This is an electronic version of an article published in Communications in Statistics – Theory and Methods, 1532-415X, Volume 29, Issue 11, 2000, Pages 2605-2619. The published article is available online at: http://www.tandfonline.com/doi/abs/10.1080/03610920008832626?journalCode=lsta20#.UcrRRhzGozM

A CONDITIONALLY-UNBIASED ESTIMATOR OF POPULATION SIZE BASED ON PLANT-CAPTURE IN CONTINUOUS TIME

I.B.J. Goudie and J. Ashbridge

School of Mathematics and Statistics,
University of St Andrews, St Andrews, Fife, KY16 9SS, Scotland.

Key Words: factorial series distribution; harmonic mean estimator; inhomogeneous Poisson process; mark-recapture; maximum likelihood; non-central Stirling number.

ABSTRACT

This paper proposes an estimator of the unknown size of a target population to which has been added a planted population of known size. The augmented population is observed for a fixed time and individuals are sighted according to independent Poisson processes. These processes may be time-inhomogeneous, but, within each population, the intensity function is the same for all individuals. When the two populations have the same intensity function, known results on factorial series distributions suggest that the proposed estimator is approximately unbiased and provide a useful estimator of standard deviation. Except for short sampling times, computational results confirm that the proposed population-size estimator is nearly unbiased, and indicate that it gives a better overall performance than existing estimators in the literature. The robustness of this performance is investigated in situations in which it cannot be assumed that the behaviour of the plants matches that of individuals from the target population.

1. INTRODUCTION

A planted population, consisting of a known number $R \geq 0$ of individuals, is added to a target population of unknown size $N \geq 0$, where $N + R > 0$. This augmented population, which is assumed to be closed, is observed continuously throughout a predetermined time period $[0, \tau]$. The times at which any member of the target population is seen follow a Poisson process with intensity function $\lambda(t)$, whilst those for any planted individual follow a Poisson process with intensity function $\mu(t)$. Thus for each population there is a common intensity function. All $N + R$ processes are independent. On the first occasion on which a member of the target population is seen it is tagged so that it can be distinguished from other members of the population. The R planted individuals are assumed to have already been marked in this way. This paper introduces a new estimator of the size N of the target population.

Despite the difference in context, the motivation for the estimator which we propose can be traced to the work of Harris (1968). On the basis of a sample of known size, he wished to estimate the number of equiprobable classes in a multinomial population. He exhibited a ratio of Stirling numbers of the second kind which is the unique unbiased estimator of the number of classes, provided this number does not exceed the sample size. Berg (1974) extended this result to estimating the size of a homogeneous population using a multiple-recapture census with a fixed number of samples and with the sample sizes regarded as constants. He presented an estimator which is minimum variance unbiased provided the population size does not exceed the sum of the sample sizes.

In the analogous problems in continuous time, the sequence of observed individuals mirrors the data set obtained by taking samples of size one in discrete-time, except that, under truncated sampling, the experimenter does not choose the total number of observations. Since, in the discrete-time problems, it is indeed a condition on the sample size that ensures that unbiased estimators exist, it might be judged surprising if such estimators could be found for models in continuous time. This viewpoint has previously been confirmed in the special case $R = 0$ of our model for which Nayak (1989) showed that no unbiased estimator of N exists.

The two models for estimating the size of a population of butterflies discussed by Craig (1953) can also be derived from the special case $R = 0$ of our model that arises when no plants are used. The maximum likelihood estimator for Craig's models is more readily calculated in the manner of Nayak (1988). Note also that this estimator is equivalent to the optimal estimator in the class obtained by Yip et al (1993) using a martingale-based approach.

Following the work of Laska and Meisner (1993), a number of papers in the last few years have addressed the benefits of deploying plants in population size problems. Their use in population size estimation was considered by Yip (1996), who provided an extension of the martingale-based approach, whilst Goudie (1995) explored how they can assist in the related problem of deciding when all individuals in the target population have been seen. In earlier work on the model discussed in the present paper, Goudie, Pollock and Ashbridge (1998) concluded that the harmonic mean estimator of N, introduced by Joe & Reid (1985), is to be preferred to both the maximum likelihood estimator and to a Petersen-type estimator.

