

---

## ESTIMATING ABUNDANCE AND DENSITY: ADDITIONAL METHODS

(Version 4, 14 March 2013)

Page

<b>3.1 EXPLOITED POPULATION TECHNIQUES .....</b>	<b>91</b>
<b>3.1.1 Change-in-Ratio Methods .....</b>	<b>91</b>
<b>3.1.2 Planning Change-in-Ratio Studies .....</b>	<b>96</b>
<b>3.1.3 Eberhardt's Removal Method .....</b>	<b>100</b>
<b>3.1.4 Catch-Effort Methods .....</b>	<b>103</b>
<b>3.2 RESIGHT METHODS.....</b>	<b>109</b>
<b>3.3 ENUMERATION METHODS .....</b>	<b>116</b>
<b>3.4 ESTIMATING DENSITY .....</b>	<b>118</b>
<b>3.4.1 Boundary Strip Methods .....</b>	<b>119</b>
<b>3.4.2 Nested Grids Method.....</b>	<b>123</b>
<b>3.4.3 Spatial Methods for Multiple Recaptures .....</b>	<b>125</b>
<b>3.5 SUMMARY.....</b>	<b>130</b>
<b>SELECTED READING .....</b>	<b>131</b>
<b>QUESTIONS AND PROBLEMS .....</b>	<b>132</b>

Several methods have been developed for population estimation in which the organisms need to be captured only one time. The first set of these were developed for exploited populations in which the individuals were removed as a harvest from the population, and they are loosely described as *removal methods*. These methods were first developed in the 1940s for wildlife and fisheries management to get estimates of the population under harvest. A discussion of removal methods forms the first section of this chapter.

A second set of methods are of much more recent development and are based on the principle of resighting animals that have been marked. They require an individual animal to be captured only once, and then all subsequent "recaptures" are

from sighting records only - the individual never has to be physically captured again. These newer methods were developed for animals with radio-collars but can be use for any kind of mark which is visible to a distant observer. We shall discuss these resighting methods in the second section of this chapter.

Mark-recapture methods have developed into very complex statistical methods during the last 10 years. While many of these methods are beyond the scope of this book, we can still use these methods for field populations once we understand their assumptions. In the third section of this chapter I will provide an overview of how we can convert abundance estimates to density estimates.

### **3.1 EXPLOITED POPULATION TECHNIQUES**

A special set of techniques have been developed for estimating population size in exploited populations. Many of these techniques are highly specific for exploited fish populations (Ricker 1975, Seber 1982) but some are of general interest because they can be applied to wildlife and fisheries problems as well as other field situations. I will briefly describe two types of approaches to population estimation that can be used with exploited populations.

#### **3.1.1 Change-in-Ratio Methods**

The idea that population size could be estimated from field data on the change in sex ratio during a hunting season was first noted by Kelker (1940). When only one sex is hunted, and the sex ratio before and after hunting is known, as well as the total kill, Kelker showed that one could calculate a population estimation from the simple ratio:

$$\left. \begin{array}{l} \text{Fraction of males} \\ \text{in population} \\ \text{after hunting} \\ \text{removals} \end{array} \right\} = \frac{\left( \begin{array}{l} \text{Number of males befor} \\ \text{hunting} \end{array} \right) - \left( \begin{array}{l} \text{Number of males} \\ \text{removed} \end{array} \right)}{\left( \begin{array}{l} \text{Total population size} \\ \text{before hunting season} \end{array} \right) - \left( \begin{array}{l} \text{Total number of} \\ \text{animals killed} \\ \text{by hunters} \end{array} \right)}$$

Several investigators discovered and rediscovered this approach during the last 50 years (Hanson 1963), and only recently has the general theory of change-in-ratio

estimators been pulled together (Paulik and Robson 1969, Seber 1982). All change-in-ratio estimators are based on two critical assumptions:

1. The population is composed of *two types* of organisms, such as males and females, or adults and young.
2. A *differential change* in the numbers of the two types of organisms occurs during the observation period.

I will use here the general terminology of Paulik and Robson (1969), calling the two types of organisms *x-types* and *y-types*. To make the situation more concrete, you can think of *x-types* as females and *y-types* as males.

We define the following symbols:

$N_1$  = Total population size at time 1

$N_2$  = Total population size at time 2

$X_1, X_2$  = Number of *x-type* organisms in the population at times 1 and 2

$Y_1, Y_2$  = Number of *y-type* organisms in the population at time 1 and 2

$p_1 = X_1 / N_1$  = Proportion of *x-types* in population at time 1

$p_2 = X_2 / N_2$  = Proportion of *x-types* in population at time 2

$R_x = X_2 - X_1$  = Net change in numbers of *x-type* organisms between times 1 and 2 (may be + or -)

$R_y = Y_2 - Y_1$  = net change in numbers of *y-type* organisms between times 1 and 2 (may be + or -)

$R = R_x + R_y$  = net addition (+) or net removal (-) to the total population between times 1 and 2

Given these symbols, we can re-state the verbal model of Kelker (1940) given above by the formula:

$$\hat{N}_1 = \frac{R_x - \hat{p}_2 R}{\hat{p}_2 - \hat{p}_1} \quad (3.1)$$

where  $p_1$  and  $p_2$  are replaced by their estimated values from the observed data:

$$\hat{p}_i = \frac{x_i}{n_i}$$

where  $n_1$  = Total sample size used to estimate the ratio  $p_1$  at time 1

$n_2$  = Total sample size used to estimate the ratio  $p_2$  at time 2

This is the generalized change-in-ratio estimator of population size. We can illustrate this method with some hypothetical data on ringed-neck pheasants used by Paulik and Robson (1969). During a pre-season survey 800 of 1400 adult birds were females, and after hunting was over 1800 of 2000 birds were females. The total kill was estimated to be 8000 male pheasants and 500 females. Thus:

$$\hat{p}_1 = \frac{800}{1400} = 0.571428$$

$$\hat{p}_2 = \frac{1800}{2000} = 0.900000$$

$$R_x = -500$$

$$R_y = -8000$$

$$R = R_x + R_y = -8500$$

Thus from equation (3.1):

$$\hat{N}_1 = \frac{-500 - (0.900000)(-8500)}{(0.900000) - (0.571428)} = 21,761 \text{ pheasants}$$

Note that because of the structure of this formula it is desirable to keep the estimated proportions to 6 or more decimal places so rounding errors are minimal.

It follows that:

$$\begin{aligned} \hat{X}_1 &= \hat{p}_1 \hat{N}_1 \\ &= (0.571428)(21,761) = 12,435 \text{ female pheasants at time 1} \end{aligned}$$

and

$$\begin{aligned} \hat{N}_2 &= \hat{N}_1 + R \\ &= 21,761 + (-8500) = 13,261 \text{ pheasants alive at time 2} \end{aligned}$$

Confidence intervals for this population estimate can be calculated in two different ways, depending on sample size (Paulik and Robson 1969).

**Large samples:** when approximately 500 or more individuals are sampled to estimate  $p_1$  and an equal large number to estimate  $p_2$ , you should use the normal approximation:

$$\text{Variance}(\hat{N}_1) = \frac{\hat{N}_1^2 (\text{variance}(\hat{p}_1)) + \hat{N}_2^2 (\text{variance}(\hat{p}_2))}{(\hat{p}_1 - \hat{p}_2)^2} \quad (3.2)$$

$$\text{where } \text{Variance}(\hat{p}_1) = \frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} \quad (3.3)$$

$$\text{Variance}(\hat{p}_2) = \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2} \quad (3.4)$$

This variance formula assumes binomial sampling with replacement and is a reasonable approximation to sampling without replacement when less than 10% of the population is sampled (Seber 1982 p. 356). It also assumes that  $R_x$  and  $R_y$  (the removals) are known exactly with no error.

For the pheasant example above:

$$\text{Variance of } \hat{p}_1 = \frac{(0.571428)(1 - 0.571428)}{1400} = 0.0001749$$

$$\text{Variance of } \hat{p}_2 = \frac{(0.9)(1 - 0.9)}{2000} = 0.000045$$

Thus:

$$\begin{aligned} \text{Variance}(\hat{N}_1) &= \frac{(21,761)^2(0.0001749) + (13,261)^2(0.000045)}{(0.900000 - 0.571428)^2} \\ &= 840,466 \end{aligned}$$

The standard error is given by:

$$\begin{aligned} \text{Standard error of } \hat{N}_1 &= \sqrt{\text{Variance}(\hat{N}_1)} \\ &= \sqrt{840,446} = 916.8 \end{aligned}$$

The 95% confidence limits follow from the normal distribution:

$$\hat{N}_1 \pm 1.96[\text{S.E.}(\hat{N}_1)]$$

For these data:

$$21,761 \pm 1.96(916.8)$$

or 19,964 to 23,558 pheasants.

**Small samples:** When less than 100 or 200 individuals are sampled to estimate  $p_1$  and  $p_2$ , you should use an alternative method suggested by Paulik and Robson (1969). This method obtains a confidence interval for the reciprocal of  $\hat{N}_1$  as follows:

$$\text{Variance of } (1/\hat{N}_1) = \frac{(R_x - \hat{p}_1 R)^2}{(R_x - \hat{p}_2 R)^4} \left( \text{Variance of } \hat{p}_2 \right) + \frac{1}{(R_x - p_2 R)^2} \left( \text{Variance of } \hat{p}_1 \right) \quad (3.5)$$

where variance of  $p_1$  = as defined in equation (3.3)

variance of  $p_2$  = as defined in equation (3.4)

and all other terms are as defined above (page 92).

This formula assumes that  $R_x$  and  $R_y$  are known without error. If we apply this formula to the pheasant data used above:

$$\begin{aligned} \text{Variance of } (1/\hat{N}_1) &= \frac{[-8000 - (0.428571)(-8500)]^2}{[-8000 - (0.10)(-8500)]^4} (0.000045) \\ &\quad + \frac{1}{[-8000 - (0.10)(-8500)]^2} (0.0001749) \\ &= 3.7481 \times 10^{-12} \end{aligned}$$

$$\begin{aligned} \text{Standard error of } (1/\hat{N}_1) &= \sqrt{\text{Variance}(1/\hat{N}_1)} \\ &= \sqrt{3.7481 \times 10^{-12}} = 1.9360 \times 10^{-6} \end{aligned}$$

The 95% confidence interval is thus:

$$\begin{aligned} \frac{1}{\hat{N}_1} \pm 1.96 [\text{S.E.}(1/\hat{N}_1)] \\ \frac{1}{21,761} \pm 1.96 (1.9360 \times 10^{-6}) \end{aligned}$$

or  $4.2159 \times 10^{-5}$  to  $4.9748 \times 10^{-5}$

Inverting these limits to get confidence limits for  $\hat{N}_1$ :

$$\text{Lower 95\% confidence limit} = \frac{1}{4.9748 \times 10^{-5}} = 20,101 \text{ pheasants}$$

$$\text{Upper 95\% confidence limit} = \frac{1}{4.2159 \times 10^{-5}} = 23,720 \text{ pheasants}$$

Note that these confidence limits are asymmetrical about  $\hat{N}_1$  and are slightly wider than those calculated above using the (more appropriate) large sample formulas on these data.

The program RECAP (Appendix 2, page 000) can do these calculations for the generalized change-in-ratio estimator of population size.

