, \tag{2}$$

proposed by Robson & Regier (1964), is unbiased when $n_1+n_2 \geqslant N$, and approximately unbiased otherwise. In this latter case

$$E(N^*\,|\,n_1,n_2) \doteqdot N(1-e^{-\mu}),$$

where $\mu = E(m_2\,|\,n_1,n_2) = n_1 n_2/N$. Seber (1970) and Wittes (1972) showed independently that an estimate of $\operatorname{var}(N^*\,|\,n_1,n_2)$, the variance of N^*, with similar properties is given by

$$v^* = \frac{(n_1+1)(n_2+1)(n_1-m_2)(n_2-m_2)}{(m_2+1)^2(m_2+2)} \tag{3}$$

$$= \frac{(n_1+1)^2(n_2+1)^2}{(m_2+1)^2} - \frac{(n_1+1)(n_1+2)(n_2+1)(n_2+2)}{(m_2+1)(m_2+2)} + \frac{(n_1+1)(n_2+1)}{m_2+1} \tag{4}$$

$$= (N^*+1)^2 - A + (N^*+1), \tag{5}$$

say, where A is $(N^*+1)^2$ expressed as factorials instead of powers. A key part in the proof of approximate unbiasedness is to show that

$$E(A\,|\,n_1,n_2) \doteqdot (N+1)(N+2)\{1-(\mu+1)e^{-\mu}\} \doteqdot (N+1)(N+2)$$

for moderate μ. Also by a Taylor expansion, the so-called delta method, it can be shown that

$$\operatorname{var}(N^*\,|\,n_1,n_2) \doteqdot \frac{(n_1+1)^2(n_2+1)^2}{(n_1 n_2+N)^4} n_1 n_2 N(N-n_1)(N-n_2) - \frac{(N-n_1)^2(N-n_2)^2}{(n_1 n_2+N)^2},$$

where the second term will be of smaller order than the first for moderate μ. A reasonable large sample working approximation is

$$\operatorname{var}(N^*\,|\,n_1,n_2) \doteqdot \operatorname{var}(\hat{N}\,|\,n_1,n_2) \doteqdot N^3/(n_1 n_2) = N^2/\mu. \tag{6}$$

Use of (1) emphasizes the fact that it is basically the activity of the experimenter that brings about random sampling. However another approach in which randomness is related to the activity of the individuals leads to random sample sizes. If n_1 and n_2 are regarded as independent binomial random variables, then the unconditional probability function, $f_2(m_2, n_1, n_2)$, is the multinomial model proposed by Darroch (Seber, 1973, eqn (4·1), $s = 2$). We can also use an 'intermediate' model in which n_1 is fixed but n_2 is binomial with parameters (N, p_2): we obtain a product of two independent binomial distributions, namely

$$f_3(m_2, u_2\,|\,n_1) = \binom{n_1}{m_2} p_2^{m_2} q_2^{n_1-m_2} \binom{N-n_1}{u_2} p_2^{u_2} q_2^{N-n_1-u_2}. \tag{7}$$

In practice there is little difference between the three models. For example taking expectations with respect to n_2, or n_1 and n_2, we find that N^* and v^* are approximately unbiased for all three models. Also, to the order of approximation used, the results obtained below hold for all three models. However we shall use model (7) because exact results can be found. We note that

$$\operatorname{var}(N^*\,|\,n_1) = E_{n_2}\{\operatorname{var}(N^*\,|\,n_1, n_2)\} + \operatorname{var}_{n_2}\{E(N^*\,|\,n_1, n_2)\} \doteq \operatorname{var}(N^*\,|\,n_1, n_2 = Np_2). \tag{8}$$

The following lemma will be used often.

LEMMA. *If Y is binomial with parameters (N,p):*
(i)

$$E\left\{\frac{(N+1)(N+2)\dots(N+k)}{(Y+1)(Y+2)\dots(Y+k)}\right\} = \frac{1}{p^k}\left\{1-\sum_{u=0}^{k-1}\binom{N+k}{u}p^u q^{N+k-u}\right\}; \tag{9}$$

and (ii) $E\{(1-\pi)^Y\} = (1-p\pi)^N$.