In section 2 below we show that the conditional distribution of the number X of distinct members of the target population seen by time τ, given the total number Z of sightings of members of either population, is a factorial series distribution (FSD), in the sense of Berg (1974). By applying his results to this FSD, we obtain, in section 3, a new estimator of the population size N. When sightings of a planted individual follow the same Poisson process as that for a member of the target population, this estimator is unbiased with respect to the conditional distribution of X given Z, provided Z is sufficiently large. In section 4 we discuss other possible estimators of N for this situation, and compare their performance with that of the conditionally-unbiased estimator in section 5. Estimation of the standard deviation of the conditionally-unbiased estimator is considered in section 6. Since the plant-capture methodology requires the major assumption that detection of the target and planted individuals occurs at the same rate, we examine the sensitivity of the new estimator to this assumption in section 7.

2. PROBABILISTIC RESULTS

Standard results show that the number of times each member of the target population is seen by time τ has a Poisson distribution with mean θ, where

$$\theta = \int_0^{\tau} \lambda(t)\, dt .$$

The number of times each planted individual is seen by time τ is therefore also Poisson, and we denote its mean by $K\theta$. By independence, the total number Z of sightings of individuals of either type has a Poisson distribution with mean $(N + \tilde{R})\theta$, where $\tilde{R} = KR$. It follows from the results given in Goudie, Pollock and Ashbridge (1998) that, given $Z = z$, the conditional distribution for the number X of distinct members of the target population seen by time τ has probability function

$$(N)_x \, S(x, z; \tilde{R}) \left(N + \tilde{R}\right)^{-z} \qquad x = 0, 1, \ldots, \min(N, z). \qquad (1)$$

where $(N)_x = x! \binom{N}{x}$ and $S(x, z; \tilde{R})$ is the non-central Stirling number of the second kind, defined by

$$S(x, z; \tilde{R}) = \frac{1}{x!} \left[\Delta^x \left(N + \tilde{R}\right)^z \right]_{N=0} = \frac{1}{x!} \sum_{i=0}^{x} \binom{x}{i} (-1)^i \left(x - i + \tilde{R}\right)^z .$$

Note, in particular, that

$$S(0, z; \tilde{R}) = \tilde{R}^z \qquad (2)$$

and that use of standard results for the difference operator Δ implies that

$$S(x, z; \tilde{R}) = 0 \qquad x > z , \qquad (3)$$

$$S(x, x; \tilde{R}) = 1 \qquad x = 0, 1, 2, \ldots . \qquad (4)$$

It follows from (1) that the joint probability function of X and Z is

$$(N)_x \, S(x, z; \tilde{R}) \, \theta^z \, e^{-(N + \tilde{R})\theta} \Big/ z! \qquad \begin{array}{l} x = 0, 1, \ldots, \min(N, z); \\ z = x, x+1, \ldots . \end{array} \qquad (5)$$

The conditional distribution (1) can be written in the form of a FSD, as defined by Berg (1974), taking the series function to be $A(N) = \left(N + \tilde{R}\right)^z$. Use of equation (3) shows that the range of X may be reduced from the set $\{0, 1, \ldots, N\}$ given in Berg's general formulation to that shown in equation (1). Note also that the validity of the distribution (1) is not limited to the situation in which R is positive. In fact throughout this paper we can subsume within our general

treatment the case $R = 0$, in which no plants are used. The number $S(x, z; 0)$ that arises in this case is the usual Stirling number of the second kind, which we denote by $S(x, z)$.

The conditional probability function (1) also has the same form as a probability function obtained by Charalambides (1981) in the context of a sampling problem. Expressed in our notation, he took a single sample of size z from an augmented population comprising a target population of size Nt, consisting of t individuals of each of N types, together with a control class of s individuals. He showed that, if s/t tends to $\tilde{R}$ as t tends to infinity, the limit of the probability function of the number x of the N types that are seen is given by (1).

It is also useful to consider the embedded process obtained by restricting attention to the epochs at which a sighting occurs. Let X_z denote the number of unmarked individuals seen by the z^{th} epoch. The process $\{X_z\}$ is a Markov chain such that if $X_z = x$ then X_{z+1} equals $x+1$ with probability $(N - x)/(N + \tilde{R})$ or x with probability $(x + \tilde{R})/(N + \tilde{R})$. If $p_{x|z}$ denotes the probability of being in state x of the embedded chain after the z^{th} epoch, then it is clear that $(N + \tilde{R})p_{x|z}$ is given by

$$\begin{cases} (x + \tilde{R})p_{x|(z-1)} + (N - x + 1)\,p_{(x-1)|(z-1)} & x = 1 + \max(0, 1 - R), \ldots, \min(N, z) \\ & z = x, x + 1, x + 2, \ldots \\ (x + \tilde{R})\,p_{x|(z-1)} & x = \max(0, 1 - R) \\ & z = x, x + 1, x + 2, \ldots \end{cases} \tag{6}$$

subject to $p_{0|0} = 1$.