### 3.1.2 Planning Change-in-Ratio Studies

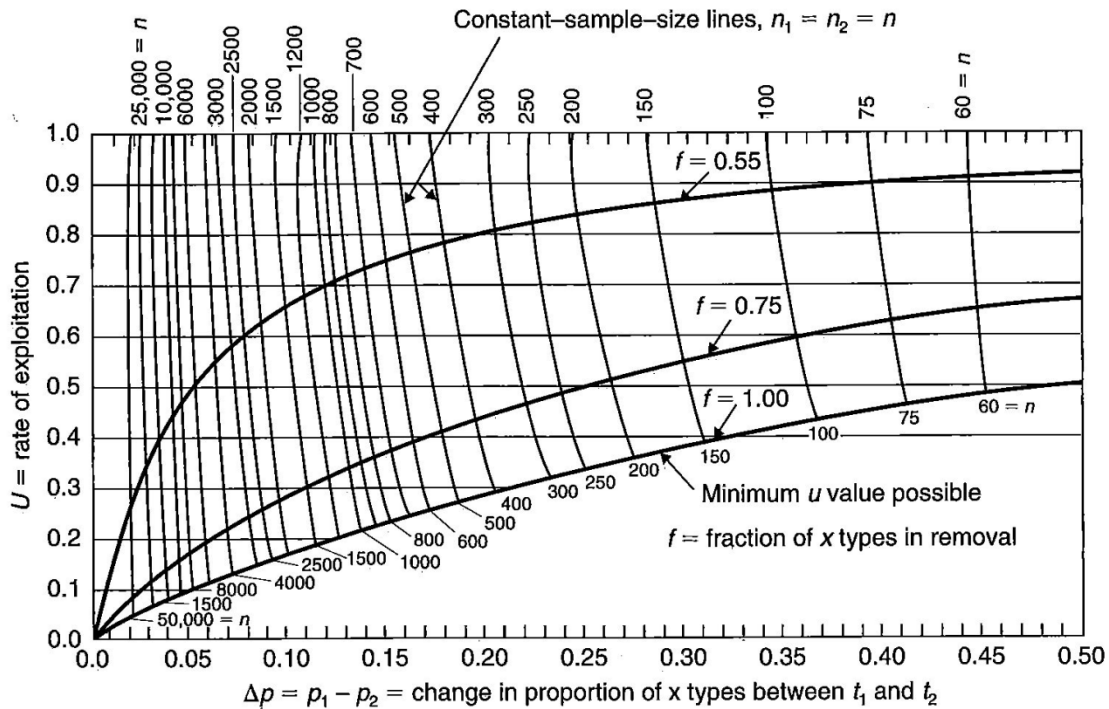
If you propose to use the change-in-ratio estimator (eq. 3.1) to estimate population size, you should use the approach outlined by Paulik and Robson (1969) to help plan your experiment. Five variables must be guessed at to do this planning:

1.  $\Delta p = p_1 - p_2 =$  Expected change in the proportion of x-types during the experiment
2.  $u =$  Rate of exploitation  $= R / N_1$ , which is the fraction of the whole population that is removed
3.  $f = R_x / R =$  fraction of x-types in the removals
4. Acceptable limits of error for the  $\hat{N}_1$  estimate ( $\pm 25\%$  might be usual) see page 000)
5. Probability  $(1 - \alpha)$  of achieving the acceptable limits of error defined in (4); 90% or 95% might be the usual values here.

Figure 3.1 shows the sample sizes required for *each* sample ( $n_1, n_2$ ) for combinations of these parameters with the limits of error set at  $\pm 25\%$  and  $(1 - \alpha)$  as 90%. To use Figure 3.1, proceed as follows:

1. Estimate two of the three variables plotted - the change in proportion of x-types, the rate of exploitation, and the fraction of x-types in the removal.
2. From these two variables, locate your position on the graph in Figure 3.1, and read the nearest contour line to get the sample size required for each sample.

For example, if 60% of the whole population will be harvested and you expect 75% of the harvested animals to be x-types, Figure 3.1 shows that sample size should be about 100 for the first sample and 100 for the second sample (to achieve error limits of  $\pm 25\%$  with a  $1 - \alpha$  of 0.90).



**Figure 3.1** Sample sizes required for change-in-ratio estimation of population size with acceptable error limits of  $\pm 25\%$  with  $(1 - \alpha)$  of 0.90 when the initial proportion of x-types is 0.50. The contour lines represent the sample sizes needed in each of the time 1 and time 2 samples. The required sample sizes are only slightly affected by changes in the rate of exploitation ( $u$ ) but are greatly affected by changes in  $\Delta p$ . (Source: Paulik and Robson, 1969).

Two general points can be noted by inspecting Figure 3.1, and we can formulate them as rules of guidance:

1.  $\Delta p$ , the expected change in proportions, is the critical variable affecting the required sample sizes of change-in-ratio experiments. The rate of exploitation ( $u$ ) is of minor importance, as is the fraction of x-types in the removals ( $f$ ).
2. For  $\Delta p$  less than 0.05, this method requires enormous sample sizes and in practice is not useful for field studies. If  $p$  is less than 0.10, large sample sizes



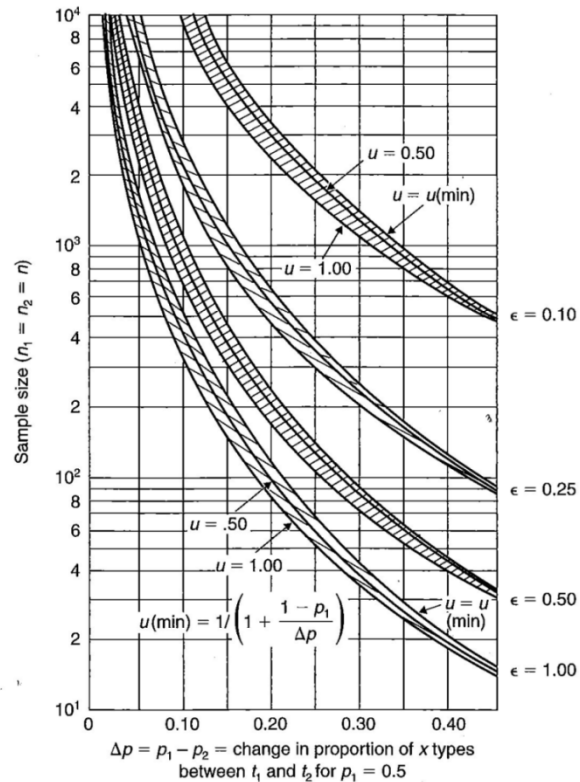
are needed and it is especially critical to test all the assumptions of this approach.

Using a series of graphs like Figure 3.1, Paulik and Robson (1969) synthesized Figure 3.2 from which one can read directly the required sample sizes for change-in-ratio experiments. We can illustrate the utility of this graph with an example.

Suppose you are planning a change-in-ratio experiment for a deer population and you expect a change in the sex-ratio by at least 0.30 over the hunting season. If you desire to estimate population size with an accuracy of  $\pm 25\%$ , from Figure 3.2 you can estimate the necessary sample size to be approximately 200-250 deer to be sexed in *both* the before-hunting sample *and* the after-hunting sample. If you want  $\pm 10\%$  accuracy, you must pay for it by increasing the sample size to 1050-1400 deer in each sample. Clearly, if your budget will allow you to count and sex only 40 deer in each sample, you will get an estimate of  $N_1$  accurate only to  $\pm 100\%$ . If this is inadequate, you need to reformulate your research goals.

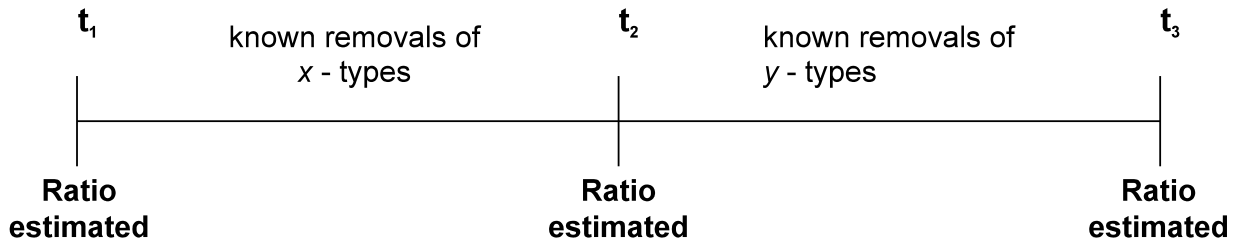
Figures 3.1 and 3.2 are similar in general approach to Figures 2.3 and 2.4 (page 000) which allow one to estimate samples required in a Petersen-type mark-recapture experiment. Paulik and Robson (1969) show that Petersen-type marking studies can be viewed as a special case of a change-in-ratio estimator in which  $x$ -type animals are *marked* and  $y$ -type *unmarked*.

All the above methods for change-in-ratio estimation are designed for *closed* populations and are based on the assumption that all individuals have an equal chance of being sampled in both the first and in the second samples. Seber (1982, Chapter 9) discusses more complex situations for open populations and techniques for testing the assumptions of this method.



**Figure 3.2** Sample sizes required for a change-in-ratio estimate of population size at several possible acceptable error limits ( $\epsilon$  as a proportion of the population size) with  $(1 - \alpha)$  of 0.95. The initial proportion of x-types is assumed to be 0.50. Error limits ( $\epsilon$ ) ranging from  $\pm 10\%$  of the population size up to  $\pm 100\%$  are indicated on the right side of the graph. These estimates are affected only very slightly by changes in the initial proportion of x-types between 0.05 and 0.50, and the critical variable is  $\Delta p$  (X-axis). The upper limits of the shaded zones are for situations with minimal exploitation, and the lower limits of the shaded zones are for situations with maximal exploitation. By reading sample sizes off the top of the curves, one can operate conservatively. (Source: Paulik and Robson, 1969).

If the x-types and the y-types of animals do not have the same probability of being counted, one of the critical assumptions of the change-in-ratio method is violated. Pollock *et al.* (1985) have proposed a new procedure for estimating population size when there is unequal catchability. This procedure operates by having two periods of removal in which only x-types are removed in the first period and only y-types in the second period. This occurs in some states, for example, where there is a season for male deer followed by a season for females. The procedure can be diagrammed as follows:



Given that you can estimate the proportion of x-types in the population at these three times, Pollock *et al.* (1985) show how to estimate the initial population size. One advantage of this procedure is that the proportion of x-types in the population need not change from time 1 to time 3 (it must obviously change at time 2 if the method is to work!). But the main advantage of the Pollock design is that it is robust to unequal catchability of the two types of animals. It clearly involves more work however to obtain an estimate of population size.

### 3.1.3 Eberhardt's Removal Method

A simpler use of removal data to estimate population size was suggested by Eberhardt (1982) and has been called the "index-removal" method. If an index of population size (like roadside counts) can be made before and after the removal of a known number of individuals, it is possible to use the indices to estimate absolute density. This method does not require you to classify individuals into x-types and y-types as does the change-in-ratio method, and there is no need to identify individuals. Eberhardt (1982) discusses how this removal method compares with methods based on mark-recapture.

To use Eberhardt's method, an index of population size is measured before and after the removals. This index is assumed to have some constant but unknown relationship to population size. For example, a roadside count may see 15% of the deer in an area. If you know the number of animals removed from the population after the first index is taken, you can calculate an estimate of population size by the equation:

$$\hat{N} = \frac{x_1 R}{x_1 - x_2} \quad (3.6)$$

where

$\hat{N}$  = Estimated population size at time 1  
 $x_1$  = Index count at time 1  
 $x_2$  = Index count at time 2  
 $R$  = Number of animals removed

The proportion of animals removed can be estimated from the ratio:

$$\hat{p} = \frac{(x_1 - x_2)}{x_1} \quad (3.7)$$

where  $\hat{p}$  = Estimated proportion of animals removed

Eberhardt (1982) derives a variance estimate for population size as:

$$s_{\hat{N}}^2 = \left( \frac{1 - \hat{p}}{\hat{p}} \right)^2 \left( \frac{1}{x_1} + \frac{1}{x_2} \right) (\hat{N}^2) \quad (3.8)$$

where  $s_{\hat{N}}^2$  = Variance of estimated population size

and the other terms are defined above. From this variance you can construct the 95% confidence limits in the usual manner:

$$\hat{N} \pm 1.96[\text{S.E.}(\hat{N})]$$

where  $\text{S.E.}(\hat{N}) = \sqrt{s_{\hat{N}}^2}$  as defined in equation (3.8).

