United States
Department of
Agriculture

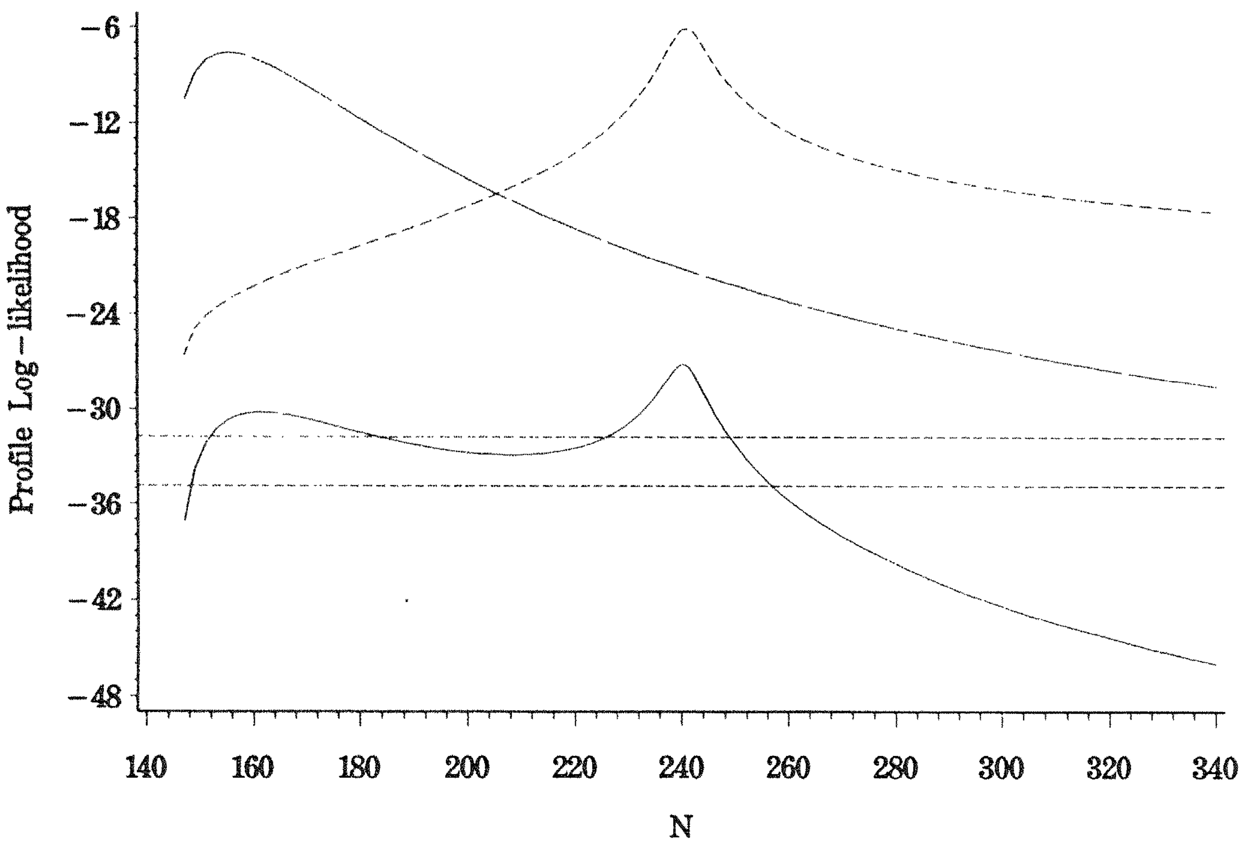
Forest Service

**Northeastern Forest
Experiment Station**

General Technical
Report NE-199

Some Results on the Combined Removal and Signs-of-Activities Estimators for Sampling Closed Animal Populations

Jeffrey H. Gove
Ernst Linder
Walter M. Tzilkowski



Abstract

The possibility of a bimodal log-likelihood function arises with certain data when the combined removal and signs-of-activities estimator is used. This possibility may present an inference problem—yielding disjoint confidence intervals for certain confidence levels. One option is to weight the combined likelihood in favor of the more reliable sample component (e.g., removal or signs-of-activities data) when bimodality does arise. Simulations exploring the effect of model assumptions on estimation and inference showed that violation of removal model assumptions by way of unequal capture probability influenced the frequency of bimodal likelihoods; similarly, extreme parameter values for probability of capture influenced the number of excessively large confidence intervals produced. The simulations suggest that the signs-of-activities estimator should be used in lieu of the traditional removal estimator or the more complicated combined estimator under these circumstances. The signs-of-activities estimator may be the best overall for combined estimation, though sensitivity to its assumptions was not explored.

The Authors

JEFFREY H. GOVE is a research forester with the research work unit on Methods for Measurement, Analysis, and Modeling of Forest Growth and Structure at the Northeastern Forest Experiment Station at Durham, New Hampshire. He received A.A.S., B.S., and M.S. degrees in forestry from the University of New Hampshire, and an M.A. degree in statistics and a Ph.D. degree in forest biometrics from the Pennsylvania State University.

ERNST LINDER is an associate professor at the University of New Hampshire. He received a B.S. degree in mathematics from the Federal Institute of Technology (ETH), Zurich, Switzerland, and a Ph.D. degree in statistics from The Pennsylvania State University. He has collaborated on research on ecological risk assessment, marine fisheries, forestry, global climate modeling, and gamma-ray astronomy.