In particular we use

$$E\left(\frac{N+1}{Y+1}\right) = \frac{1}{p}(1-q^{N+1}), \quad E\left\{\frac{(N+1)(N+2)}{(Y+1)(Y+2)}\right\} = \frac{1}{p^2}\{1-q^{N+2}-(N+2)pq^{N+1}\}. \tag{10}$$

2. Effect of tag loss

Suppose that assumption (e) is false, Let $\theta = 1-\theta_0$ be the probability that a mark or tag is retained between samples, and let m_T and $m_0 = m_2 - m_T$ be the numbers caught in sample 2 which have respectively retained and lost their tags. Replacing m_2 by m_T, the observed number of tagged individuals, in N^* and v^*, we have estimates N_T^* and v_T^*, say, and we now investigate their properties. We note that (3) still equals (4) when m_2 is replaced by m_T.

Assuming that individuals lose their tags independently of one another, we have that, given m_2, m_T is binomial with parameters (m_2, θ). Since m_2 is binomial with parameters (n_1, p_2) we see that, given n_1, m_T is binomial with parameters $(n_1, \theta p_2)$. Using the suffix 1 to denote 'conditional on n_1', we have, from Appendix 1, with the recommendation that $m_T \geqslant 7$,

$$E_1(N_T^*) \doteq (N+1)/\theta - 1 = N + (N+1)\,\theta_0/(1-\theta_0) = N + \beta N,$$

say, and

$$E_1(v_T^*) \doteq \operatorname{var}_1(N_T^*) - (N+1)\,\theta_0/(1-\theta_0)^2 = \operatorname{var}_1(N_T^*)\,(1-\gamma),$$

say, where β and γ are the proportional biases of the two estimates, respectively. Using the same order of approximation that lead to (6) and assuming n_1 and n_2 are small compared to N, we find, using the delta method and conditional arguments like that leading to (8), that

$$\operatorname{var}_1(N_T^*) \doteq \operatorname{var}_1(n_1 n_2/m_T) \doteq N^2/(\theta^3 n_1 p_2) = N^2/\{\theta^2 E_1(m_T)\}, \quad \gamma \doteq \theta_0 E_1(m_T)/N.$$

In general γ will be negligible for large N and all θ_0, but β may be unacceptably large for moderate θ_0. For example if the rate of tag loss is 20%, $\theta_0 = 0{\cdot}2$ and $\beta \doteq \theta_0/(1-\theta_0) = 0{\cdot}25$. As suggested by Cochran (1977, pp. 12–15), the bias can be ignored in confidence intervals if it is less than 10% of the standard deviation, i.e. if $\theta_0 < 0{\cdot}1/\{\sqrt{E_1(m_T)}\}$.

If double tagging is used the situation is, as expected, much better. The above results still apply but with θ_0 replaced by θ_{00}, the probability of losing both tags. Thus

$$E_1(N_T^*) \doteqdot N+(N+1)\,\theta_{00}/(1-\theta_{00}),$$

and if we simply ignore those which have lost both tags then the bias in N_T^* may be small. For example if the tags are identical and independent then $\theta_{00}=\theta_0^2$ and $\theta_{00}/(1-\theta_{00}) < \theta_0\,\theta_0/(1-\theta_0)$. If $\theta_0 = 0{\cdot}2$ then $\theta_{00}/(1-\theta_{00}) = 0{\cdot}042$.

3. Correction for tag loss

3·1. *Two distinguishable tags*

Suppose all individuals caught in the first sample are given two tags denoted by A and B. Let m_x $(x = A, B)$, m_{AB}, $m_T = m_A+m_B+m_{AB}$ and $m_0 = m_2-m_T$ be the numbers in sample 2 with only tag x, with both tags, with at least one tag, and with both tags missing, respectively. Let θ_x $(x = A, B)$, θ_{AB}, θ_{A0}, θ_{0B} and $\theta_{00} = 1-\theta_{A0}-\theta_{0B}-\theta_{AB}$ be the respective probabilities of retaining between samples tag x, both tags, only tag A, only tag B and neither tag. Assuming that individuals are independent with respect to tag loss, we have the multinomial distribution

$$f(m_A, m_B, m_{AB}\,|\,m_2) = \frac{m_2!}{m_A!\,m_B!\,m_{AB}!\,m_0!}\,\theta_{A0}^{m_A}\,\theta_{0B}^{m_B}\,\theta_{AB}^{m_{AB}}\,\theta_{00}^{m_0}. \tag{11}$$