As with mark-recapture estimates, Eberhardt's removal method works best when a high fraction of the population is seen and a high fraction is removed. Table 3.1 gives the expected coefficient of variation of population estimates made by Eberhardt's method. Unless the percentage of the population seen is above about 40% and the percentage removed above 20%, the method is not very precise.

**Table 3.1** COEFFICIENTS OF VARIATION OF EBERHARDT'S INDEX REMOVAL POPULATION ESTIMATE FOR VARIOUS VALUES OF THE PROPORTION OF THE POPULATION COUNTED AND THE PROPORTION REMOVED<sup>a</sup>.

Percentage removed	Proportion seen			
	0.2	0.4	0.6	0.8

10%	0.92	0.65	0.53	0.46
20%	0.42	0.30	0.25	0.21
30%	0.26	0.18	0.15	0.13
40%	0.17	0.12	0.10	0.09
50%	0.12	0.09	0.07	0.06
60%	0.09	0.06	0.05	0.04
70%	0.06	0.05	0.04	0.03
80%	0.04	0.03	0.03	0.02

<sup>a</sup> Source: Eberhardt (1982)

Box 3.1 illustrates the use of Eberhardt's index removal method.

### Box 3.1 EBERHARDT'S INDEX-REMOVAL METHOD OF POPULATION ESTIMATION

Feral horses were counted before and after a removal program in Oregon with the following results:

$$x_1 = 301 \text{ horses counted before removals}$$

$$x_2 = 76 \text{ horses counted after removals}$$

$$R = 357 \text{ horses removed}$$

Assuming that the visual counts are in direct proportion to population size, we can use equation (3.6) to estimate population size at time 1:

$$\hat{N} = \frac{x_1 R}{x_1 - x_2} = \frac{(301)(357)}{301 - 76} = 477.6 \text{ horses}$$

In this example an estimated 63% of the total population were counted in the first index count (301/477.6). The percentage removed is also high (75%). From Table 3.1 the standard error of this population estimate should be about 3-4% of the estimate, so the confidence interval should be approximately  $\pm 6-8\%$  of the population size.

The variance of this estimate is, from equation (3.7):

$$\begin{aligned} s_{\hat{N}}^2 &= \left( \frac{1 - \hat{p}}{\hat{p}} \right)^2 \left( \frac{1}{x_1} + \frac{1}{x_2} \right) (\hat{N}^2) \\ &= \left( \frac{1 - 0.74686}{0.74686} \right)^2 \left( \frac{1}{301} + \frac{1}{76} \right) (477.6)^2 \\ &= 432.566 \end{aligned}$$

From this variance we obtain the 95% confidence limits as:

$$\hat{N} \pm 1.96[\text{S.E.}(\hat{N})]$$

$$477.6 \pm 1.96 \left[ \sqrt{432.566} \right]$$

or 437 to 518

These confidence limits are  $\pm 8.5\%$  of population size.

Eberhardt's removal method is most useful in management situations in which a controlled removal of animals is undertaken and it is not feasible or economic to mark individuals. It must be feasible to count a significant fraction of the whole population to obtain results that have adequate precision.

### 3.1.4 Catch-Effort Methods

In exploited populations it may be possible to estimate population size by the decline in catch-per-unit-effort with time. This possibility was first recognized by Leslie and Davis (1939) who used it to estimate the size of a rat population that was being exterminated by trapping. DeLury (1947) and Ricker (1975) discuss the method in more detail. This method is highly restricted in its use because *it will work only if a large enough fraction of the population is removed so that there is a decline in the catch per unit effort*. It will not work if the population is large relative to the removals. The following assumptions are also critical for this method:

1. The population is closed
2. Probability of each individual being caught in a trap is constant throughout the experiment
3. All individuals have the same probability of being caught in sample  $i$

The data required for catch-effort models are as follows:

$c_i$  = Catch or number of individuals removed at sample time  $i$

$K_i$  = Accumulated catch from the start up to the beginning of sample time  $i$

$f_i$  = Amount of trapping effort expended in sample time  $i$

$F_i$  = accumulated amount of trapping effort from the start up to the

Table 3.2 gives an example of these type of data for a fishery operating on blue crabs.

**TABLE 3.2** CATCH-EFFORT DATA FOR A POPULATION OF MALE BLUE CRABS (*CALLINECTES SAPIDUS*) FOR A 12-WEEK PERIOD<sup>A</sup>

Week no., $i$	Catch (pounds), $c_i$	Effort (lines per day), $f_i$	Catch per unit effort, $c_i/f_i = Y_i$	Accumulated catch, $K_i$	Accumulated effort, $F_i$
1	33,541	194	172.9	0	0
2	47,326	248	190.8	33,541	194
3	36,460	243	150.0	80,867	442
4	33,157	301	110.2	117,327	685
5	29,207	357	81.8	150,484	986
6	33,125	352	94.1	179,691	1343
7	14,191	269	52.8	212,816	1695
8	9,503	244	38.9	227,007	1964
9	13,115	256	51.2	236,510	2208
10	13,663	248	55.1	249,625	2464
11	10,865	234	46.4	263,288	2712
12	9,887	227	43.6	274,153	2946

<sup>a</sup> Data from Fischler, 1965.

Under the assumptions listed above, the catch-per-unit-effort is directly proportional to the existing population size. Leslie and Davis (1939) showed that, because the population must be declining from time to time by an amount equal to the catch, a regression plot of:

X-axis      accumulated catch ( $K_i$ )

Y-axis      catch-per-unit-effort ( $c_i / f_i$ )

should be a straight line. Figure 3.3 illustrates this plot for the blue crab data in Table 3.1. This graph is easy to grasp because the x-intercept (that point where  $y=0$ , or the catch-per-unit-effort falls to zero) is the initial population size ( $N$ ), since it represents the exhaustion of the catch. Secondly, the slope of the line is an estimate of the

*catchability* of the individuals, the probability that a given individual will be caught with one unit of effort. With this regression (Fig. 3.3) we can estimate these parameters by eye, or to be more precise, use linear regression techniques as follows:

$$\text{Catchability} = \hat{C} = \frac{-\sum_{i=1}^s Y_i(K_i - \bar{K})}{\sum_{i=1}^s (K_i - \bar{K})^2} \quad (3.9)$$

$$\text{Population size} = \hat{N} = \bar{K} + \left( \frac{\bar{Y}}{\hat{C}} \right) \quad (3.10)$$

where  $Y_i = \text{Catch-per-unit-effort} = c_i / f_i$   
 $\bar{K} = \text{Mean value of } K_i \text{ (accumulated catch)} = (\sum K_i) / s$   
 $s = \text{Total number of samples } (i = 1, 2, 3 \dots s)$

For example, using the data in Table 3.1:

$$\bar{Y} = \frac{\sum Y_i}{s} = \frac{172.9 + 190.8 + \dots}{12} = 90.65$$

$$\bar{K} = \frac{\sum K_i}{s} = \frac{0 + 33,541 + 80,867 + \dots}{12} = 168,775.75$$

Thus:

$$\hat{C} = \frac{-[172.9(0 - 168,775.75) + 190.8(33,541 - 168,775.75) + \dots]}{[(0 - 168,775.75)^2 + (33,541 - 168,775.75)^2 + \dots]}$$

$$= 0.0005614$$

$$\hat{N} = 168,775.75 + \left[ \frac{90.65}{(0.0005614)} \right] = 330,268 \text{ pounds}$$

The variance of this population estimate is given by:

$$\text{Variance of } (\hat{N}) = \frac{s_{yx}}{C^2} \left[ \frac{1}{s} + \frac{(\hat{N} - \bar{K})^2}{\sum (K_i - \bar{K})^2} \right] \quad (3.11)$$

where  $s_{yx} = \text{Variance about regression} = \sum [Y_i - \hat{C}(\hat{N} - K_i)]^2 / (s - 2)$  as defined in  
 $s = \text{Number of samples}$



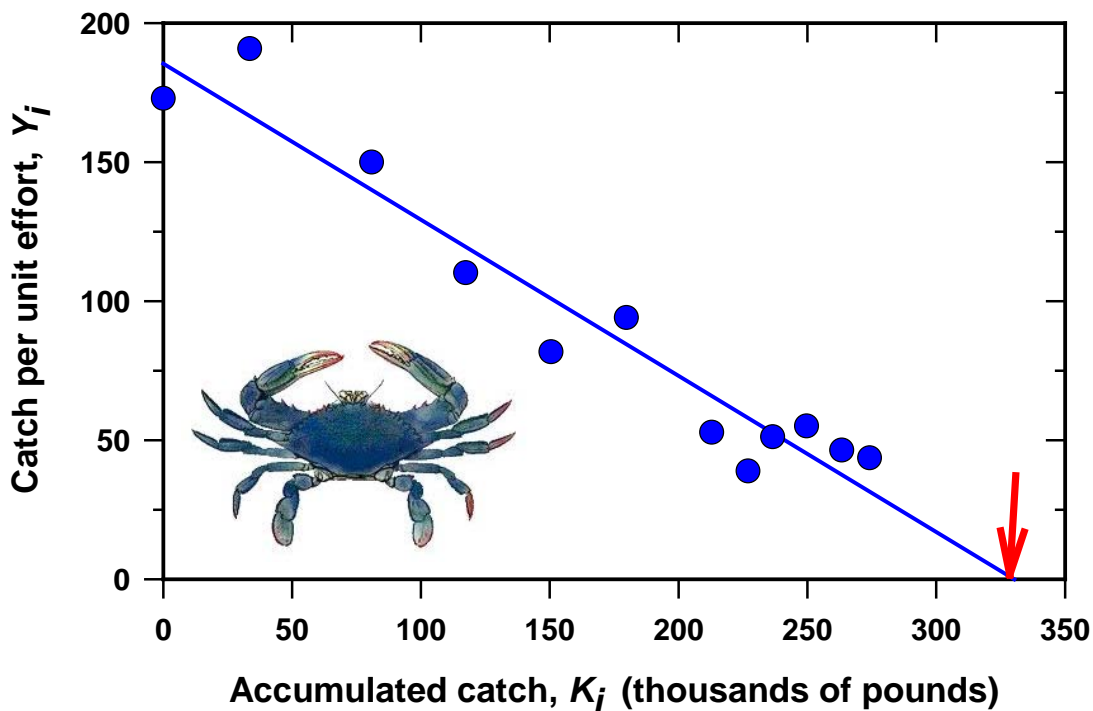
When the number of samples is large ( $s > 10$ ), approximate 95% confidence limits are obtained in the usual way:

$$\text{Standard error of } \hat{N} = \sqrt{\text{Variance of } \hat{N}}$$

$$95\% \text{ confidence limits} = \hat{N} \pm 1.96[\text{S.E.}(\hat{N})]$$

When the number of samples is small, use the general method outlined in Seber (1982, pg. 299) to get confidence limits on  $\hat{N}$ .

The plot shown in Figure 3.3 provides a rough visual check on the assumptions of this model. If the data do not seem to fit a straight line, or if the variance about the regression line is not constant, the data violate some or all of the assumptions and this model is not appropriate. Ricker (1975) discusses how to deal with certain cases in which the assumptions of the model are not fulfilled.