WALTER M. TZILKOWSKI is an associate professor of wildlife science at The Pennsylvania State University. He received M.S. and Ph.D. degrees in wildlife management from West Virginia University and the University of Massachusetts, respectively.

Manuscript received for publication 4 May 1994

USDA FOREST SERVICE
5 RADNOR CORP CTR STE 200
P.O. BOX 6775
RADNOR, PA 19087-8775

January 1995

Introduction

The removal method for sampling wildlife populations is commonly used to determine population size. In simplest terms, animals in a closed population that are caught on M attempts are physically removed from the population, while uncaught animals simply remain in the population. This dichotomy forms the basis for using the binomial probability distribution as a model for analyzing such experiments. In addition to the assumption of closure, the other assumptions required for this model are constant probability of capture for all individuals and the use of constant effort in catching individuals by the experimenter. When these assumptions are violated, the appropriateness of the removal method becomes suspect; investigators have proposed other models for such situations (see White et al. (1982) for a discussion of these assumptions and alternative methods).

Routledge (1989) presented a modification to the classical removal estimator that considered auxiliary information in the form of signs-of-activities and which he called the *combined* estimator. The combined estimator allows the investigator to gather other evidence of animal activity that may be related to population size. These data are used in conjuction with traditional removal data to yield a more refined estimate of population size. The signs-of-activities estimator is based on the normal distribution and uses similar techniques (i.e., maximum likelihood) in solving for the population size as the removal estimator.

This study grew out of a desire to use this new estimator in the field. The first step was to replicate Routledge's results using his data from a combined removal and signs-of-activities experiment with cockroaches. This proved successful, though in the process we discovered that the cockroach data used by Routledge produced a bimodal profile log-likelihood, and that the possibility of generating a profile log-likelihood of this form had not been reported. Therefore, the purpose of our original field study took on a new dimension: to investigate the occurence of bimodality in the combined estimator and its possible relation to violations of the assumptions in the removal model. This report presents results of simulation experiments designed to explore this question. Included are appendices of computer code and spreadsheet programs to analyze removal and signs-of-activities data, or conduct similar simulations. Also included are appendices that contain the derivation of the likelihood functions and extensions to test statistics used throughout.

The Combined Estimator of Routledge

Routledge combined the log-likelihoods λ_1 and λ_2 ((1) and (2) that follow) of a removal model and a signs-of-activities model, respectively, into the sum $\lambda = \lambda_1 + \lambda_2$. From this the "combined" maximum likelihood estimate (MLE) of the population size (N) is computed. The combined estimator has a much tighter sampling distribution than the estimator based on the removal model alone, particularly when the catchability rate p is not large (< 0.5), as Routledge illustrated in his Table 1.

Following the notation of Routledge (1989, p. 112–113), we let

- $M =$ the number of removal attempts;
- $R_i =$ the number removed on the ith attempt;
- $T_i = \sum_{j=1}^{i} R_j =$ the total number of individuals removed by the end of the ith attempt, for $i = 1, 2, \ldots, M$;
- $S = \sum_{i=1}^{M}(N - T_{i-1}) =$ the sum of all the numbers of animals available for capture at each stage, with $T_0 = 0$;
- $Y_0 =$ the number of signs of activity observed over some time interval of length t, prior to the first removal attempt;
- $Y_i =$ the number of signs of activity observed over some time interval of length t, between the ith and the $(i+1)$th removal attempts, for $i = 1, 2, \ldots, M-1$;
- $Y_M =$ the number of signs of activity observed over some time interval of length t, after the last (Mth) removal attempt;
- $N =$ the initial population size.

The removal estimator seeks an estimate of N assuming that $R_1 \sim \text{Binomial}(N, p)$ on the first attempt and $(R_i | \sum_{j=1}^{i-1} R_j) \sim \text{Binomial}(N - \sum_{j=1}^{i-1} R_j, p)$ on all other attempts. Therefore, an estimate $\hat{N}_R$ of N is found by maximizing the log-likelihood

$$\lambda_1 = \log(N!) - \sum_{i=1}^{M} \log(R_i!) - \log((N - T_M)!) + T_M \log(p) + (S - T_M) \log(1 - p). \tag{1}$$

Setting $\partial\lambda_1/\partial p = 0$ yields the overall success rate, $\hat{p} = T_M/\hat{S}$, upon substituting $\hat{N}_R$ for N (see Appendix I for derivation).

The signs-of-activities model assumes that some sign, such as animal tracks or fecal material, is left by the animals independently of each other and that removal of individuals does "...not affect the rate (per remaining individual) of deposition or detection of these signs of activity" (Routledge 1989, p. 113). A further assumption is required: that all old signs of activity can be "...erased or removed at each sampling occasion or, equivalently, that new signs can be distinguished from old signs" (Routledge 1989, p. 113). The signs-of-activities model assumes that $Y_i \sim \text{Normal}(b(N - T_i), \sigma^2(N - T_i))$, and, therefore, $b = \text{E}[Y_i]/(N - T_i)$ and $\sigma^2 = \text{Var}(Y_i)/(N - T_i)$. Then an estimate $\hat{N}_S$ of N is found by maximizing the log-likelihood

$$\lambda_2 = -\frac{1}{2}\left((M+1)\log(2\pi\sigma^2) + \sum_{i=0}^{M} \log(N - T_i) + \frac{1}{\sigma^2}\sum_{i=0}^{M} \frac{(Y_i - b(N - T_i))^2}{N - T_i}\right). \tag{2}$$