Substituting the conditional probabilities (1) into equations (6) shows that the non-central Stirling number $S(x, z; \tilde{R})$ satisfies a 'triangular' recurrence relation. Explicitly, $S(x, z; \tilde{R})$ is given by

$$\begin{cases} (x + \tilde{R})S(x, z - 1; \tilde{R}) + S(x - 1, z - 1; \tilde{R}) & x = 1 + \max(0, 1 - R), \ldots, \min(N, z) \\ & z = x,\ x + 1, x + 2, \ldots \\ (x + \tilde{R})\,S(x, z - 1; \tilde{R}) & x = \max(0, 1 - R) \\ & z = x,\ x + 1, x + 2, \ldots \end{cases} \tag{7}$$

In the case $R = 0$, these equations reduce to the well-known 'triangular' recurrence relation for Stirling numbers of the second kind.

3. A CONDITIONALLY-UNBIASED ESTIMATOR OF *N*

Assuming for the moment that the ratio K of the mean numbers of sightings of planted and target individuals is known, consider the estimator defined for $z = x, x+1, \ldots$ by

$$\tilde{N}_U(x, z; K) = \begin{cases} \dfrac{S(x-1, z; \tilde{R})}{S(x, z; \tilde{R})} + x & x = 1, 2, 3, \ldots \\ 0 & x = 0. \end{cases} \qquad (8)$$

The results of Berg (1974) show that, if $z \geq N$, this is the unique unbiased estimator of N, with respect to the conditional distribution of X given Z. As in the comparable situation discussed by Goudie (1995), it is intuitively reasonable that all the accessible information about N is obtained via this conditional distribution, since Z has a Poisson distribution with mean $(N + \tilde{R})\theta$, and the information about N contained in Z cannot be disentangled from that about θ. Equations (7) and (2) show that equation (8) can be re-expressed as

$$\tilde{N}_U(x, z; K) = \frac{S(x, z+1; \tilde{R})}{S(x, z; \tilde{R})} - \tilde{R} \qquad \begin{matrix} z = x, x+1, \ldots \\ x = 0, 1, 2, \ldots \end{matrix} \qquad (9)$$

In the case $R = 0$, in which no plants are used, $\tilde{N}_U$ reduces to $S(x, z+1)/S(x, z)$. This ratio of Stirling numbers of the second kind is the well-known estimator that Harris (1968) proposed for the classical occupancy problem.

Since the size of the non-central Stirling numbers of the second kind grows rapidly, computation of the estimator $\tilde{N}_U$ may, in practice, prove difficult. The problem can, however, be circumvented by generalising Property 5 of Berg (1975) to obtain a recursive relationship, again of triangular form, for $\tilde{N}_U$. More explicitly, an expression for $\tilde{N}_U(x, z; K)$ is given by the relationship

$$\frac{\tilde{N}_U(x, z; K) - x}{\tilde{N}_U(x, z-1; K) - x} = \frac{\tilde{N}_U(x-1, z-1; K) + \tilde{R}}{\tilde{N}_U(x, z-1; K) + \tilde{R}},$$

which can readily be verified by using equations (8) and (9) to express each side as ratios of non-central Stirling numbers of the second kind. The recursion is started by recalling from (8) that $\tilde{N}_U(0, z; K) = 0$, and also noting that

$$\tilde{N}_U(x, x; K) = x(2\tilde{R} + x + 1)/2. \qquad (10)$$

To verify this value for $\tilde{N}_U(x, x; K)$, observe first that, by the properties of the non-central Stirling numbers of the second kind labelled (4) and (2), repeated use

of equation (7) gives

$$S(x, x+1; \tilde{R}) = \tilde{R} + (\tilde{R}+1) + \ldots + (\tilde{R}+x) = (x+1)(2\tilde{R}+x)/2 ,$$

and then use equations (9) and (4). Notice also from (10) that the estimator remains finite when $Z = X$.

In practice we wish to deploy this estimator in situations in which the value of K is not known, but is believed to equal unity. Thus, as the target population size N is integer valued, attention is restricted from here on to the estimator $\hat{N}_U \equiv \hat{N}_U(x, z; 1) = [\tilde{N}_U(x, z; 1) + 0.5]$, where the square brackets denote the integer part. We will refer to $\hat{N}_U$ as a conditionally unbiased estimator, though strictly the appellation is more appropriate for $\tilde{N}_U$.