**Figure 3.3** Leslie plot of catch-effort data of Fischler (1965) for a population of male blue crabs. Using the Leslie model, one can estimate initial population size ( $N$ ) by extrapolating the linear regression to the  $X$ -axis. In this case (arrow)  $\hat{N}$  is about  $330 \times 10^3$  pounds. (Original data in Table 3.1.)

DeLury (1947) and Ricker (1975) provide two alternative models for analyzing catch-effort data based on a semi-logarithmic relationship between the log of the catch-per-unit-effort (y-axis) and the accumulated fishing effort ( $F_i$ ) on the x-axis (e.g. Fig. 3.4). The calculations are again based on linear regression techniques. For the more general Ricker model: defining  $z_i = \log(Y_i)$  we have -

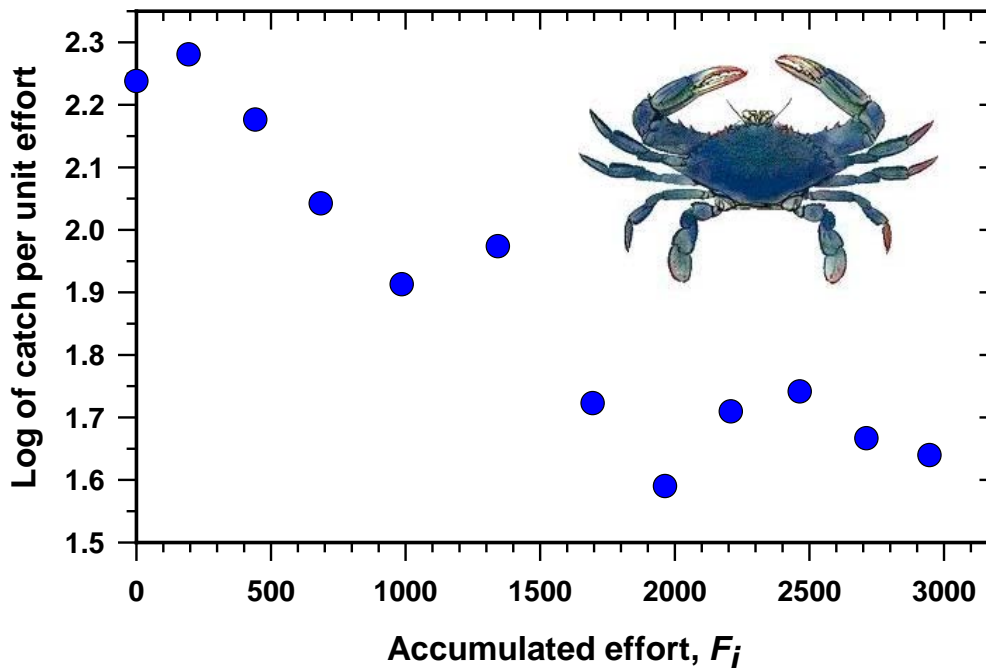
$$\log(1-\hat{C}) = \frac{\sum_{i=1}^s z_i (F_i - \bar{F})}{\sum_{i=1}^s (F_i - \bar{F})^2} \quad (3.12)$$

and

$$\log \hat{N} = \bar{z} - \bar{F} \log(1-\hat{C}) - \log \hat{C} \quad (3.13)$$

and parameter estimates of catchability ( $\hat{C}$ ) and population size ( $\hat{N}$ ) are obtained by taking antilogs. Seber (1982, p. 302) and Ricker (1975) show how confidence intervals may be calculated for these semi-logarithmic models.

Since the Leslie model (eq. 3.10) and the Ricker model (eq. 3.13) are based on somewhat alternative approaches, it is desirable to plot both the Leslie regression (e.g. Fig. 3.3) and the Ricker semi-log regression (Fig. 3.4) as checks on whether the underlying assumptions may be violated.



**Figure 3.4** Ricker plot of catch-effort data of Fischler (1965) for male blue crabs. This is a semilogarithmic plot of the data in Table 3.2 in which the y-axis is the log of the catch per unit effort. The graph shows a curved line and not a straight line, and consequently the Ricker catch-effort model is not supported for these data

The program RECAP in Appendix 2 (page 000) can do these calculations for the catch-effort models of Leslie and Ricker described in this section and calculates confidence intervals for population size for both estimators.

Population estimation by the removal method is subject to many pitfalls. Braaten (1969) showed by computer simulation that the DeLury (1947) semi-logarithmic model produces population estimates that are biased by being too low on average. This bias can be eliminated by regressing the log of the catch-per-unit-effort against the accumulated fishing effort ( $F_i$ ) plus half the effort expended in the  $i$ -interval ( $\frac{1}{2} f_i$ ) (Braaten 1969). The more serious problem of variable catchability has been discussed by Schnute (1983) who proposes a new method of population estimation for the removal method based on maximum-likelihood models. In particular, Schnute (1983) provides a method of testing for constant catchability and

for fitting an alternate model in which catchability is higher in the first sample, and lower in all subsequent samplings. Otis *et al.* (1978) also discuss in detail the problem of removal estimation in which the probability of capture varies with time. The critical assumption of constant catchability should always be examined when you use catch-effort methods to estimate population size.

### **3.2 RESIGHT METHODS**

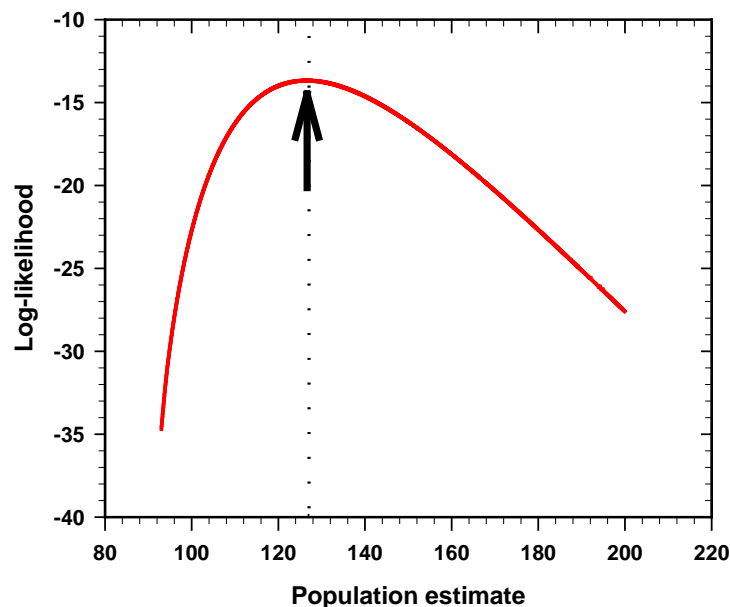
In recent work on vertebrate population, individuals are often marked with radio transmitters, and for these types of studies a range of new population estimators are available (Arnason *et al.* 1991, Neal *et al.* 1993, White 2008). Radio-tagged individuals are often expensive to capture and tag, but after they are released, they can be easily re-sighted using radio-telemetry equipment (Kenward 1987, White and Garrott 1990). The methods that have developed are an important extension of Petersen-type population estimators for closed populations. Resight methods do not require radio-telemetry, of course, and any marking method that allows one to identify an individual at a distance could be used with this approach. Animals marked with color tags or neck bands can be sighted and recorded without capture, and consequently resight methods have a wide applicability.

Mark-resight methods are of considerable utility because they can often be considerably cheaper than conventional mark-recapture with traps. They are also important because they reduce the stress associated with live trap captures. Animals with individual marks can be observed at a distance with minimal disturbance (McClintock 2010<sup>{#15989}</sup>). As such these methods are analogous to the standard mark-recapture methods we discussed in Chapter 2. Resight methods can be applied to open or closed populations and these data can be used in the robust design (page 77). I discuss here two of the simpler models that apply to a closed population. Resight methods can also be applied to open populations but these methods are not covered here (see Cooch and White 2010, page 646).

Four estimators of population size are available for mark-resight data, and White (2008) has implemented these estimators in Program MARK. I will discuss here two of these four estimators, the<sup>{#15417}</sup> joint hypergeometric maximum

likelihood estimator (JHE) and Bowden's estimator. The statistical theory behind the JHE estimator is beyond the scope of this book, and is discussed in detail in Cooch and White (2010). I would like to provide a general understanding of how maximum likelihood estimators are derived since they form the core of much of modern estimation theory for mark-recapture estimates.

Maximum likelihood is a method of statistical inference in which a particular model is evaluated with reference to a set of data. In this case we have a set of observations based on the sighting frequencies of marked and unmarked animals. If we assume that re-sightings occur at random, we can connect through probability theory the observed data and the likely value of population size  $N$ . For simple situations, such as the Petersen method, in which only two sampling times are involved the resulting estimator for  $N$  will be a simple equation like equation (2.2). For more complex situations, involving several sampling times, the resulting estimator cannot be written as a simple equation with a solution and one must search to find the best value of  $N$  by trial and error. In mathematical jargon one searches for the maximum value of a function called the likelihood function (Edwards 1972{ #4877}). Figure 3.5 illustrates the maximum likelihood approach for equation (3.14) and the data used in Box 3.2.



**Figure 3.5** Likelihood ratio for the mountain sheep example given in Box 3.2. Given the observed data, estimated population sizes are substituted in equation (3.10) and the likelihood calculated. The most probable value of  $N$  is that which gives the highest likelihood (arrow), in this case  $\hat{N} = 127$  (dashed line and arrow) at which point the log likelihood is -13.6746 (likelihood =  $1.1513 \times 10^{-6}$ ). Because likelihoods cover a broad numerical range it is easier to graph them as log-likelihoods ( $\log_e$  or  $\ln$ ) as we have done here.

For the JHE estimator used by White (1996{ #4206}), we search for the value of  $N$  which maximizes the following function:

$$L(N \mid M, n_i, m_i) = \prod_{i=1}^k \frac{\binom{M}{m_i} \binom{N-M}{n_i - m_i}}{\binom{N}{n_i}} \quad (3.14)$$

where  $L(N \mid M, n_i, m_i)$  = likelihood of  $N$  conditional on the observed values of  $M$ ,  $n_i$  and  $m_i$

$N$  = population size

$M$  = number of marked animals in the study zone

$n_i$  = total number of animals seen in the  $i$ -th sample survey ( $i = 1, 2, 3, \dots, k$ )

$m_i$  = number of marked animals sighted in the  $i$ -th survey

and the operator  $\binom{M}{m_i}$  indicates the number of possible combinations of  $M$

items taken  $m_i$  at a time =  $\frac{M!}{m_i! (M - m_i)!}$

Equations of this type for many maximum likelihood estimators can be solved only by trial-and-error, by substituting values of  $N$  into the equation and finding the value of  $N$  that maximizes the likelihood. For this reason all these methods are computer-intensive and were not feasible before the advent of modern computers.

JHE estimators can be viewed as a weighted overall average of a series of Petersen estimates for each sampling period. Box 3.2 illustrates the application of the JHE method to sample data on radio-collared mountain sheep.

Confidence intervals for the JHE estimator can be determined by the profile likelihood method described by Venzon and Moolgavkar (1988){#4816}. These

methods are computer intensive and will not be described here. They are available in Program MARK.