In a similar manner, setting $\partial\lambda_2/\partial b = 0$ yields $\hat{b} = \bar{Y}/(\hat{N}_S - \bar{T})$ and $\partial\lambda_2/\partial\sigma^2 = 0$ yields $\hat{\sigma}^2 = (M+1)^{-1}\sum_{i=0}^{M}(Y_i - \hat{b}(\hat{N}_S - T_i))^2/(\hat{N}_S - T_i)$, with $\bar{Y} = \sum_{i=0}^{M} Y_i/(M+1)$, $\bar{T} = \sum_{i=0}^{M} T_i/(M+1)$, and $T_0 = 0$ (see Appendix I for derivation). Note that substituting the MLEs of p, b, and σ^2 into (1) and (2) yields two functions that can be solved numerically in terms of N alone to obtain the removal ($\hat{N}_R$) or the signs-of-activities ($\hat{N}_S$) estimate. In addition, because the R_i and the Y_i are assumed conditionally independent, Routledge combined the two log-likelihoods by addition (i.e., multiplying the likelihoods), yielding the combined log-likelihood

$$\lambda = \lambda_1 + \lambda_2, \tag{3}$$

which also can be solved numerically for $\hat{N}$ alone.

Routledge also presented a method for finding confidence interval endpoints based on the likelihood ratio test (LRT) statistic. The presence of the nuisance parameters p (catchability rate), b, and σ^2 (slope and error variance, respectively, per number of individuals remaining in the signs-of-activities model) in the likelihoods requires the use of the profile log-likelihood for constructing confidence intervals based on the LRT. The profile log-likelihoods are obtained by substituting the maximum likelihood solutions of the nuisance parameters p, b, and σ^2 into the log-likelihood function for any given value of N. The profile log-likelihoods $\lambda_1(N)$ and $\lambda_2(N)$ for the cockroach data used in Routledge (1989) are plotted in Figure 1, and in Figure 2 for $\lambda(N)$. Therefore, an approximate $1 - \alpha$ confidence interval for N based on the removal model alone can be calculated by considering all points N_0 with a profile log-likelihood value that is at least within $1/2\chi^2_\alpha$ of the maximum value $\lambda_1(\hat{N}_R)$. This follows from the LRT statistic $-2\log\Lambda_R = 2\lambda_1(\hat{N}_R) - 2\lambda_1(N_0)$ with asymptotic distribution of χ^2_α for the test $\mathcal{H}_0 : N = N_0$; here, χ^2_α is the $1-\alpha$ quantile of the $\chi^2(1)$ distribution and $\Lambda_R = \mathcal{L}(\underline{r}; N_0, p)/\mathcal{L}(\underline{r}; \hat{N}_R, p)$ is the likelihood ratio for the removal estimate.

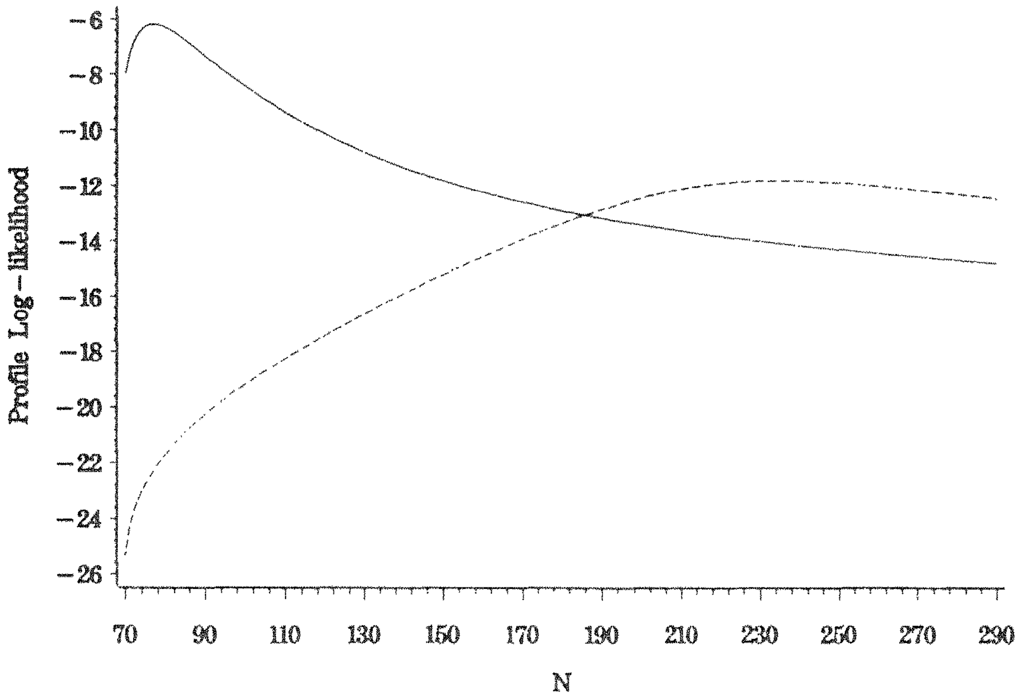


Figure 1.—Profile log-likelihood for removal (solid line) and signs-of-activities (dashed line) models using cockroach data from Routledge (1989, p. 117).