4. OTHER ESTIMATORS OF *N*

One of the estimators of N considered by Goudie, Pollock and Ashbridge (1998) was the maximum likelihood estimator $\hat{N}_M$ based on the likelihood for N and θ obtained from (5) with $K = 1$, or equivalently $\tilde{R} = R$. From the corresponding profile likelihood for N given by $L(N) = (N)_x (N+R)^{-z}$ for $N = x, x+1, \ldots$, they showed that $\hat{N}_M = 0$ for $x = 0, z > 0$, $\hat{N}_M = \infty$ for $z = x > 0$, whilst, for $z > x > 0$, $\hat{N} = k$ if and only if $h_{x,k+1} < z < h_{x,k}$, where

$$h_{x,k} = \left\{ \log\left(1 - \frac{x}{k}\right) \right\} \Big/ \log\left(1 - \frac{1}{k+R}\right) \qquad k = x+1, x+2, \ldots ,$$

and $h_{x,x} = \infty$. The same paper also discussed a Petersen-type estimator. This was a modified version of the estimator obtained by scaling up the observed number of distinct target individuals using the inverse of the proportion of planted individuals seen. For the case where plants are present, it was found that, except for short sampling times, this Petersen-type estimator is almost unbiased, but, in comparison to other possible estimators, has a large variance and consequently a large mean square error. We therefore discuss it no further in the present paper.

It was, however, the harmonic mean estimator which Goudie, Pollock and Ashbridge (1998) recommended. If c is a constant in the interval $(0, 1)$, let $n_1 = \inf\{N \geq x : L(N) \geq c\, L(\hat{N}_M)\}$ and $n_2 = \sup\{N \geq x : L(N) \geq c\, L(\hat{N}_M)\}$ denote the lower and upper endpoints respectively of the likelihood interval for N obtained from the profile likelihood $L(N)$ for the case where $K = 1$. The recommended harmonic mean estimator used the likelihood interval for the case $c = 0.5$ and

was given by $\hat{N}_H = \left[0.5 + \left\{2n_1n_2/(n_1 + n_2)\right\}\right]$, where the square brackets again denote the integer part.

5. THE PERFORMANCE OF THE NEW ESTIMATOR

Our primary interest in this paper is how the conditionally unbiased estimator $\hat{N}_U$ performs when $K = 1$, or equivalently, sightings of both planted individuals and members of the target population follow the same Poisson process. Under this assumption we compare in this section the performance of $\hat{N}_U$ with that of the maximum likelihood estimator $\hat{N}_M$ and the harmonic mean estimator $\hat{N}_H$. More precisely, in Tables I and II we take target populations of sizes $N = 25$ and 50 respectively, and consider the cases where the number of planted individuals $R = 10$ or 25. The values of θ we take are 0.11, 0.36, 0.69, 1.2 and 2.3, which correspond approximately to detection probabilities of 0.1, 0.3, 0.5, 0.7 and 0.9 respectively for any particular individual.

In these tables the performance of each estimator is indicated by the statistics B, S and M which are respectively the bias, standard deviation and mean square error conditional on the event $C = \{Z > X\}$. Although this conditioning is not required for the estimators $\hat{N}_H$ and $\hat{N}_U$, as both are finite with probability one, it is, however, necessary for the estimator $\hat{N}_M$ if the statistics B, S and M are to be finite. The conditioning thus enables us to make a precise comparison. The conditional distributions, given the event C, of each of the three estimators can readily be calculated using the joint probability function (5) of X and Z, in the case $\tilde{R} = R$. All the moments of estimators that are given in this paper are conditional on the event C, but, for ease of exposition and to avoid confusion with the conditioning on Z, we will, for instance, use the term 'mean' rather than 'conditional mean' on each occasion.

For the cases considered in Tables I and II, it can be seen that, when $\theta = 0.69$, 1.2 and 2.3, the absolute bias, given the event C, of the estimator $\hat{N}_U$ is always less than that of the other two estimators. Indeed, for these values of θ, the conditionally unbiased estimator is almost unbiased. Moreover, in only one of the twenty cases covered by these tables is the standard deviation of $\hat{N}_U$, given

TABLE I. Conditional biases (B), standard deviations (S), and mean square errors (M) of the three estimators when $N = 25$.