### Box 3.2 RESIGHT METHOD OF POPULATION ESTIMATION

Neal *et al.* (1993) report the following data from radio-collared bighorn sheep (*Ovis canadensis*) from Trickle Mountain, Colorado from 14 January to 1 February 1989:

Date	Number of marked sheep $M_i$	Number animals seen during survey $n_i$	Marked animals seen $m_i$
January 14	25	40	9
January 17	25	63	11
January 18	25	66	14
January 19	25	57	11
January 27	25	52	10
January 28	25	61	9
February 1	25	87	19

There were a total of 426 recaptures in the sampling period of which 83 sheep had radio-collars and 343 sheep were not marked. We can use each of these daily samples as Petersen estimates and obtain a series of Petersen estimates from equation (2.2). For example, from January 19 with  $M = 25$ ,  $C = 57$ , and  $R = 11$  using equation (2.3) for sampling with replacement we obtain a Petersen estimate as follows:

$$\hat{N} = \frac{M(C+1)}{(R+1)} = \frac{25(57+1)}{(11+1)} = 120.8 \text{ sheep}$$

To use the JHE estimator, we compute the likelihood of a range of possible values of  $N$  and find the one that has the maximum likelihood. This is tedious and I will illustrate only one calculation. Use as a preliminary estimate  $\hat{N} = 95$ . From equation (3.14):

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

Since  $M = 25$ , we have seven terms to evaluate as follows:

$$L(95 | 25, n_i, m_i) = \frac{\binom{25}{9} \binom{95-25}{40-9}}{\binom{95}{40}} \cdot \frac{\binom{25}{11} \binom{95-25}{63-11}}{\binom{95}{63}} \cdot \frac{\binom{25}{14} \binom{95-25}{66-14}}{\binom{95}{66}} \cdot \dots$$

$$= (0.14632) (0.00522) (0.04817) \dots = 1.2318 \times 10^{-13}$$

To avoid these very small numbers, we can express these likelihoods as log-likelihoods:

$$\log_e (1.2318 \times 10^{-13}) = -29.7251$$

Repeating these calculations for a range of values of N from 96 to 200 gives us the data plotted in Figure 3.5. The maximum likelihood occurs at  $\hat{N} = 127$ , which is the best estimate of population size given these resighting data. The 95% confidence interval for this estimate from Program MARK is 112 to 146.

For the Bowden estimator, we require the sighting frequencies of each of the 25 mountain sheep that had radio-transmitters. For these 25 animals these values ( $f_i$ ) were: 6, 4, 3, 6, 5, 3, 0, 5, 4, 6, 1, 0, 0, 5, 6, 0, 6, 1, 5, 5, 2, 2, 0, 4, and 4. From the usual statistical calculations and the variance equation above:

$$\bar{f} = 3.320 \text{ resightings per animal}$$

$$\text{VAR}(f_i) = 4.8576$$

From equation (3.15) we obtain for these data:

$$\hat{N} = \frac{\left( \frac{(u_T + m_T)}{\bar{f}} \right) + \left( \frac{s_f^2}{\bar{f}^2} \right)}{\left[ 1 + \left( \frac{s_f^2}{T \bar{f}^2} \right) \right]} = \frac{\left( \frac{343 + 83}{3.320} \right) + \left( \frac{4.8576}{(3.320)^2} \right)}{\left[ 1 + \left( \frac{4.8576}{25 (3.320)^2} \right) \right]} = 126.52$$

with variance from equation (3.16):

$$(\text{Var}(\hat{N})) = \frac{\hat{N}^2 \left( \frac{1}{T} - \frac{1}{\hat{N}} \right) \left( \frac{s_f^2}{\bar{f}^2} \right)}{\left( 1 + \frac{s_f^2}{T \bar{f}^2} \right)^2} = \frac{126.52^2 \left( \frac{1}{25} - \frac{1}{126.52} \right) \left( \frac{4.8576}{(3.320)^2} \right)}{\left( 1 + \frac{4.8576}{25 (3.320)^2} \right)^2} = 218.658$$

The 95% confidence limits from equations (3.17) and (3.18) are, with  $t_{\alpha} = 2.064$  for 24 degrees of freedom:



$$\text{Lower confidence limit} = \frac{\hat{N}}{\exp\left(t_{\alpha} \text{CV}(\hat{N})\right)} = \frac{126.52}{e^{\left(2.064\left[\frac{\sqrt{218.658}}{126.52}\right]\right)}} = 99$$

$$\text{Upper confidence limit of } \hat{N} = \left(\hat{N}\right)\left(\exp\left(t_{\alpha} \text{CV}(\hat{N})\right)\right) = (126.52)e^{\left(2.064\left[\frac{\sqrt{218.658}}{126.52}\right]\right)} \\ = 161$$

Note that these confidence limits are slightly wider than those given for the JHE estimator above. This is the price one pays for the less restrictive assumption that individual sheep may have different probabilities of being sighted on the study area.

These calculations are done by Program RECAP (page 000) and by Program MARK from White (2008).

The JHE method assumes that all the marked animals remain on the study area during the surveys. Thus the number of marked animals is constant, and the probability of re-sighting must be constant from sample to sample to obtain a valid closed population estimate. Day to day variation in sightability can not be accommodated in this model but can be allowed in the next model (McClintock 2010).

Bowden (1995) developed an estimate for population size with mark-resight data that allows us to relax the assumption that all individuals in the population have the same probability of resighting. To compute this estimator, we need to have the resighting frequency for each individually-tagged individual in the population. Some animals may not be seen at all, and some may be seen at every one of the sampling times. The Bowden estimator of population size is given by:

$$\hat{N} = \frac{\left(\frac{(u_T + m_T)}{\bar{f}}\right) + \left(\frac{s_f^2}{\bar{f}^2}\right)}{\left[1 + \left(\frac{s_f^2}{T\bar{f}^2}\right)\right]} \quad (3.15)$$

where  $u_T$  = total number of sightings of unmarked animals over all periods  
 $m_T$  = total number of sightings of marked animals over all time periods  
 $\bar{f}$  = mean resighting frequency of marked animals =  $m_T / T$

$T = M =$  number of marked animals in the population at the time of the surveys

$$s_f^2 = \text{variance of sighting frequencies of marked animals} = \frac{\sum_{j=1}^T (f_j - \bar{f})^2}{T}$$

where  $f_j = \frac{\text{number of sampling times in which individual } j \text{ was sighted}}{\text{total number of sampling times}}$

Note that the variance estimate for sighting frequencies has  $T$  in the denominator (and not  $T-1$ ) because it is a complete sample of all the population of tagged animals.

This estimate (3.15) has variance:

$$\hat{\text{Var}}(\hat{N}) = \frac{\hat{N}^2 \left( \frac{1}{T} - \frac{1}{\hat{N}} \right) \left( \frac{s_f^2}{\bar{f}^2} \right)}{\left( 1 + \frac{s_f^2}{T \bar{f}^2} \right)^2} \quad (3.16)$$

Confidence intervals for the Bowden estimator are obtained from a log-transformation as

$$\text{Lower confidence limit} = \frac{\hat{N}}{\exp\left(t_\alpha \text{CV}(\hat{N})\right)} \quad (3.17)$$

$$\text{Upper confidence limit of } \hat{N} = \left( \hat{N} \left( \exp\left(t_\alpha \text{CV}(\hat{N})\right) \right) \right) \quad (3.18)$$

where  $t_\alpha = t$ -value with  $(T-1)$  degrees of freedom for the specified confidence level

$$\text{and } \text{CV}(\hat{N}) = \text{coefficient of variation of } (\hat{N}) = \frac{\sqrt{\text{Var}(\hat{N})}}{\hat{N}} \quad (3.19)$$

The Bowden estimator usually gives somewhat wider confidence limits for the population estimate because it does not make the restrictive assumption that all individuals have equal sightability. It can be computed with Program RECAP (Appendix 2, page 000) or in Program MARK from White (2008).

Resight methods can be also developed for open populations by the use of the robust design (Figure 2.7, page 77). McClintock (2010) describes the details of how to carry out this approach using Program MARK. In addition if the marked animals have radio-collars, it may be possible to determine which individuals are outside of the study area either permanently or temporarily. These data would allow emigration and immigration estimates to be determined. Details are given in McClintock (2010). The movements of radio-collared animals are not relevant for obtaining estimates of abundance but are critical for estimates of density, as we will discuss below.

The technological innovation of adding a GPS (global positioning system) to radio collars has greatly enhanced the ability of ecologists to get frequent movement data on wide-ranging animals. For density estimates these movement data are of limited use, but they can greatly assist in the estimation of survival rates and reproductive rates (Conroy et al. 2000), as well as habitat selection for large carnivores and other threatened species (Hebblewhite and Haydon 2010). As always in ecological studies we should be clear about the question being asked and not have technology rule what we study.

### 3.3 ENUMERATION METHODS

Why estimate when you can count the entire population? This is one possible response to the problems of estimating density by mark-recapture, removal, or resight methods. All of these methods have their assumptions, and you could avoid all this statistical hassle by counting all the organisms in the population. In most cases this simple solution is not possible either physically or financially, and one must rely on a sampling method of some type. But there are a few situations in which enumeration is a possible strategy.

Enumeration methods have been used widely in small mammal studies (Krebs 1966) where they have been called the *minimum-number-alive* method (MNA). The principle of this method is simple. Consider a small example from a mark-recapture study (0 = not caught, 1 = caught):

Tag no.	Time 1	Time 2	Time 3	Time 4	Time 5
A34	1	0	1	1	0

A38	1	1	1	<b>0</b>	1
A47	1	<b>0</b>	<b>0</b>	1	1
A78	1	1	<b>0</b>	1	1
A79	0	0	1	<b>0</b>	1
A83	1	1	1	<b>0</b>	1
<hr/>					
No. caught	5	3	4	3	5
Minimum number alive	5	5	6	6	5

Individuals miss being caught at one or more sample periods (bold **0** in table), but if they are known to be present before and after the time of sampling, we assume they were present but not caught during the intervening sample times. We assume they did not move off the study area and then come back later. These individuals are added to the number actually caught to give the minimum number alive on the study site. Simple arithmetic, no complicated equations, and one might think an ecologists dream!

Enumeration methods have been used for birds, small mammals, reptiles, and amphibians (Pollock *et al.* 1990) and their simplicity makes them popular. But there is one overwhelming complaint about enumeration methods—they all suffer from a negative bias. Numbers enumerated are always less than or equal to the true population size (Jolly and Dickson 1983). If catchability is high, so that most individuals are caught in most of the sampling periods, this negative bias is small. But if catchability is low, the negative bias is very large. Also they operate with the underlying assumption that the population is closed in space and in time, so that emigration from the study area cannot be recognized.

The net result is that enumeration methods should not be used when it is practical to use one of the unbiased mark-recapture methods we have discussed in the last two chapters. At the very least one should compute both, for example, the Cormack-Jolly-Seber population estimates and the MNA estimates and show that these are nearly the same. Cormack-Jolly-Seber estimates can never be below MNA estimates, and these two will converge only when all animals are caught each sampling time.

In a few cases enumeration methods are needed and can be justified on the principle that a negatively biased estimate of population size is better than no estimate. In some cases the numbers of animals in the population are so low that recaptures are rare. In other cases with endangered species it may not be possible to mark and recapture animals continuously without possible damage so sample sizes may be less than the 5-10 individuals needed for most estimators to be unbiased. In these cases it is necessary to ask whether a population estimate is needed, or whether an index of abundance is sufficient for management purposes. It may also be possible to use the methods presented in the next two chapters for rare and endangered species. The important point is to recognize that complete enumeration is rarely a satisfactory method for estimating population size.

The fact that most animals do not satisfy the randomness-of-capture assumption of mark-recapture models is often used to justify the use of enumeration as an alternative estimation procedure. But Pollock *et al.* (1990) show that even with heterogeneity of capture the Cormack-Jolly-Seber model is less biased than MNA based enumeration procedures. A closed population estimator or the Cormack-Jolly-Seber model should be used in almost all cases instead of enumeration.