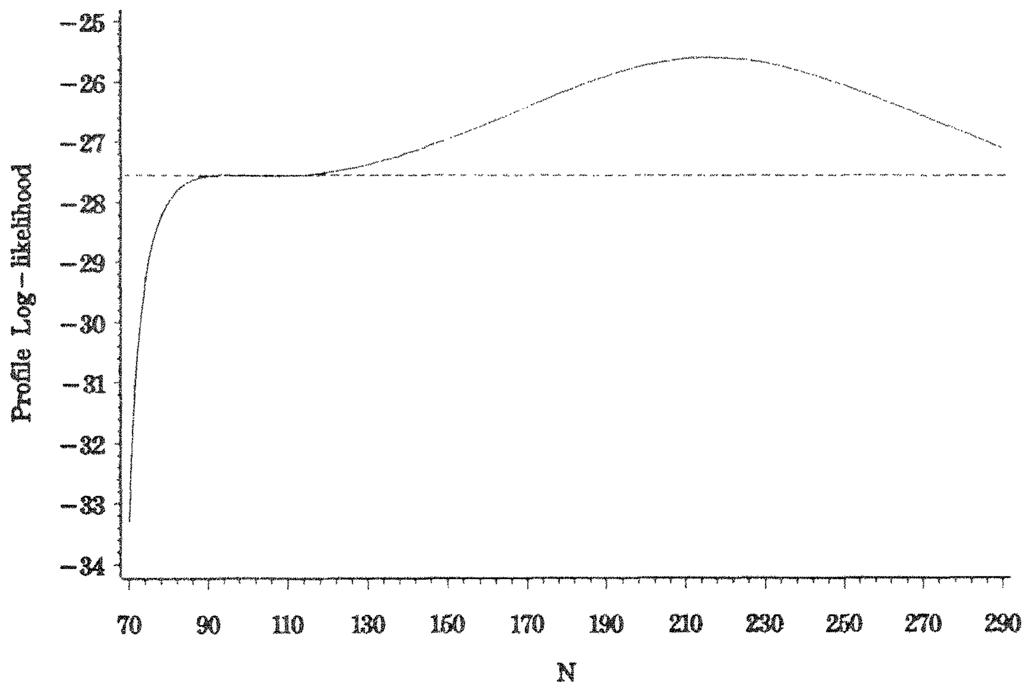


Figure 2.—Profile log-likelihood for combined model using cockroach data from Routledge (1989 p. 117). The dashed line shows the 79.31% confidence level with the LRT statistic at equality.

Routledge showed that under the signs-of-activities model, one could replace the chi-square quantile χ^2_α by $\tilde{F}_\alpha$, where

$$\tilde{F}_\alpha = (M+1)\log\left(1 + \frac{F_\alpha}{M-1}\right), \tag{4}$$

and F_α is the $1-\alpha$ quantile of the $\mathrm{F}(1, M-1)$ distribution. Based on this approximation, $1-\alpha$ confidence intervals for N consist of all values N_0 for which

$$-2\log\Lambda_S = 2\lambda_2(\hat{N}_S) - 2\lambda_2(N_0) \leq \tilde{F}_\alpha, \tag{5}$$

where $\hat{N}_S$ is the signs-of-activities estimate of N and $\Lambda_S = \mathcal{L}(\underline{y}; N_0, b, \sigma^2)/\mathcal{L}(\underline{y}; \hat{N}_S, b, \sigma^2)$ is the likelihood ratio for the signs-of-activities estimate.

Simulations by Routledge showed that using this same modified chi-square approximation (5) for the combined estimator produced acceptable results, so he adopted this approach to find confidence limits on the combined estimate. A $1-\alpha$ confidence interval for N under the combined model consists of all values N_0 for which

$$-2\log\Lambda_C = 2\lambda(\hat{N}_C) - 2\lambda(N_0) \leq \tilde{F}_\alpha, \tag{6}$$

where $\hat{N}_C$ is the combined estimate of N and $\Lambda_C = \mathcal{L}(\underline{r}, \underline{y}; N_0, p, b, \sigma^2)/\mathcal{L}(\underline{r}, \underline{y}; \hat{N}_C, p, b, \sigma^2)$ is the likelihood ratio for the combined estimate.

The Likelihood and LRT Functions

We used the cockroach data presented in Routledge (1989, p. 117) and GRG2—a generalized reduced gradient optimization package (Lasdon and Waren 1979, 1986)—to find the MLEs $\hat{N}_R$, $\hat{N}_S$, and $\hat{N}_C$ for N based on the removal, signs-of-activities, and combined model, respectively. Our estimates of $\hat{N}_R = 76.700$ and $\hat{N}_S = 235.029$ agree with those of Routledge. These optima are readily apparent in the profile log-likelihoods for these models presented in Figure 1. Upon maximizing the combined likelihood of the roach data, however, we arrived at an estimate of $\hat{N}_C = 94.092$, using a starting value (*viz.* 75.1) close to the original removal estimate. However, Routledge's estimate for $\hat{N}_C$ was 216.32.

A plot of the combined profile log-likelihood (Fig. 2) shows that the reason for this disagreement is that the profile log-likelihood is multimodal for the roach data. There is a local optimum at $\hat{N}_C = 94.092$, strong enough that the Kuhn-Tucker stationary conditions (Wismer and Chattergy 1978) were satisfied to within 1.0×10^{-7} at this point. However, the global optimum (also satisfying the Kuhn-Tucker conditions) was correctly found to be at $\hat{N}_C = 216.324$ (Routledge's estimate) when we started GRG2 from a point (*viz.* 175.1) closer to this value.