To obtain a correction for tag loss we now make the assumption of independent tags, that is $\theta_{A0} = \theta_A(1-\theta_B)$, $\theta_{AB} = \theta_A\,\theta_B$, etc. Substituting in (11), we obtain the following maximum likelihood estimates of θ_A, θ_B and m_2 (Seber, 1973, p. 95 with π replaced by $1-\theta$):

$$\hat\theta_A = m_{AB}/(m_{AB}+m_B), \quad \hat\theta_B = m_{AB}/(m_{AB}+m_A),$$

$$\hat m_2 = (m_A+m_{AB})\,(m_B+m_{AB})/m_{AB} = m_T+(m_A\,m_B)/m_{AB}.$$

A natural estimate of N is $\hat N_{AB} = n_1\,n_2/\hat m_2$ and, using the delta method, we find that

$$\operatorname{var}_1(\hat N_{AB}) \doteqdot \frac{N^3}{n_1(Np_2)}\left\{\frac{(1-\theta_A)\,(1-\theta_B)}{\theta_A\,\theta_B}+1\right\}. \tag{12}$$

Incidentally the formula (3·25) given by Seber (1973) is incorrect. By analogy with (2), we consider a slight modification of $(m_1+1)\,(n_2+1)/(\hat m_2+1)-1$, namely

$$N_{AB}^* = \frac{(n_1+1)\,(n_2+1)}{(m_T+1)}\left\{1-\frac{m_A\,m_B}{m_T(m_{AB}+1)}\right\}-1. \tag{13}$$

Using the assumption of independent tags we find, from Appendix 2, that

$$E_1(N_{AB}^*) \doteqdot \frac{1-2\theta_{00}}{(1-\theta_{00})^2}\,(N+1)-1 = N-(N+1)\,\theta_{00}^2/(1-\theta_{00})^2.$$

In general the bias will be negligible: for example if $\theta_{00} = \theta_0^2$ and $\theta_0 < 0{\cdot}3$ then $\theta_{00}^2/(1-\theta_{00})^2 < 0{\cdot}01$.

An approximately unbiased estimate of $\operatorname{var}_1(N_{AB}^*)$ can be constructed along the lines of (5), namely

$$v_{AB}^* = (N_{AB}^*+1)^2-A_1+(N_{AB}^*+1),$$

where A_1 is $(N^*_{AB}+1)^2$ expressed as factorials, that is

$$A_1 = \frac{(n_1+1)(n_1+2)(n_2+1)(n_2+2)}{(m_T+1)(m_T+2)}\left\{1-\frac{2m_A m_B}{m_T(m_{AB}+1)}+\frac{m_A(m_A-1)m_B(m_B-1)}{m_T(m_T-1)(m_{AB}+1)(m_{AB}+2)}\right\}.$$

Using the method of Appendix 2, we find that

$$E_1(A_1) \simeq \frac{(N+1)(N+2)}{(1-\theta_{00})^2}\left(1-\frac{\theta_{00}}{1-\theta_{00}}\right)^2,$$

$$E_1(v^*_{AB}) \simeq \operatorname{var}_1(N^*_{AB})+(N+1)\left\{1-\frac{\theta_{00}^2}{(1-\theta_{00})^2}\right\}\frac{\theta_{00}^2}{(1-\theta_{00})^2}.$$

Since $\operatorname{var}_1(N^*_{AB}) \simeq \operatorname{var}_1(\hat{N}_{AB})$ we see, from (12), that the above bias is negligible and v^*_{AB} is approximately unbiased.