### **3.4 ESTIMATING DENSITY**

We have been discussing methods of estimating abundance, and yet many ecologists want to know the *density* of the population, the number of individuals per unit area or unit volume. If you know the size of the study area, it is easy to determine the density once you know the abundance of a species. But in many cases the translation from *abundance* to *density* is far from simple. Consider for example a grouse population in an extensive area of forest or moorland. Individuals captured and marked on the study area will move on and off the area, and for vertebrates their home ranges may be only partly on the study area. In practice this means that the size of the study area is larger than the physical area trapped and the true population density is less than the biased density estimate obtained by dividing the population estimate by the physical area of study. Figure 3.6 illustrates the problem

schematically. How can we determine how large an effective area we are studying so that we can obtain estimates of true density?

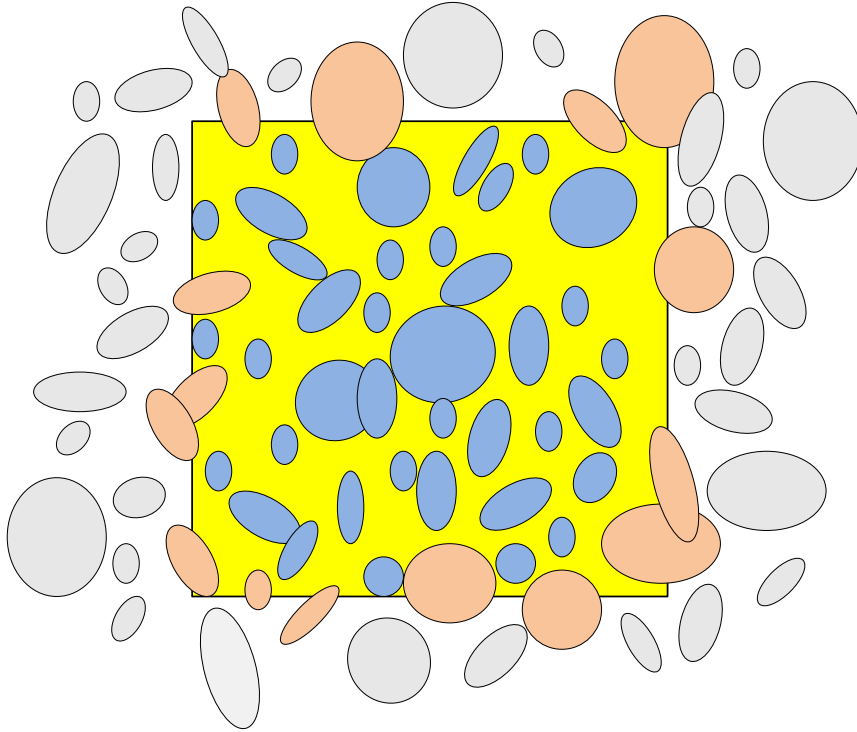
Several methods have been employed to estimate the effective size of a trapping area.

### **3.4.1 Boundary Strip Methods**

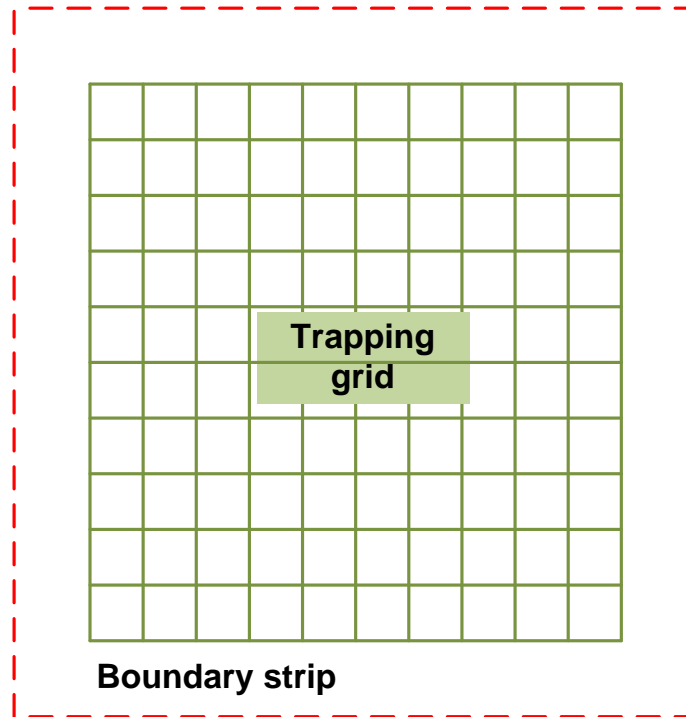
The simplest methods add a boundary strip around the trapping area to estimate the effective size of area trapped. Figure 3.7 illustrates the principle of a boundary strip. The width of the boundary strip can be estimated in several ways (Stenseth and Hansson 1979). The simplest procedure is to add a strip one-half the movement radius of the animals under study. For example, in mammals one can determine the average distance moved between trap captures and use this as a likely estimate of movement radius. The problem with this approach is that it is highly dependent on the spacing between the capture points, and the number of recaptures. While this approach is better than ignoring the problem of a boundary strip, it is not completely satisfactory (Otis *et al.* 1978).

Boundary strip width can be estimated in several ways. Two are most common and both rely on multiple recaptures in order to estimate a correlate of the home range size of the animals being studied (Wilson and Anderson 1985). The mean-maximum-distance-moved (MMDM) estimator calculates for each individual the maximum distance between capture points and obtains a mean of these values (Jett and Nichols 1987). This value of MMDM can then be divided by two to estimate the width of the boundary strip. An alternative that used the same data is to calculate the asymptotic range length (ARL) from the same data by fitting an asymptotic regression to the number of captures (X) versus the observed range length (Y) for each animal. The assumption is that as you obtain more and more captures on an individual, the home range length will increase up to some maximum, which is then used as an estimate of boundary strip width. Both MMDM and ARL can be calculated for a set of closed population data in Program DENSITY described below.

Another approximate boundary strip method was devised by Bondrup-Nielsen (1983) who used home range size to estimate effective size of a sampling area so that density could be estimated. He made the important point to adjust



**Figure 3.6** The trapping grid (yellow square) includes all (blue), part (tan), or none (grey) of the home ranges of the animals in the study zone. Some of the animals whose home ranges overlap the trapping grid will be captured, marked and released. The effective size of the trapping grid is thus larger than its physical area. If home ranges are very small, the effective grid area is only slightly larger than its physical area. (Modified from White *et al.* 1983).



**Figure 3.7** Illustration of a study grid of 10 x 10 checkerboard shape with a boundary strip added to delineate the effective study area. Population size would be estimated for the checkerboard and the total area including the boundary strip would be used to estimate population density.

sampling grid size to minimize the boundary strip effect by defining the required size of the sampling area. This method proceeds in three steps:

1. Calculate the average home range size for the animal under study. Methods for doing this are presented in Kenward (1987), White and Garrott (1990), and Seaman et al. (1998), and evaluated by Downs et al. (2012).
2. Compute the ratio of grid size to area of the average home range. In a hypothetical world with square home ranges that do not overlap, the overestimation of density is shown by Bondrup-Nielsen (1983) to be:

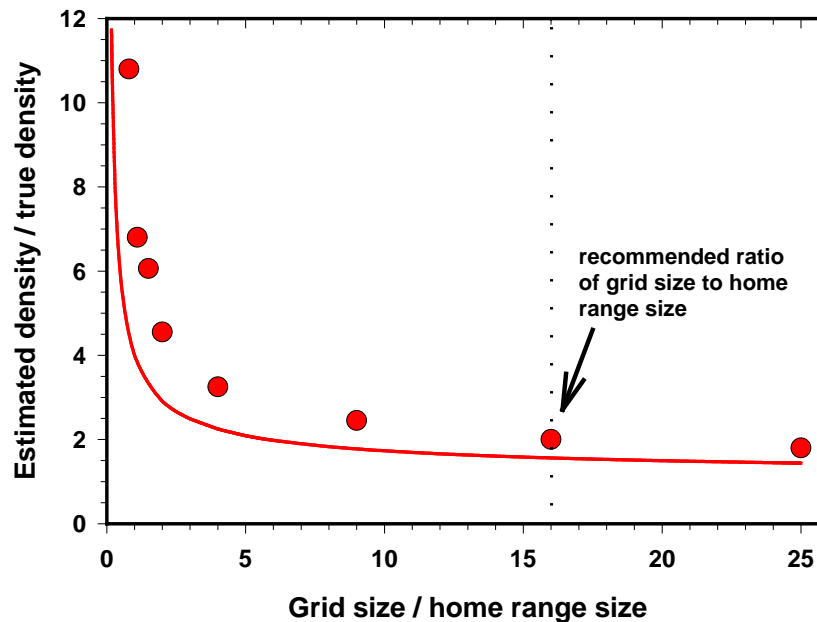
$$\frac{\text{Estimated density}}{\text{True density}} = \frac{(\sqrt{A} + 1)^2}{A} \quad (3.25)$$

$$\text{where } A = \frac{\text{area of study grid}}{\text{average home range size}} \quad (3.26)$$

3. Assume that the home range is elliptical and in a computer “throw” home ranges at random on a map that delineates the study zone as part of a large area of habitat.



Figure 3.8 shows the results of this simulation. On the basis of this graph, Bondrup-Nielsen (1983{ #801}) suggested that the trapping grid size should be at least 16 times the size of the average home range of the species under study to minimize the edge effect.



**Figure 3.8** Expected relationship between relative grid size (grid size / home range size) and the overestimation of density if no boundary strip is used to calculate population density. The theoretical curve is shown for square home ranges that do not overlap, and the red points show the average results of computer simulations with elliptical home ranges that are allowed to overlap. (Modified from Bondrup-Nielsen 1983).

Given that we have data on study area size and average home range size, we can use Figure 3.8 as a guide to estimating true density of a population. For example, in a study of house mouse populations in wheat fields live-traps were set out on a 0.8 ha grid, and the Petersen population estimate was 127 mice. The average home range size of these house mice was 0.34 ha. We can calculate from equation (3.26):

$$A = \frac{\text{Grid size}}{\text{Home range size}} = \frac{0.8}{0.34} = 2.35$$

From Figure 3.8 or equation (3.25) the expected overestimate is

$$\frac{\text{Estimated density}}{\text{True density}} = \frac{(\sqrt{A} + 1)^2}{A} = \frac{(\sqrt{2.35} + 1)^2}{2.35} = 2.73$$

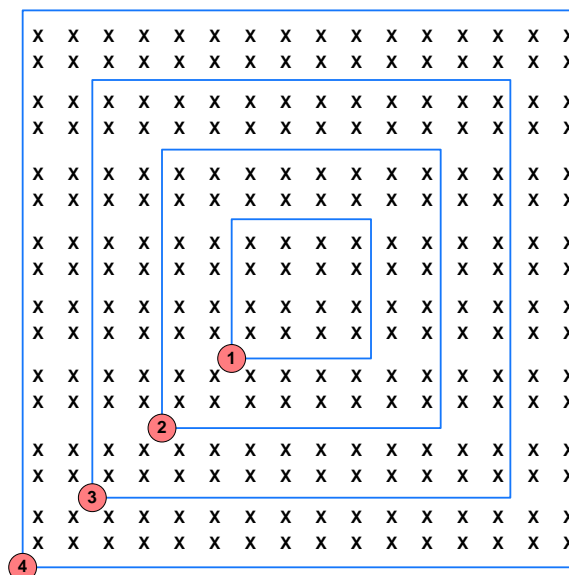
This the biased density estimated from the area of the grid:

$$\begin{aligned} \text{Biased density estimate} &= \frac{\text{Population estimate}}{\text{Size of study area}} = \frac{127 \text{ mice}}{0.8 \text{ ha}} \\ &= 159 \text{ mice per ha} \\ \text{Corrected density estimate} &= \frac{159}{2.73} = 58 \text{ mice per ha} \end{aligned}$$

This correction factor is only approximate because home ranges are variable in size and not exactly elliptical and there is some overlap among individuals, but the adjusted density estimate is closer to the true value than is the biased estimate that does not take into account the effective grid size.