		$R = 10$			$R = 25$		
		$\hat{N}_M$	$\hat{N}_H$	$\hat{N}_U$	$\hat{N}_M$	$\hat{N}_H$	$\hat{N}_U$
$\theta = 0.11$	B	-2.87	-10.87	-11.27	6.09	-4.53	-3.90
	S	17.48	11.87	9.47	28.59	18.32	15.64
	M	313.9	259.1	216.8	854.3	356.3	259.9
$\theta = 0.36$	B	4.54	0.52	-0.61	1.66	-0.61	-0.05
	S	19.99	14.41	11.96	12.94	11.20	10.64
	M	420.3	207.8	143.4	170.3	125.9	113.2
$\theta = 0.69$	B	0.86	0.28	0.00	0.11	-0.34	-0.01
	S	8.86	8.03	7.50	6.68	6.43	6.39
	M	79.3	64.5	56.2	44.6	41.5	40.9
$\theta = 1.2$	B	-0.19	-0.02	0.01	-0.33	-0.25	0.01
	S	4.39	4.27	4.25	3.91	3.85	3.86
	M	19.3	18.2	18.0	15.4	14.9	14.9
$\theta = 2.3$	B	-0.46	-0.42	0.00	-0.47	-0.46	0.00
	S	1.89	1.95	1.88	1.83	1.90	1.81
	M	3.8	4.0	3.5	3.6	3.8	3.3

TABLE II. Conditional biases (B), standard deviations (S), and mean square errors (M) of the three estimators when $N = 50$.

		$R = 10$			$R = 25$		
		$\hat{N}_M$	$\hat{N}_H$	$\hat{N}_U$	$\hat{N}_M$	$\hat{N}_H$	$\hat{N}_U$
$\theta = 0.11$	B	-2.59	-16.46	-20.68	12.98	-4.14	-6.94
	S	31.32	21.12	16.11	49.31	32.37	25.75
	M	987.4	717.0	687.0	2599.8	1064.8	710.9
$\theta = 0.36$	B	7.50	2.53	-0.43	3.12	0.32	-0.02
	S	33.01	25.06	20.78	20.59	18.36	17.20
	M	1146.1	634.2	432.1	433.8	337.3	295.9
$\theta = 0.69$	B	1.25	0.68	-0.01	0.52	-0.06	0.02
	S	12.67	12.01	11.47	10.35	10.07	9.90
	M	162.0	144.8	131.6	107.3	101.3	98.1
$\theta = 1.2$	B	-0.07	0.05	0.01	-0.23	-0.15	0.01
	S	6.42	6.36	6.27	5.87	5.85	5.79
	M	41.2	40.5	39.4	34.5	34.3	33.6
$\theta = 2.3$	B	-0.45	-0.24	0.00	-0.47	-0.31	0.00
	S	2.70	2.72	2.69	2.62	2.62	2.61
	M	7.5	7.5	7.2	7.1	7.0	6.8

the event C, larger than that of either $\hat{N}_M$ or $\hat{N}_H$, and in no case does the corresponding mean square error of $\hat{N}_U$ exceed that of the other two estimators. For $\theta = 0.69$, 1.2 and 2.3, the performance of the conditionally unbiased estimator $\hat{N}_U$ for population sizes in this range is thus evidently better than that of the two earlier estimators.

The situation is less clear-cut for the shorter sampling times, $\theta = 0.11$ and 0.36. For the latter of these, the estimator $\hat{N}_U$ has the smallest absolute bias except for the case $N = 25$ and $R = 10$, when the harmonic mean estimator $\hat{N}_H$ is less biased. For $\theta = 0.11$, the performance of $\hat{N}_U$ appears to be more dependent on the number of plants. When θ takes this value, the estimator $\hat{N}_U$ is the most biased of the three estimators for $R = 10$, but, when $R = 25$, it gives results which are better than those of $\hat{N}_M$ and, in view of the smaller standard deviation of $\hat{N}_U$, broadly comparable to those of $\hat{N}_H$. Goudie, Pollock and Ashbridge (1998), however, argued that a truncated sampling procedure is only likely to be adopted in practice when the experimenter has at least some vague knowledge of the rates at which sightings occur, and that this knowledge could be used to avoid premature termination of sampling. This argument suggests that the relative performance of the estimators for these shorter sampling times is of much less significance than that for the larger values of θ.