The key assumption underlying the above theory is the assumption of independence, that is $\theta_{00} = (1-\theta_A)(1-\theta_B)$. We now consider briefly what effect departures from this assumption have on the estimates. Using the method of Appendix 2 with (A4) instead of (A5), we find that

$$E_1(N^*_{AB}) \simeq (N+1)\alpha_{AB}-1 = N-(N+1)(1-\delta)\theta_{00}^2/(1-\theta_{00})^2$$

where

$$\alpha_{AB} = \frac{1}{1-\theta_{00}}\left\{1-\frac{\theta_{A0}\,\theta_{0B}}{(1-\theta_{00})\,\theta_{AB}}\right\},$$

$$\delta = (\theta_{00}\,\theta_{AB}-\theta_{A0}\,\theta_{0B})/(\theta_{00}^2\,\theta_{AB}) = (\theta_{AB}-\theta_A\,\theta_B)/(\theta_{00}^2\,\theta_{AB})$$

and δ is zero if and only if the tags are independent. Define the pair of indicator variables (X, Y) taking values $(1,1)$, $(1,0)$, $(0,1)$ and $(0,0)$ with respective probabilities θ_{AB}, θ_{A0}, θ_{0B} and θ_{00}. Then $\operatorname{cov}(X, Y) = \theta_{AB}-\theta_A\,\theta_B$ and

$$\theta_{AB} = \rho\{\theta_A(1-\theta_A)\,\theta_B(1-\theta_B)\}^{\frac{1}{2}}+\theta_A\,\theta_B.$$

Writing $1-\theta_{00} = \theta_A+\theta_B-\theta_{AB}$, we see that δ can be expressed as a function of θ_A, θ_B and ρ, the correlation between X and Y, as in Table 1. Also calculated is $\delta_1 = -(1-\delta)\,\theta_{00}^2/(1-\theta_{00})^2 = \alpha_{AB}-1$, the approximate proportional bias of N^*_{AB}, for $\delta_1 \leqslant 0{\cdot}05$. The effect of the correlation increases as the retention rates θ_A and θ_B go up. However provided these rates are reasonably high, a fair degree of correlation can be tolerated without causing appreciable bias.

Using similar arguments it can be shown that the bias term for v^*_{AB} is now $(N+1)\,\alpha_{AB}(1-\alpha_{AB}) = -(N+1)\,\delta_2$ and δ_2 is included in Table 1. We note that δ_2 increases with ρ and, from a more extensive table, it is seen that $|\delta_2| < 1$ when $\theta_A, \theta_B \geqslant 0{\cdot}6$ and $\rho \leqslant 0{\cdot}9$. Since $\operatorname{var}_1(N^*_{AB})$ is of order $N^2/E_1(m_2)$, the above bias term can therefore be neglected for any $\rho \leqslant 0{\cdot}9$ if $N/E_1(m_2)$ is reasonably large.

3·2. *Two indistinguishable independent tags*

Assuming indistinguishable tags we now have $\theta_A = \theta_B = \theta = 1-\theta_0$ say. After substitution in (11), the maximum likelihood estimates of θ and m_2 are (Seber, 1973, p. 96)

$$\tilde{\theta} = 2m_{AB}/(m_C+2m_{AB}),\quad \tilde{m}_2 = \tfrac{1}{4}(m_C+2m_{AB})^2/m_{AB} = m_T+\tfrac{1}{4}m_C^2/m_{AB},$$

Table 1. *The effect of tag dependence on estimator bias*

θ_A	θ_B	ρ	δ	δ_1	δ_2	θ_A	θ_B	ρ	δ	δ_1	δ_2
0·6	0·6	0·1	1·85	0·043	0·045	0·8	0·8	0·1	7·78	0·024	0·024
	0·7	0·1	2·50	0·041	0·043			0·2	9·18	0·493	0·052
	0·8	0·1	3·95	0·036	0·038	0·8	0·9	0·1	16·01	0·016	0·017
	0·9	0·1	8·86	0·026	0·027			0·2	16·66	0·033	0·034
								0·3	15·18	0·050	0·052
0·7	0·7	0·1	3·34	0·036	0·038						
	0·8	0·1	5·17	0·030	0·031	0·9	0·9	0·1	30·44	0·011	0·011
	0·9	0·1	11·15	0·021	0·022			0·2	27·73	0·022	0·023
		0·2	12·65	0·043	0·045			0·3	23·56	0·033	0·034
								0·4	20·11	0·040	0·046