### 3.4.2 Nested Grids Method

A more rigorous method for estimating density was developed by Otis *et al.* (1978). If a large area is sampled, it may be possible to break the data up into a series of nested grids (Figure 3.9). At least 4 nested grids are needed to get good estimates, and this will mean a 15 x 15 checkerboard of sample traps or capture points. The principle of the method is simple: the biased density estimate (population estimate / size of sampling area) will decline as the sampling area

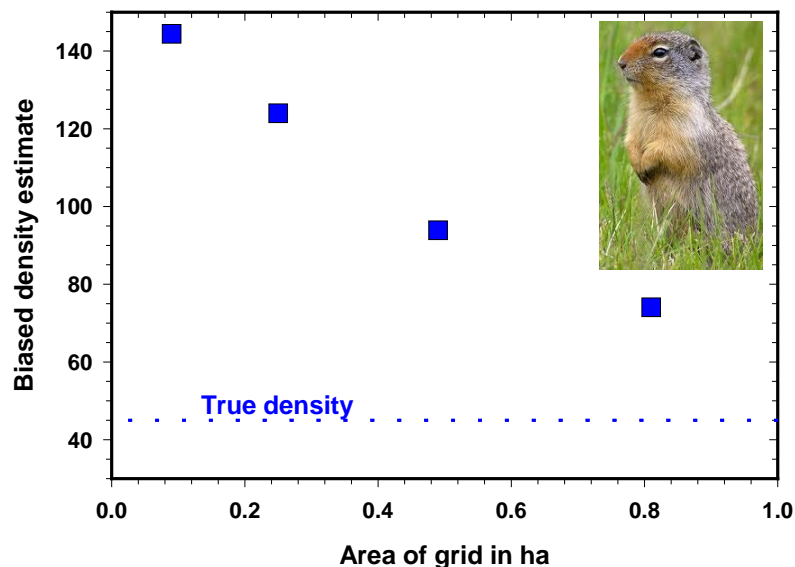


**Figure 3.9** An example of a set of nested grids for population density estimation. The entire study area is a 16 x 16 checkerboard, and 4 nested subgrids are shown by the blue lines. The nested grids are 4 x 4, 8 x 8, 12 x 12, and 16 x 16, and if the sample points are 10 m apart (for example), the areas of these four subgrids would be 0.16 ha, 0.64 ha, 1.44 ha, and 2.56 ha.

increases in size. Each nested grid will have a larger and larger area of boundary strip as grid size increases (Figure 3.9), and we can use the change in observed density to estimate the width of the boundary strip for the population being studied.

The estimation problem is difficult and the details are not given here (see Otis *et al.* 1978). Given the area and the population estimate for each subgrid, the problem is to estimate two parameters — boundary strip width and population density. If the boundary strip is very wide, the biased density estimates will change very slowly with grid size. Figure 3.10 illustrates an example from Richardson's ground squirrels in which biased density falls rapidly with grid size. The estimation procedures are available in Program MARK (described in Chapter 2).

There is some controversy about the utility of the nested-grid approach to estimating population density. Few trials have been done on field populations and some of these have provided accurate density estimates (Wilson and Anderson 1985). Computer-simulated populations however typically show a large positive bias in the density estimate because the strip width is underestimated (Wilson and Anderson 1985).



**Figure 3.10** Illustration of the approach to density estimation from nested grids. Data on Richardson's ground squirrels live-trapped on a 10 x 10 checkerboard grid with 10 m spacing. Four nested subgrids of 4 x 4, 6 x 6, 8 x 8, and 10 x 10 were used to generate four separate estimates of population size. The biased density estimate (population estimate / area trapped) declines with grid size. The best estimate of true density for these data was 45 squirrels per hectare, with a boundary strip width of 12.4 m. These estimates were obtained using Program CAPTURE. (Data from White *et al.* 1983).

The nested grid approach to density estimation is very data-intensive, and it will work best when a high fraction of the population can be marked and recaptured and when density is moderate to high. It assumes that the density of the population is uniform over all the area of the nested grids, and that the population is closed during the time of sampling. If the traps used are attractive to animals so that individuals are attracted into the study area, the assumption that the population is closed will be violated and this method should not be used. If a removal study is being conducted, and the removal extends over a long period so that immigrants enter the population, again the assumption of a closed population is violated and this method is not useful.

### **3.4.3 Spatial Methods for Multiple Recaptures**

So far we have discussed estimating the abundance of a mobile animal for a closed population. But if we know abundance we do not automatically know population density because, as we have just seen, we require an estimate of the area sampled, the effective trapping area. A new method of estimating density from mark-recapture of closed populations of mobile animals was developed by Borchers and Efford (2008), Efford (2004), and Efford *et al.* (2009). Efford's approach was to fit a simple spatial model of animal trapping that estimated the probability that an animal is caught in a trap at a given distance from its home range centre. This probability is assumed to follow a 2-parameter spatial detection function. This approach assumes that animals occupy stationary home ranges whose centers are a realization of a homogeneous random spatial point process with density  $D$ . An individual-based model uses the assumption that the probability of detecting an individual is a decreasing function of the distance between its home range centre and a live trap. The simplest detection function used by Efford (2004) has 2 parameters that correspond to a measure of home range size ( $\sigma$ , sigma), and the probability of

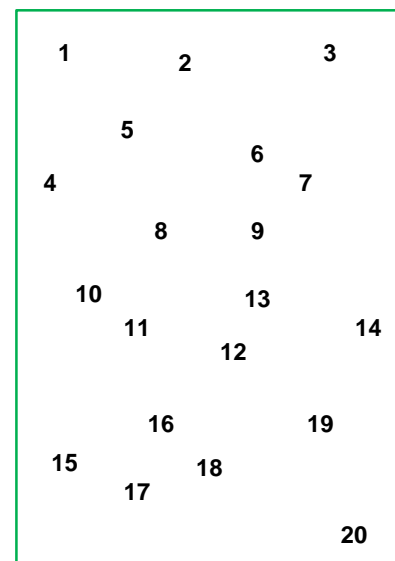
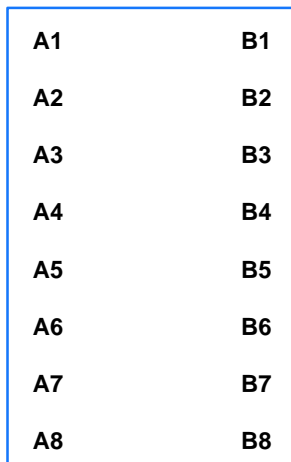
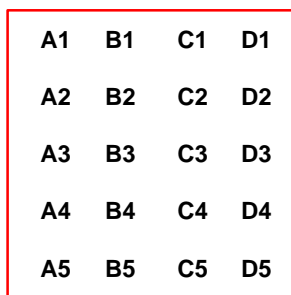
capture at the centre of the home range ( $g_0$ ). These 2 parameters can be estimated from the distances between recaptures of marked animals and the frequency of capture. The approach used in Efford's method assumes that the spatial point process is Poisson, and that the shape of the detection function is either half of a normal distribution, a negative exponential, or a hazard function.

Two possible methods of estimation of the spatial model have been developed. Efford (2004) developed the first approach through inverse prediction (IP) and simulation. A second approach was developed by Borchers and Efford (2008) and utilizes maximum likelihood methods to derive a better density estimate than is achieved by inverse prediction. Efford (2009) and Borchers and Efford (2008) used simulation to demonstrate that the maximum likelihood method gave the most accurate estimate of true density in comparison to other methods using the boundary strip approach. The details of the estimation problem are given in Efford (2004) and Borchers and Efford (2008) and implemented in Program DENSITY 4 (<http://www.otago.ac.nz/density/>).

The data required from a closed population for Program DENSITY are a series of data on the capture locations for a set of animals (Table 3.3). There must be two or more capture or live trapping sessions because the key data are measures of the scale of movements the individuals make. If there are no movements of marked individuals, the spatial model cannot be used, and the larger the sample of movements, the more precise the estimated density. A spatial map of the detectors (e.g. live traps) must also be available. The detectors can be in any spatial format as long as you have a spatial map of potential detectors (Figure 3.11). In particular the detectors such as live traps do not need to be in a checkerboard grid pattern like that shown in Figure 3.7.

**Table 3.3** An example of the type of data required for a population density estimate for a closed population of mobile animals from Program DENSITY. Marked individuals at several checkerboard points on a regular grid in this example. • indicates not caught.

Individual tag	Sampling Time				
	1	2	3	4	5
355	A1	•	A3	B6	•
392	A4	A4	A4	•	A5
416	B7	•	B8	B8	B8
422	•	C4	D2	•	•
536	•	•	E3	E3	D5
413	•	F9	E9	•	G7
212	H5	H5	H5	•	H5
188	•	J7	•	J9	J10



**Figure 3.11** Illustration of three spatial maps of detectors that could be used in Program DENSITY to estimate population density from movement data. The key is to know the position of each detector in Cartesian space. (See Efford et al. 2005.) {#15416}

The approach to density estimation via spatial models is useful for field ecologists because it allows a great variety of different data collections to be

analysed. The proximity detectors may be cameras, DNA caught in hair traps, multiple live-catch traps or single live-catch traps.

I will not try to detail the calculations used in these spatial models because the mathematics are beyond the scope of this book. Interested students are referred to Efford et al. (2009){#15413}. Box 3.3 illustrates the data and results from an analysis of stoat (weasel) DNA collections in New Zealand.

### Box 3.3 EFFORD'S SPATIAL ANALYSIS FOR POPULATION DENSITY ESTIMATION

Stoats (*Mustela erminea*) were sampled in a New Zealand beech forest by obtaining hair samples in sticky tubes and subjecting the hair to microsatellite DNA analysis to identify individuals (Efford et al. 2009). They obtained these capture data:

Spatial encounter histories of 20 stoats collected over 7 days on 7 parallel lines (A to G) of sampling stations with spacing of 500 m between lines and 250 m along the lines, Matakītaki, New Zealand. (– = not captured).