The fact that the profile log-likelihood is bimodal for the roach data is not surprising when one considers simply adding λ_1 and λ_2 in Figure 1. In terms of sufficiency, Kendall and Stuart (1967 p. 38) noted that if there is no single sufficient statistic t for a parameter θ, then a single unique maximum is not guaranteed. Such is the case under the combined model.

The possible bimodality of the profile log-likelihood function (bi- or unimodality depends on the observed data) may present a problem in maximum likelihood estimation of the combined model in determining the global maximum. However, insight can be gained into the shape of the profile log-likelihood function by plotting it over the range of N for which estimates are likely, as in Figures 1 and 2. Indeed, Fisher (1956, p. 71) and Kendall and Stuart (1967, p. 62) both recommended this procedure for any likelihood function.

Table 1.—Simulated removal experiment data set with $N = 200$, $p_M = 0.7$ (male catchability rate), $p_F = 0.1$ (female catchability rate), $\sigma = 0.2$, $b = 1.0$, and $M = 3$ removals; an equal sex ratio was assumed (i.e., $N_M = N_F = 100$)

i	R_i	Y_i
0	—	207
1	85	117
2	20	97
3	9	87

In the case of a bimodal profile log-likelihood, one also may begin the numerical optimization algorithm at several different starting points and see where they converge. Kendall and Stuart (1967, p. 38) recommended choosing the largest of the local maxima as the global optimum. They also discussed the case that might arise for certain combinations of samples where the likelihood function has two *equal* maxima (Kendall and Stuart 1967, p. 41). Such a combination of samples, they noted, has a low probability of occurrence, and is not of practical concern.

The possible bimodality of $\lambda(N)$ does have interesting ramifications with respect to the interpretation of confidence intervals constructed using the LRT statistic technique. Inasmuch as there are two maxima satisfying $\partial^2\lambda/\partial N^2 < 0$, say $\hat{N}_{C1} < \hat{N}_{C2}$, there must be a minimum for which $\partial^2\lambda/\partial N^2 \,|_{N=\hat{N}_{C3}} \geq 0$ such that $\hat{N}_{C1} < \hat{N}_{C3} < \hat{N}_{C2}$. This being the case, there is some $100(1-\alpha)\%$ confidence level such that (6) has more than two solutions under strict equality (that is, more than the "traditional" two endpoints are possible for the confidence interval). In the case of the roach data, one such confidence level is the 79.31% level with points $N_0 = 90.00, 101.93, 108.46$, and 309.40, all satisfying strict equality in (6). In this case, the actual 79.31% confidence interval is the *union* of the two intervals: $[90.00, 101.93] \cup [108.46, 309.40]$, and any points *viz.* $(101.93, 108.46)$ between these sets are not in the interval, just as points < 90.00 and > 309.40 are outside the interval.

The bimodality and hence the α level(s) where disjoint intervals may occur depend on the observed data; for any given data set, bimodality may occur for one of the more popular levels of α, such as .05 or .01. Figure 3 shows a combined log-likelihood profile that arose from the simulated data in Table 1. This profile log-likelihood shows two distinct maxima with the "removal peak" as the global maximum. The simple removal and signs-of-activities estimates for N were $\hat{N}_R = 116.0$ and $\hat{N}_S = 196.9$, respectively. When the combined log-likelihood function is used with these data, it is apparent that two vastly different estimates for N are possible depending on whether the solution algorithm is started at $\hat{N}_R$ or $\hat{N}_S$. The estimates $\hat{N}_{C(\hat{N}_R)} = 118.5$ and $\hat{N}_{C(\hat{N}_S)} = 196.9$ correspond to the two peaks in Figure 3 when the reduced gradient algorithm was begun at $\hat{N}_R$ and $\hat{N}_S$, respectively. It is clear that the estimate $\hat{N}_{C(\hat{N}_R)}$ is a poor estimate of the true population size of $N = 200$; yet it is the global maximum for the log-likelihood function. The question is, then, given no prior knowledge of the population size, which estimate do we choose? The 95% confidence intervals calculated using the modified chi-square LRT procedure are equally unenlightening in this case. For the estimate $\hat{N}_{C(\hat{N}_R)}$, the intervals are $[114.19, 135.83] \cup [194.07, 199.27]$; those for the estimate $\hat{N}_{C(\hat{N}_S)}$ are $[114.11, 139.35] \cup [193.11, 199.95]$. If the recommendations given previously were followed, $\hat{N}_{C(\hat{N}_R)}$ would be chosen as our estimate of N for this experiment. Yet this profile log-likelihood is close to the "equal maxima" case mentioned by Kendall and Stuart (1967, p. 41). Clearly, neither the estimates nor the confidence intervals provide direction as to which estimate of N to favor in this situation. If these were real data, the ramifications could be serious depending on the empirical situation that generated them.