Table III gives the corresponding results for the case $R = 0$, in which no plants are used. The relative performances of the three estimators here is similar to that for positive R. For the longest three sampling times, the conditionally unbiased estimator $\hat{N}_U$ provides the smallest values of both of the statistics B and S, and hence also of M. For the shortest two sampling times, $\hat{N}_U$ still gives the smallest values of S, but, as θ gets smaller, its absolute bias increases, leaving it the most biased of the three estimators when $\theta = 0.11$.

6. THE STANDARD DEVIATION OF THE NEW ESTIMATOR

The results of Berg (1974) on factorial series distributions show that, for fixed $z \geq N$, there is a unique unbiased estimator of the variance of $\tilde{N}_U$. In the present context, it appears appropriate to consider the variance estimator

TABLE III. Conditional biases (B), standard deviations (S), and mean square errors (M) of the three estimators when $R = 0$.

		$N = 25$			$N = 50$		
		$\hat{N}_M$	$\hat{N}_H$	$\hat{N}_U$	$\hat{N}_M$	$\hat{N}_H$	$\hat{N}_U$
$\theta = 0.11$	B	-18.02	-17.97	-19.47	-30.96	-33.50	-36.78
	S	5.71	4.73	3.78	13.20	9.80	7.72
	M	357.2	345.1	393.2	1132.6	1218.6	1412.3
$\theta = 0.36$	B	-1.38	-4.33	-7.64	8.29	0.52	-6.70
	S	14.35	10.59	8.21	34.74	24.61	18.44
	M	207.8	130.9	125.7	1275.3	605.9	384.9
$\theta = 0.69$	B	3.04	1.61	-0.80	3.20	2.15	-0.06
	S	14.65	11.23	8.84	19.02	16.31	14.12
	M	223.9	128.7	78.8	372.2	270.6	199.4
$\theta = 1.2$	B	0.29	0.61	-0.01	0.15	0.45	0.00
	S	5.94	5.51	5.09	7.23	7.16	6.91
	M	35.4	30.8	25.9	52.2	51.4	47.7
$\theta = 2.3$	B	-0.44	-0.23	0.00	-0.43	-0.19	0.01
	S	2.02	2.08	2.00	2.79	2.83	2.77
	M	4.3	4.4	4.0	8.0	8.0	8.0

TABLE IV. Comparison of the mean $\hat{S}$ of the estimated standard deviation of the estimator $\hat{N}_U$ with the true value S for the situations considered in the previous tables.

		$N = 25$			$N = 50$		
		$R = 0$	$R = 10$	R =25	$R = 0$	$R = 10$	$R = 25$
$\theta = 0.11$	S	3.78	9.47	15.64	7.72	16.11	25.75
	$\hat{S}$	2.86	10.59	17.12	7.19	19.55	29.31
$\theta = 0.36$	S	8.21	11.96	10.64	18.44	20.78	17.20
	$\hat{S}$	8.56	12.61	10.64	20.89	21.49	17.23
$\theta = 0.69$	S	8.84	7.50	6.39	14.12	11.47	9.90
	$\hat{S}$	9.29	7.53	6.36	14.24	11.47	9.90
$\theta = 1.2$	S	5.09	4.25	3.86	6.91	6.27	5.79
	$\hat{S}$	5.08	4.23	3.85	6.89	6.26	5.79
$\theta = 2.3$	S	2.00	1.88	1.81	2.77	2.69	2.61
	$\hat{S}$	1.97	1.86	1.79	2.76	2.67	2.60

$\hat{V}_U(x,z) = \left\{\hat{N}_U(x,z;1) - x\right\}\left\{\hat{N}_U(x,z;1) - \hat{N}_U(x-1,z;1)\right\}$, which differs from the one in the class given by Berg only in the use of $\hat{N}_U$ rather than $\tilde{N}_U$. Even without this minor modification the estimator $\hat{V}_U(x,z)$ would not be exactly unbiased with respect to the joint distribution of X and Z since the condition that $z \geq N$ is not satisfied with probability one.

Nonetheless the results given in Table IV indicate that $\hat{V}_U(x,z)$ is a useful estimator of variance. In that table we have used $\hat{V}_U(x,z)$ to estimate the standard deviation of the estimator $\hat{N}_U$ under each of the scenarios covered by the previous three tables. In each case the mean $\hat{S}$ for the estimated standard deviation is compared with the true value S for that situation. Table IV shows that, when plants are used, whilst there is a small positive bias in the estimated standard deviation for the shorter sampling times, the correspondence in the case of the longer sampling times is very close. If no plants are used, Table IV suggests that there is a tendency to underestimate the standard deviation when $\theta = 0.11$, but the variance estimator again performs well for the longer sampling times.