Animal #	Day of Sampling						
	1	2	3	4	5	6	7
1	A9	–	–	–	–	–	–
2	A12	A12	–	–	–	–	–
3	–	–	C6	B5	–	–	–
4	–	–	G3	–	F3	–	–
5	–	–	–	F2	E2	–	F1
6	–	–	–	–	–	E8	–
7	–	–	F5	–	–	G7	–
8	F6	–	–	–	–	–	–
9	–	A4	–	–	–	–	–
10	–	C5	–	–	–	–	–
11	–	D4	–	–	–	–	–
12	–	–	D7	–	–	–	–
13	–	–	E5	–	–	E4	–
14	–	–	F1	–	–	G3	–
15	–	–	F9	–	–	–	–
16	–	–	G13	–	F13	G13	–
17	–	–	–	G9	–	–	–
18	–	–	–	–	–	F1	–
19	–	–	–	–	–	G9	–
20	–	–	–	–	–	–	G8

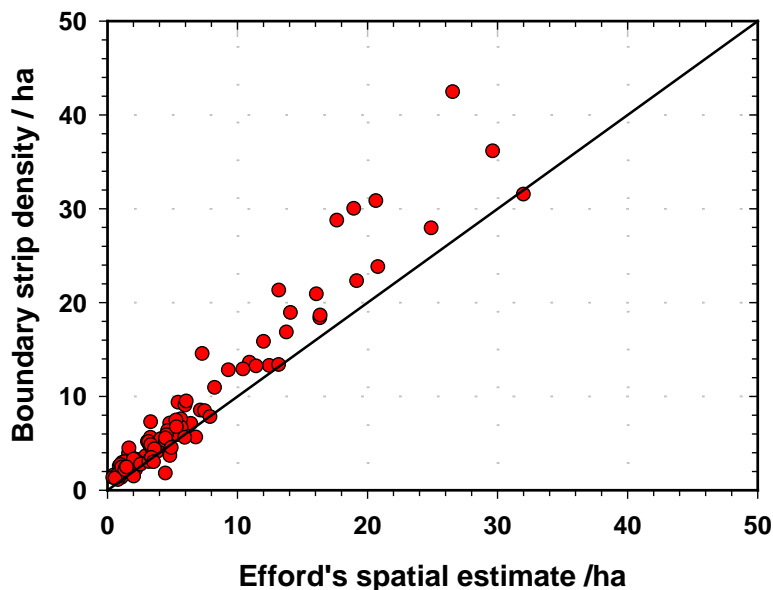
Seven of the 20 individuals were recaptured a total of 10 times. The longest movement was 707 m.

Efford et al. (2009) calculated density for this sampling period with each of the three possible detection functions, with these results:

Detection function	$\hat{g}_0$	$\hat{\sigma}$	Density per km <sup>2</sup>	95% confidence interval for density
Half-normal	0.045	255	2.5	1.4-4.5
Negative exponential	0.113	162	2.4	1.4-4.6
Hazard function	0.024	482	2.6	1.4-4.7

The density estimate for the stoats was not affected by the choice of the detection function. Confidence limits were broad because of the low rate of capture. The probability of being caught in a 'trap' was only 2-11% per day ( $\hat{g}_0$ ).

One general problem in evaluating mark-recapture models in all their forms is that we have few natural populations in which we know the true density. We can evaluate models like the spatial model by simulation, and they typically perform well when many movements are detected and a large fraction of the population is marked. But we do not know if these models are accurate in the field. Figure 3.12 illustrates a comparison of two density estimators, the boundary strip method of Figure 3.7 and Efford's spatial estimator for small rodents in northern Canada.



**Figure 3.12** Density estimates for small rodents in the Yukon boreal forest. Estimates were obtained from Program CAPTURE and density calculated from the boundary strip method and compared with density calculated by the spatial estimator of Efford et al. (2009) in Program DENSITY. If the estimates were identical we would expect them to fall along the



diagonal line. They do not and boundary strip methods overestimate density relative to spatial estimates. Since we do not know the true value of density for these populations, we do not know which estimator is more accurate. (Modified from Krebs et al. 2011). {#15430}

Depending on the method used, quite different values for density could be obtained. Unfortunately because we do not know the true density for these populations so we do not know which estimator produces results closer to the true density. There is a need for more studies with field populations of known density to evaluate these different models of density estimation.

### 3.5 SUMMARY

Three special techniques are presented in this chapter that may be useful for population estimation with exploited populations. If there are 2 types of organisms in a population, like males and females, and hunting removes more of one type than the other, the resulting *change-in-ratio* may be used to estimate population size. Large samples are usually required for change-in-ratio estimators, and you need to plan your sampling program thoroughly. Alternatively, if removals are not sex or age specific, you can use Eberhardt's index-removal method to estimate population size. This method depends on index counts that must enumerate about 20% or more of the total population to achieve precise estimates of population size.

If harvesting removes a reasonable fraction of the population, the *catch-per-unit-of-effort* may decline with time. If catchability is constant and this decline is linear, regression methods may be used to estimate population size at the time when exploitation begins. Fisheries managers have used this method extensively, but the assumption of constant catchability is vital if the population estimates are to be accurate.

Recent developments in mark-recapture methods have led to a series of models for closed population estimation. These models can accommodate to variation in catchability over time as well as among individuals, and thus circumvent the randomness of capture requirement of simpler methods. These new methods are computer intensive and have been implemented in two programs for personal computers. They are restrictive in assuming closed populations and at least 3 (and

typically 5) capture periods, and they require individual capture histories to be recorded for all animals.

Enumeration methods have a long history in population ecology and their simplicity makes them attractive. But all methods of enumeration suffer from a negative bias— population estimates are always less than or equal to the true population — and they should be used only as a last resort. The critical assumption of enumeration methods is that you have captured or sighted all or nearly all of the animals in the population, so that the bias is minimal. Enumeration methods may be required for endangered species which cannot be captured without trauma.

All of the methods in this chapter and the previous chapter estimate population size, and to convert this to population density you must know the area occupied by the population. This is simple for discrete populations on islands but difficult for species that are spread continuously across a landscape. Nested grids may be used to estimate density, or a boundary strip may be added to a sampling area to approximate the actual area sampled. Density estimates are more reliable when the sampling area is large relative to the home range of the animal being studied, and when a large fraction of the individuals can be captured.

Spatial models have been developed that utilize the movements of marked individuals in a closed population to estimate density. These methods assume that individuals have a constant home range size and move independently. They require a series of 5-10 or more movements for individuals with a trapping session. They could provide a less biased estimate of density than boundary strip methods, but more evaluation on field populations of known size is required to determine which density estimation method is most accurate.

### **SELECTED READINGS**

Efford, M.G., Borchers, D.L., and Byrom, A.E. 2009. Density estimation by spatially explicit capture-recapture: likelihood-based methods. *In* Modeling Demographic Processes in Marked Populations. *Edited by* E.G.C. D.L. Thomson, M.J. Conroy. Springer, New York. pp. 255-269.

- Foster, R.J., and Harmsen, B.J. 2012. A critique of density estimation from camera-trap data. *Journal of Wildlife Management* **76**(2): 224-236.
- Lebreton, J. D. (1995). The future of population dynamics studies using marked individuals: a statistician's perspective. *Journal of Applied Statistics* **22**, 1009-1030.
- Matlock, R. B. J., Welch, J. B. & Parker, F. D. (1996). Estimating population density per unit area from mark, release, recapture data. *Ecological Applications* **6**, 1241-1253.
- Miller, S. D., White, G. C., Sellers, R. A., Reynolds, H. V., Schoen, J. W., Titus, K., Barnes, V. G., Smith, R. B., Nelson, R. R., Ballard, W. B. & Schwartz, C. C. (1997). Brown and black bear density estimation in Alaska using radiotelemetry and replicated mark-resight techniques. *Wildlife Monographs* **61**, 1-55.
- Obbard, M.E., Howe, E.J., and Kyle, C.J. 2010. Empirical comparison of density estimators for large carnivores. *Journal of Applied Ecology* **47**(1): 76-84.
- Rosenberg, D. K., Overton, W. S. & Anthony, R. G. (1995). Estimation of animal abundance when capture probabilities are low and heterogeneous. *Journal of Wildlife Management* **59**, 252-261.
- Wilson, D.J., Efford, M.G., Brown, S.J., Williamson, J.F., and McElrea, G.J. 2007. Estimating density of ship rats in New Zealand forests by capture-mark-recapture trapping. *New Zealand Journal of Ecology* **31**: 47-59.
- Wilson, K. R. & Anderson, D. R. (1985). Evaluation of two density estimators of small mammal population size. *Journal of Mammalogy* **66**, 13-21  
 {#15420} {#15937} {#15413} { #15990}

### QUESTIONS AND PROBLEMS

- 3.1.** For the catch-effort regression estimation technique of Leslie, discuss the type of bias in estimated population size when there is:
- Change in catchability during the experiment
  - Natural mortality during the experiment.
  - Mortality caused by the marking procedure or the tag itself.
  - Emigration of animals from the population.
- 3.2.** Leslie and Davis (1939) set 210 rat traps (break-back type) in 70 houses in Freetown, Sierra Leone in 1937. They caught over 18 days:

Day	No. of <i>Rattus rattus</i> caught
-----	------------------------------------

1	49
2	32
3	31
4	34
5	16
6	33
7	22
8	27
9	17
10	19
11	18
12	16
13	18
14	12
15	14
16	12
17	17
18	7

210 traps were set each day. Estimate the total population size of *Rattus rattus* at the start of the study. What sources of error might affect this particular estimate?

**3.3.** During a severe winter in 1938-39 in Utah mule deer populations suffered heavy losses from starvation. The ratio of fawns to adults was 83 fawns and 100 adults before the severe weather ( $n=69$  and 83 observed), and 53 fawns and 100 adults after the winter ( $n=38$  and 72). Over the study area 248 dead fawns and 60 dead adults were found and these were believed to represent the whole of the losses (Rasmussen and Doman 1943). Estimate population size for this deer population at the start and end of this winter. Calculate 95% confidence intervals for your estimates.

**3.4.** Dunnet (1963) sampled quokkas (small marsupials) on Rottnest Island off Western Australia over three sampling periods with these results for 1957:

	Mature animals	Immature animals
Number marked at time 1	32	35
Number marked at time 1 and never seen again	22	20
Number marked caught at	5	1

times 1 and 2 only		
Number marked caught at times 1 and 3 only	4	11
Number marked caught at times 1, 2 and 3	1	3
Number marked first at time 2 and never seen again	38	28
Number marked first at time 3	26	23
Number marked caught at times 2 and 3 only	4	7

Recast these data in the form of an  $X$ -matrix, and use Program CAPTURE to estimate the size of the mature population and the immature population of this marsupial. What model is selected by Program CAPTURE, and how different are the estimates from the different models available in this program? Does the tests in Program CAPTURE suggest any evidence for unequal catchability in these data?

- 3.5.** Plan a study that will use the change-in-ratio estimator of population size. You would like to get an estimate within  $\pm 10\%$  of the true value at an  $\alpha$  of 5%. You expect a change in the proportion of males in your population to be about 10-15%. Describe your proposed sampling program in detail. How is it affected by uncertainty about the initial sex ratio (which might be 50% males up to 70% males)? How is it affected by the rate of exploitation?
- 3.6** White-footed mice were live-trapped in northern Michigan by Blair (1942). He trapped a large area of 7.4 ha. The population estimate was 39 males and 31 females for this trapping area. Male white-footed mice have home ranges that average 0.93 ha while females have smaller ranges (0.56 ha). What are the biased density estimates for the two sexes of mice for this particular area, and what are the best estimates of the true density for each sex? What size of live-trapping area would you recommend for this species?
- 3.7** An aerial index count of feral water buffalo in the Northern Territory of Australia was carried out before and after a culling program in 1984. In the first index

count 2786 buffalo were counted, and in the second count 1368 were seen. The cull occurred over two months and 12,890 buffalo were shot. Can you estimate population size for this population? What proportion of the population was culled? What assumptions must you make to use these estimates?

- 3.8.** During 7 aerial flights to locate 19 radio-collared caribou in southern British Columbia over a two week period a total of 144 caribou were seen of which 54 had radio-collars and the remaining 90 were not marked. The frequency of resighting of the 19 collared caribou were as follows: 7, 5, 1, 2, 3, 0, 5, 3, 1, 3, 0, 2, 4, 6, 1, 5, 1, 3, and 2. Estimate the population size of this caribou herd and the 90% confidence limits for this estimate.
- 3.9** Discuss from a practical point of view the advantages and disadvantages of using a trapping web design versus a nested grid design to estimate population density for a small mammal. What information might be lost if you adopted one design over the other?