where $m_C = m_A + m_B$, the total number in sample 2 with just one tag. The estimate $\tilde{\theta}$ has been suggested independently by Chapman *et al.* (1965, p. 340), Cormack (1968, p. 463), Caughley (1971) and Hubert *et al.* (1976, p. 165). It has been used, for example, by Best & Rand (1975) and Eberhardt, Chapman & Gilbert (1979, p. 8). The apparently different expressions for $\tilde{m}$ can be reconciled by noting the various representations

$$\tilde{m}_2 = m_{AB}/\tilde{\theta}^2 = \tfrac{1}{2}(m_C + 2m_{AB})/\tilde{\theta} = m_T/\{1-(1-\tilde{\theta})^2\}.$$

By analogy with (13) we consider the modified estimate

$$N_C^* = \frac{(n_1+1)(n_2+1)}{(m_T+1)}\left\{1 - \frac{m_C(m_C-1)}{4m_T(m_{AB}+1)}\right\} - 1.$$

Using the method of Appendix 2 we find that, for independent tags,

$$E_1(N_C^*) \simeq (N+1)\alpha_C - 1 = N-(N+1)\left\{\frac{\theta_{00}^2}{(1-\theta_{00})^2} + \frac{(\theta_A-\theta_B)^2}{4\theta_A\theta_B(1-\theta_{00})^2}\right\}, \tag{14}$$

where

$$\alpha_C = \frac{1}{1-\theta_{00}}\left\{1 - \frac{(1-\theta_{00}-\theta_{AB})^2}{4\theta_{AB}(1-\theta_{00})}\right\},$$

and $\theta_{00} = (1-\theta_A)(1-\theta_B)$. Thus when $\theta_A = \theta_B$, $\alpha_C = \alpha_{AB}$ and N_C^* has approximately the same bias as N_{AB}^*. Since $4\theta_A\theta_B(1-\theta_{00})^2$ generally exceeds 1, the second term of (14) will be negligible and $|\alpha_C - \alpha_{AB}| < 0{\cdot}01$ if $|\theta_A - \theta_B| < 0{\cdot}1$, for independent tags.

An estimate of $\text{var}_1(N_C^*)$ is given by $v_C^* = (N_C^*+1)^2 - A_2 + (N_C^*+1)$, where

$$A_2 = \frac{(n_1+1)(n_1+2)(n_2+1)(n_2+2)}{(m_T+1)(m_T+2)} \times\left\{1 - \frac{m_C(m_C-1)}{2m_T(m_{AB}+1)} + \frac{m_C(m_C-1)(m_C-2)(m_C-3)}{16m_T(m_T-1)(m_{AB}+1)(m_{AB}+2)}\right\}.$$

It can be shown that v_C^* has approximate bias $(N+1)\alpha_C(1-\alpha_C)$ and this is equal to that of v_{AB}^* when $\theta_A = \theta_B$.

The assumption H_0: $\theta_A = \theta_B = \theta$ can be tested using the binomial model $f(m_A \mid m_A + m_B)$ (Seber, 1973, p. 96). However such a test will not be very sensitive and, in the light of the above discussion, we ask if the gain in efficiency through estimating one less parameter is sufficient payoff for the uncertainty of H_0. Using the delta method we can show that

(Seber, 1973, p. 96; Eberhardt *et al.*, 1979, p. 9)

$$\text{var}_1\,(N_C^*) \doteq \text{var}_1\,(\tilde{N}_C) \doteq \frac{N^3}{n_1(Np_2)}\left\{\frac{(1-\theta)^2}{\theta^2}+1\right\},$$

which is simply (12) with $\theta_A = \theta_B = \theta$. Thus although we would expect a small gain in efficiency if we could assume H_0, it will be slight because there is no large sample gain. We therefore recommend that N_{AB}^* be used where possible. However there are situations, e.g. commercial fishing, where it may not be practicable to distinguish between the tags, and N_C^* has to be used. In this case the proportional bias of N_C^* will be of order $(\theta_A-\theta_B)^2$.

Sometimes it is possible to attach a third permanent tag to a few individuals, few because of expense, thus providing an m_0 for this group. An estimate of ρ can then be found.