Also evident in Table IV, for the case $\theta = 0.11$, is the increase in the true standard deviation S as the number of plants increases. On first sight this appears to indicate a deterioration in the performance of $\hat{N}_U$, but comparison with the earlier tables shows that this increase in S is accompanied by a large reduction in the negative bias. This feature of Table IV is thus simply another example of the tendency in population size problems for estimators with small means to also have small variances.

7. ROBUSTNESS

The results of section 5 indicated that, assuming that the expected numbers of sightings by time τ for the two types of individual are the same, the conditionally unbiased estimator $\hat{N}_U$ should be the preferred estimator, provided that the sampling time is not too short. As elsewhere in plant-capture methodology, however, due recognition must be given to the assumption that the behaviour of the plants is indistinguishable from that of the members of the target population. We therefore now examine whether our conclusions on the merits of

the estimators remain valid when there is less certainty that the detection rates are the same for the target and planted populations.

In this section we present results for the harmonic mean estimator $\hat{N}_H$, which was recommended by Goudie, Pollock and Ashbridge (1998), and for the estimator $\hat{N}_U$. We now evaluate the properties of these estimators, assuming that the joint probability function of X and Z is given by (5) with the constant K allowed to differ from unity. The results shown in Table V are based on a target population of size $N = 50$ and a planted population of size $R = 25$. We allow the rate at which plants are detected to differ by up to 20% from that of the target population, and set $K = 0.8, 0.9, 1.0, 1.1$ and 1.2.

Table V shows that in some respects the behaviour of the two estimators is similar. For the cases shown, the value of the bias B for each estimator is a decreasing function of K, and, for $\theta = 0.11$ and 0.36, the same is true of the mean square error M. Intuitive justifications for these properties are not hard to find. If the numbers of sightings for the two populations are compared in the belief that $K = 1$ when in reality $K < 1$, the proportion of target individuals seen is likely to be greater than the proportion that one would infer by comparison with the plants, leading to an over-estimate of N. We have already noted that in problems of this type it is very often true that estimators with larger means also have larger variances, and therefore, when $K < 1$, a large mean square error is also to be expected. Conversely, when $K > 1$, underestimation of N and a small mean square error is to be expected.

In terms of the standard deviation S and the mean square error M, the results given in Table V reinforce our earlier preference for using $\hat{N}_U$ rather than $\hat{N}_H$. The standard deviation of $\hat{N}_U$ never exceeds that of $\hat{N}_H$, and, except for small values of K when $\theta = 2.3$, the same is true of the mean square error. Nonetheless, when $\theta = 0.11$, although $\hat{N}_U$ yields a substantial reduction in the value of M, it should be noted that the absolute bias of $\hat{N}_H$ is consistently smaller. For larger values of θ, the conditionally unbiased estimator $\hat{N}_U$ has the smaller absolute bias when $K > 1$, but, for $K < 1$, the harmonic mean estimator $\hat{N}_H$ is usually better in this regard.

TABLE V. Conditional biases (B), standard deviations (S), and mean square errors (M) of the estimators $\hat{N}_H$ and $\hat{N}_U$ when $N = 50$ and $R = 25$.

			$K = 0.8$	$K = 0.9$	$K = 1.0$	$K = 1.1$	$K = 1.2$
$\theta = 0.11$	$\hat{N}_H$	B	1.98	-1.20	-4.14	-6.84	-9.32
		S	35.44	33.91	32.37	30.83	29.32
		M	1260.1	1151.3	1064.8	997.5	946.8
	$\hat{N}_U$	B	-2.18	-4.64	-6.94	-9.08	-11.07
		S	27.72	26.74	25.75	24.74	23.74
		M	773.2	736.7	710.9	694.5	686.2
$\theta = 0.36$	$\hat{N}_H$	B	8.13	3.88	0.32	-2.71	-5.31
		S	23.22	20.51	18.36	16.63	15.22
		M	605.1	435.7	337.3	284.0	259.9
	$\hat{N}_U$	B	7.24	3.31	-0.02	-2.88	-5.36
		S	21.12	18.96	17.20	15.75	14.53
		M	498.3	370.6	295.9	256.2	239.8
$\theta = 0.69$	$\hat{N}_H$	B	5.01	2.29	-0.06	-2.11	-3.91
		S	11.70	10.80	10.07	9.44	8.92
		M	162.0	122.0	101.3	93.6	94.8
	$\hat{N}_U$	B	4.99	2.33	0.02	-2.00	-3.78
		S	11.43	10.60	9.90	9.31	8.80
		M	155.6	117.8	98.1	90.7	91.8
$\theta = 1.2$	$\hat{N}_H$	B	2.88	1.27	-0.15	-1.40	-2.52
		S	6.43	6.12	5.85	5.62	5.41
		M	49.7	39.1	34.3	33.6	35.6
	$\hat{N}_U$	B	3.00	1.41	0.01	-1.23	-2.33
		S	6.37	6.06	5.79	5.57	5.37
		M	49.6	38.7	33.6	32.5	34.3
$\theta = 2.3$	$\hat{N}_H$	B	0.90	0.26	-0.31	-0.79	-1.22
		S	2.78	2.70	2.62	2.56	2.52
		M	8.5	7.3	7.0	7.2	7.9
	$\hat{N}_U$	B	1.17	0.55	0.00	-0.48	-0.92
		S	2.75	2.67	2.61	2.56	2.52
		M	8.9	7.5	6.8	6.8	7.2