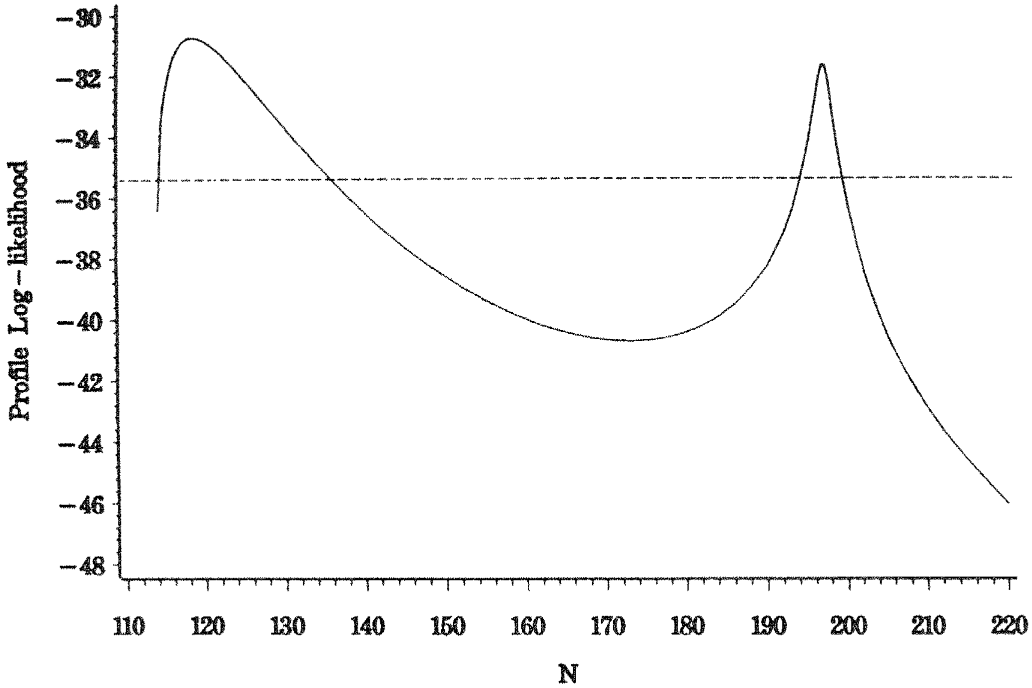


Figure 3.—Profile log-likelihood for combined model using the simulation data set in Table 1. The dashed line shows the 95% confidence level with the LRT statistic at equality when calculated from the "removal peak."

Weighted Likelihood Estimation

To someone familiar with likelihood functions and LRT statistic confidence intervals, the possibility of bimodality and associated disjoint nature of some confidence intervals arising for certain data sets is no surprise. In such cases, bimodality should alert the experienced user that certain model assumptions may have been violated. The problem then becomes one of determining which assumptions are in violation and the type of remedial measures that can be undertaken. Yet what happens in the case where standard remedial measures are not feasible? For the practitioner who is used to the concept of *continuous* confidence intervals, these results could indeed be troubling. In addition, as pointed out in the previous example, situations might arise where particular realizations of the sample data from an experiment point to no clear distinction between the two estimates in a bimodal log-likelihood.

One possible remedy is to use a weighted likelihood approach in the combined model to smooth out the bimodality issue. For instance, if we assign the weights $0 \leq \alpha_1 \leq 1$ and $0 \leq \alpha_2 \leq 1$ to each log-likelihood λ_1 and λ_2, respectively, then the combined log-likelihood becomes

$$\lambda^W = \alpha_1 \lambda_1 + \alpha_2 \lambda_2. \tag{7}$$

Two more constraints are needed to completely formulate this maximization problem in terms of the three unknowns N, α_1, and α_2. This problem now can be written as

$$\begin{array}{rl} \underset{\{N,\alpha_1,\alpha_2\}}{\text{Max}} & \alpha_1\lambda_1 + \alpha_2\lambda_2 \\ \text{St:} & \alpha_1 + \alpha_2 = 1 \\ & N - T_M > 0 \\ & \alpha_1, \alpha_2 \geq 0, \end{array} \tag{8}$$

and solved with GRG2. Unfortunately, this strategy failed for the roach data because $\lambda_1(\hat{N}_R)$ is the maximum attainable value for λ^W; that is, (8) will converge to the removal estimate with $\alpha_1 = 1$ and $\alpha_2 = 0$ because these values for α_1 and α_2 yield λ_1, the removal log-likelihood (Fig. 1). Model (8) will yield similar results for all datasets: it always will converge to $\lambda_1(\hat{N}_R)$ when $\lambda(\hat{N}_R) > \lambda_2(\hat{N}_S)$ or to $\lambda_2(\hat{N}_S)$ when the reverse is true. The only other possibility is the unlikely event that $\lambda_1(\hat{N}_R) \equiv \lambda_2(\hat{N}_S)$; it is unclear what weights will be assigned in such cases as λ_1 and λ_2 depend on the experimental data.

An alternative approach to model (8) is to use the weighted approach, but with maximization only over $\{N\}$ rather than $\{N, \alpha_1, \alpha_2\}$. The new objective function would then become

$$\underset{\{N\}}{\text{Max}} \quad \alpha_1\lambda_1 + \alpha_2\lambda_2 \tag{9}$$

with all of the constraints kept the same as in model (8). The problem then becomes one of choosing α_1 and α_2 outside of the optimization problem with some other auxilliary objective in mind. A priori choices for α_1 and α_2 seem to be one reasonable solution to this problem. The simplest approach allows the investigator to choose the weights based on intuition or prior empirical evidence. For example, if the investigator decided to weight in favor of the signs-of-activities estimator and $(\alpha_1, \alpha_2) = (0.25, 0.75)$ was chosen, then subsequent solution of (9) would yield $\hat{N}_C^W = 228.781$ as the weighted MLE of N ($N = 228$ is the actual population size).