The above theory also applies to Bailey's binomial model for the Petersen experiment (Seber, 1973, p. 61) by simply replacing both (n_1+1) and (n_1+2) by n_1.

Appendix 1

Effect of tag loss on the usual Petersen estimates

We have that

$$N_T^*+1 = (n_1+1)\left\{\frac{u_2}{m_T+1}+\frac{m_2+1}{m_T+1}\right\}.$$

From (9) we have

$$E\{1/(m_T+1)\,|\,n_1\} = \{(n_1+1)\,p_2\,\theta\}^{-1}\,\{1-(1-p_2\,\theta)^{n_1+1}\} \tag{A1}$$

since m_T is conditionally binomial with parameters $(n_1, p_2\,\theta)$, and

$$\begin{aligned} E_{m_2}\{E(m_2+1)/(m_T+1)\,|\,m_2, n_1\}] &= \theta^{-1}\{1-E_{m_2}(1-\theta)^{m_2+1}\} \\ &= \theta^{-1}\{1-(1-\theta)\,(1-p_2\,\theta)^{n_1}\}, \end{aligned} \tag{A2}$$

by Lemma (ii). Now $E(u_2\,|\,n_1) = (N-n_1)\,p_2$ and u_2 is conditionally independent of m_T so that, denoting $E(\,.\,|\,n_1)$ by E_1, we have

$$E_1(N_T^*+1) = (N+1)/\theta-b,$$

where

$$b = \theta^{-1}\{(n_1+1)\,(1-\theta)\,(1-p_2\,\theta)^{n_1}+(N-n_1)\,(1-p_2\,\theta)^{n_1+1}\}.$$

Since $(1-p_2\,\theta)^{n_1} \doteq \exp(-n_1\,p_2\,\theta) = \exp\{-E(m_T|\,n_1)\}$, we see that proportionately b will be small provided m_T is not too small. Following Robson & Regier (1964), we recommend $m_T \geqslant 7$ for negligible b.

Now,

$$v_T^* = (N_T^*+1)^2-A_T+(N_T^*+1), \tag{A3}$$

where

$$\begin{aligned} A_T &= (n_1+1)\,(n_1+2)\,(n_2+1)\,(n_2+2)/\{(m_T+1)\,(m_T+2)\} \\ &= (n_1+1)\,(n_1+2)\,\{u_2(u_2-1) \\ &\qquad +2u_2(m_2+2)+(m_2+1)\,(m_2+2)\}/\{(m_T+1)\,(m_T+2)\}. \end{aligned}$$

Using similar methods to those used for deriving (A1) and (A2), and applying the righthand equation (10), we find that

$$E_1(A_T) = (n_1+1)(n_1+2)\left\{\frac{(N-n_1)(N-n_1-1)}{(n_1+1)(n_1+2)}\frac{1}{\theta^2}+\frac{2(N-n_1)}{(n_1+1)\theta^2}+\frac{1}{\theta^2}\right\}-B$$
$$= (N+1)(N+2)/\theta^2-B,$$

where

$$B = (N-n_1)(N-n_1-1)\theta^{-2}\{1+(n_1+1)p_2\theta\}(1-p_2\theta)^{n_1+1}$$
$$+2(N-n_1)(n_1+2)\theta^{-2}\{1+n_1(1-\theta)\theta p_2-\theta^2 p_2\}(1-p_2\theta)^{n_1}$$
$$+(n_1+1)(n_1+2)\theta^{-2}(1-\theta)\{(1-p_2\theta)(1+\theta)+n_1 p_2\theta(1-\theta)\}(1-p_2\theta)^{n_1-1}.$$

Once again B will be proportionately small for moderate m_T.