8. CONCLUDING COMMENT

In section 5, we noted that, in practice, it should often be possible to avoid terminating sampling prematurely. Indeed, a reviewer of this paper has suggested that, in order to obtain reliable results, experiments of this type should be designed with the aim of achieving detection probabilities of at least 0.5 for each individual. This would reinforce our argument that the relative merits of the estimators should be primarily judged on their properties for the longer sampling times. Goudie, Pollock and Ashbridge (1998) indicated that this implied preferring the harmonic mean estimator $\hat{N}_H$ to the maximum likelihood and Petersen estimators. Application of the same argument to the results presented in this paper suggests emphasising that the conditionally unbiased estimator $\hat{N}_U$ yields a further improvement in performance for these longer sampling times, when one is confident that the behaviour of the plants matches that of the target population.

In situations in which the detection rate for the plants is liable to be less than that of members of the target population, the estimator $\hat{N}_H$ has slightly smaller bias for long sampling times, and smaller mean square error when $\theta = 2.3$. On the other hand, even in a well-designed experiment, plants may often be easier to detect than members of the target population. In this case, our results suggest that the conditionally unbiased estimator $\hat{N}_U$ is again to be preferred.

ACKNOWLEDGEMENT

The research work of Jonathan Ashbridge has been supported by the Engineering and Physical Sciences Research Council.

BIBLIOGRAPHY

Berg, S. (1974). "Factorial series distributions, with applications to capture-recapture problems," *Scand. J. Statist.* **1**, 145-152.

Berg, S. (1975). "Some properties and applications of a ratio of Stirling numbers of the second kind," *Scand. J. Statist.* **2**, 91-94.

Charalambides, C.A. (1981). "On a restricted occupancy model and its applications," *Biom. J.* **23**, 601-610.

Craig, C.C. (1953). "On the utilization of marked specimen in estimating populations of flying insects," *Biometrika* **40**, 170-176.

Goudie, I.B.J. (1995). "A plant-capture approach for achieving complete coverage of a population," *Commun. Statist. - Theor. Meth.* **24**, 1293-1305.

Goudie, I.B.J., Pollock, K.H. and Ashbridge, J. (1998). "A plant-capture approach for population size estimation in continuous time," *Commun. Statist. - Theor. Meth.* **27**, 433-451.

Harris, B. (1968). "Statistical inference in the classical occupancy problem: Unbiased estimation of the number of classes," *J. Amer. Statist. Ass.* **63**, 837-847.

Joe, H. and Reid, N. (1985). "Estimating the number of faults in a system," *J. Amer. Statist. Ass.* **80**, 222-226.

Laska, E.M. and Meisner, M. (1993). "A plant-capture method for estimating the size of a population from a single sample," *Biometrics* **49**, 209-220.

Nayak, T.K. (1988). "Estimating population size by recapture sampling," *Biometrika* **75**, 113-120.

Nayak, T.K. (1989). "A note on estimating the number of errors in a system by recapture sampling," *Statist. Prob. Letters* **7**, 191-194.

Yip, P.S.F., Fong, D.Y.T. and Wilson, K. (1993). "Estimating population size by recapture sampling via estimating function," *Stoch. Models* **9**, 179-193.

Yip, P.S.F. (1996). "Effect of plant-capture in a capture-recapture experiment," *Commun. Statist. - Theor. Meth.* **25**, 2025-2038.