This weighted approach is appealing when one considers the possible violation of the assumption of constant capture probability for all individuals in the removal method mentioned by Routledge (1989). In the roach experiment, for example, the trapping method used apparently was highly selective for males, so the female section of the population was grossly underrepresented in the roaches trapped. This led to an estimate $\hat{N}_R$ that was representative only of the males in the population. Where trapping results are able to be judged as biased because of the lop-sided sex ratio for individuals caught (an a priori assumption of equal sex ratios must be tenable to make this observation), adding more weight to the signs-of-activities log-likelihood seems justifiable, provided that there has been no obvious violation of assumptions in recording the signs-of-activities data. This a posteriori weighting simply reflects the investigator's desire to invest more in the likelihood that carries more reliable information while still not completely discounting the information in the down-weighted data.

Unfortunately, simply choosing the weights does not guarantee that the resulting weighted log-likelihood function will be unimodal. The question arises: what is the minimum weight combination needed to insure unimodality? An alternative approach that answers this question requires that the investigator choose only to weight in favor of either λ_1 or λ_2 a priori, but does not require the investigator to actually choose the weights. Assume for illustration that we chose to weight in favor of the removal estimator. Then a search technique can be used to find the auxilliary objective $\min(\alpha_1 - 0.5) > 0$ such that both starting values $\hat{N}_R$ and $\hat{N}_S$ converge to the same result under model (9) within some specified tolerance. In other words, the algorithm chooses the minimum amount of weighting associated with the favored log-likelihood (λ_1 or λ_2) that will be needed to produce a unimodal combined log-likelihood.

To illustrate the search algorithm in simple terms, assume that we desire to weight in favor of the removal log-likelihood, then on the first step the weights $(\alpha_1, \alpha_2) = (0.75, 0.25)$ are chosen and maximization is performed. If the log-likelihood is still bimodal, then the new weights (0.875, 0.125) are chosen; otherwise, the log-likelihood no longer is bimodal and less weight is applied—the weights (0.625, 0.375) are used. Maximization is again performed following this new choice of weights. At each step, the maximization problem entails solving (9) twice, with a starting value of $\hat{N}_R$ and again with a starting value of $\hat{N}_S$. The optima from these two subproblems are then compared to determine if they both converge to the same maximum. If they do, further checks are made in a neighborhood of this maximum for differences in signs of the first derivatives of λ^W with these new weights to determine if any minima still exist. The algorithm continues in this fashion, halving the difference in weights between the current and previous values for the weight that represents the favored log-likelihood, and then adding or subtracting this value to the current weights depending on the outcome of the above tests, until convergence. This modified bisection search technique is conceptually simple, but difficult to implement because of the imbedded nonlinear optimizations and the check-sums on previous steps in the weighting algorithm that must be tracked. The program *RemovSim* discussed in Appendix III has an option that performs this weighting if the user desires.

This mixed heuristic (bisection search) and classical (constrained nonlinear) optimization method should find a unimodal combined log-likelihood with global maximum $\hat{N}_C^W$, with minimal weighting for reasonably bimodal initial likelihood functions. For the roach data, weighting in favor of the removal estimator yields the optimal values $(\alpha_1, \alpha_2, \hat{N}_C^W) = (0.6325593, 0.3674407, 81.997)$; alternatively, weighting in favor of the the signs-of-activities estimator yields $(0.498154, 0.501845, 216.470)$. Notice that little weighting is needed to shift the curve toward the "signs-of-activities peak" in the combined log-likelihood because it is the dominant mode in λ for these data (Fig. 2). In addition, when weighting in favor of the signs-of-activities data, the weighted estimate of N remains 216 because there has been little change in the shape of the profile log-likelihood with these small weights. However, more substantial weighting is required in the direction of the "removal peak" because it is an almost nonexistent mode in λ (Fig. 2).

As another example, in the case of the simulated data in Table 1, weighting in favor of the signs-of-activities estimator yields $(\alpha_1, \alpha_2, \hat{N}_C^W)$ values (0.248047, 0.751953, 196.887) because a substantial amount of weight must be applied to the signs-of-activities mode with these data to produce unimodality in that direction. This weighted profile log-likelihood $(\lambda^W(N))$ shown in Figure 4, can be compared to the unweighted profile log-likelihood $(\lambda(N))$ for these data. Note that the weighting algorithm has applied just enough weight to make the "removal mode" disappear, leaving a plateau behind in its stead. With these data, this extra weighting may be intuitively reasonable if the investigator believes that the signs-of-activities data are three times as reliable as the removal data (i.e., $\alpha_2 \approx 3 \times \alpha_1$ here).