From (A3) we can find an exact expression for $E_1(v_T^*)$ in terms of $\mathrm{var}_1(N_T^*)$ and a bias term. However, ignoring b and B, we obtain

$$E_1(v_T^*) \doteq \mathrm{var}_1(N_T^*+1)+(N+1)\theta^{-2}-(N+1)(N+2)\theta^{-2}+(N+1)\theta^{-1}$$
$$\doteq \mathrm{var}_1(N_T^*)-(N+1)(1-\theta)\theta^{-2}.$$

APPENDIX 2

Expected value of a Petersen estimate corrected for tag loss

Now $f(m_A, m_B, m_{AB}|m_T)$ is multinomial with parameters $P_{xy} = \theta_{xy}/(1-\theta_{00})$ so that, if we assume independent tags,

$$E\{m_A m_B/(m_{AB}+1)|m_T\} = m_T P_{AO} P_{OB} P_{AB}^{-1}\{1-(1-P_{AB})^{m_T-1}\}$$
$$= m_T\frac{\theta_{00}}{1-\theta_{00}}\left\{1-\left(1-\frac{\theta_A\theta_B}{1-\theta_{00}}\right)^{m_T-1}\right\}, \quad \text{(A4)}$$

where $\theta_{00} = (1-\theta_A)(1-\theta_B)$. Thus

$$E_1(N_{AB}^*+1) = E_1\left(\frac{(n_1+1)(n_2+1)}{m_T+1}\left[1-\frac{\theta_{00}}{1-\theta_{00}}\{1-(1-P_{AB})^{m_T-1}\}\right]\right)$$
$$= E_1\left[(N_T^*+1)\left\{\frac{1-2\theta_{00}}{1-\theta_{00}}+\frac{\theta_{00}}{1-\theta_{00}}(1-P_{AB})^{m_T-1}\right\}\right] \quad \text{(A5)}$$
$$\doteq \frac{1-2\theta_{00}}{1-\theta_{00}}E_1(N_T^*+1) \doteq \frac{1-2\theta_{00}}{(1-\theta_{00})^2}(N+1).$$

Although, after considerable algebra, the above expectation can be found exactly, we note that the second term of (A5) is negligible. For example θ_{00} is small compared with $1-2\theta_{00}$, and from Lemma (ii)

$$E_1\{(1-P_{AB})^{m_T}\} = \{1-P_{AB}p_2(1-\theta_{00})\}^{n_1} \doteq \exp\{-E_1(m_{AB})\}.$$

REFERENCES

BEST, P. B. & RAND, R. W. (1975). Results of a pup-tagging experiment on the *Arctocephalus pusillus* rookery at Seal Island, False Bay, South Africa. *Rapp. P.-v. Réun. Cons. int. Explor. Mer.* **169**, 267–73.

CAUGHLEY, G. (1971). Correction for band loss. *Bird-banding* **42**, 220–1.

CHAPMAN, D. G., FINK, B. D. & BENNETT, E. B. (1965). A method for estimating the rate of shedding of tags from yellowfin tuna. *Bull. Int.-Am. Trop. Tuna Comm.* **10**, 335–42.

COCHRAN, W. G. (1977). *Sampling Techniques*, 3rd edition. New York: Wiley.

CORMACK, R. M. (1968). The statistics for capture-recapture methods. *Oceanogr. Mar. Bio. Ann. Rev.* **6**, 455–506.

EBERHARDT, L. L., CHAPMAN, D. G. & GILBERT, J. R. (1979). *A Review of Marine Mammal Census Methods.* Wildlife Monograph No. 63. Washington, D.C.: Wildlife Society.

HUBERT, G. F., STORM, G. L., PHILLIPS, R. L. & ANDREWS, R. D. (1976). Ear tag loss in red foxes. *J. Wildl. Manag.* **40**, 164–7.

ROBSON, D. S. (1969). Mark-recapture methods of population estimation. In *Developments in Survey Sampling*, Eds N. L. Johnson and H. Smith, pp. 120–40. New York: Wiley.

ROBSON, D. S. & REGIER, H. A. (1964). Sample size in Petersen mark-recapture experiments. *Trans. Am. Fish. Soc.* **93**, 215–26.

SEBER, G. A. F. (1970). The effects of trap response on tag-recapture estimates. *Biometrics* **26**, 13–22.

SEBER, G. A. F. (1973). *The Estimation of Animal Abundance and Related Parameters.* London: Griffin.

WITTES, J. T. (1972). On the bias and estimated variance of Chapman's two-sample capture-recapture population estimate. *Biometrics* **28**, 592–7.

[*Received March* 1980. *Revised July* 1980]