The LRT confidence interval procedure (5) and (6) needs to be changed only slightly for the weighted likelihood method (Appendix II shows the derivation for the weighted signs-of-activities LRT statistic following Routledge's asymptotic argument). In general, the test statistic becomes $-2\log\Lambda_C^W/\alpha^* \sim \tilde{F}$, where $\Lambda_C^W = \Lambda_R^{\alpha_1} \times \Lambda_S^{\alpha_2}$ is the likelihood ratio for the weighted combined estimate, and $\alpha^* = \max(\alpha_1, \alpha_2)$. When $\alpha_1 = \alpha_2 = 0.5$, it follows that α^* also equals 0.5 and $-2\log\Lambda_C^W/\alpha^*$ is simply the unweighted version (6) proposed by Routledge for the combined method. When $\alpha_2 > \alpha_1$ (i.e., weighting is in favor of the signs-of-activities estimator), it follows that $\alpha^* = \alpha_2$ and $-2\log\Lambda_C^W/\alpha^* = -2\left(\frac{\alpha_1}{\alpha_2}\log\Lambda_R + \log\Lambda_S\right)$. In this case, Routledge's F-approximation works well since the removal LRT statistic, Λ_R, is down-weighted by $\alpha_1/\alpha_2 < 1$. Similarly, when $\alpha_1 > \alpha_2$, $-2\log\Lambda_C^W/\alpha^* = -2\left(\log\Lambda_R + \frac{\alpha_2}{\alpha_1}\log\Lambda_S\right)$, in which case the F-approximation works well only if the weighting is not excessive. The reason for this is that as $\alpha_1 \to 1$, and $\alpha_2 \to 0$, then the quantity $\frac{\alpha_2}{\alpha_1}\log\Lambda_S \to 0$. Thus, we are left with only the removal likelihood ratio $-2\log\Lambda_R$ which we know is distributed as a $\chi^2(1)$, not $\tilde{F}_\alpha$. Extreme weighting (i.e., α_2 is close to zero), however, is unlikely in practice because the experimenter will not spend the extra effort of collecting signs-of-activities data and, in turn, down-weight the resulting information to near zero! In fact, it is reasonable to assume that, for the combined estimator, weighting in most cases is done in favor of the signs-of-activities estimator. Nevertheless, our weighting scheme provides as the limiting cases the re-

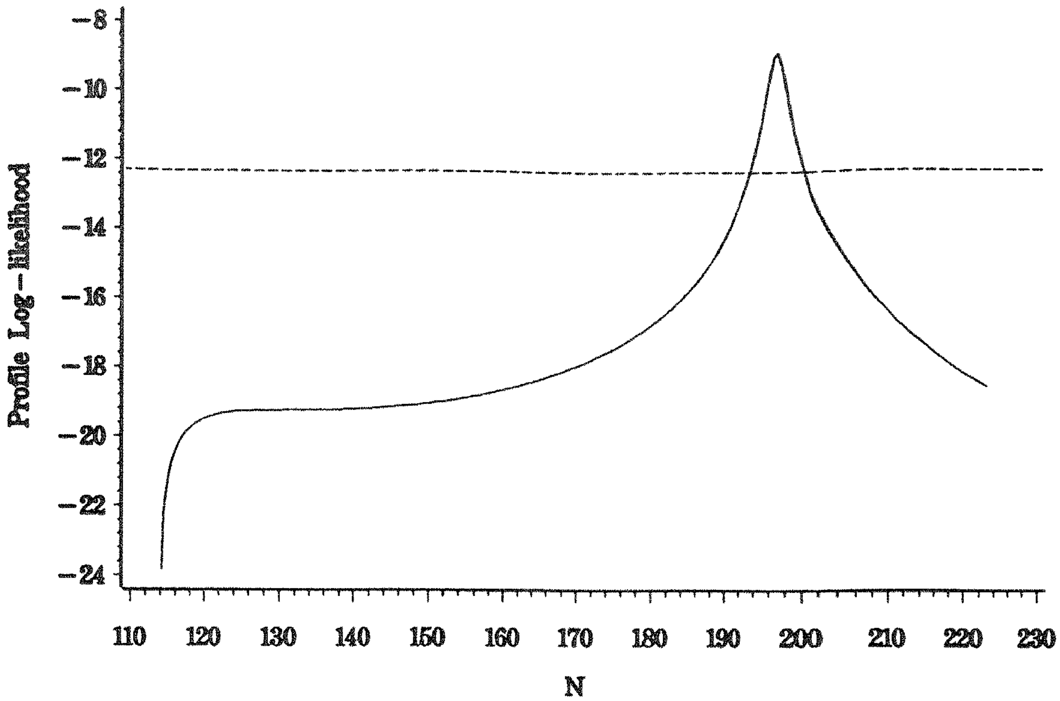


Figure 4.—Profile log-likelihood for weighted combined model using the simulation data set in Table 1. The dashed line shows the 95% confidence level with the LRT statistic at equality.

moval LRT statistic ($-2\log\Lambda_R$) when $\alpha_2 = 0$, and the signs-of-activities LRT statistic ($-2\log\Lambda_S$) when $\alpha_1 = 0$.

Simulation Results

In this section we describe the results of several simulations, the purpose of which were to: 1) investigate how often bimodality occurs under various conditions, and 2) examine the coverage of the confidence intervals corresponding to the weighted combined estimator $\hat{N}_C^W$. We are particularly interested in the effect of unequal capture probabilities for males (p_M) and females (p_F), and the effect of varying standard deviation (σ) in the signs-of-activities model on the coverage.

For each choice of capture probabilities (p_M, p_F) and signs-of-activities standard deviation (σ), we generated 1,000 removal series with corresponding signs-of-activities series for initial population size of $N_M = 114$ males and $N_F = 114$ females. The values for p_M and p_F range between 0.7 and 0.1 while σ is chosen to be 0.4, 0.7 or 1. These parameter values correspond to those in Routledge's simulations while the population size of $N = 228$ is that of the cockroach data. Weighting always is in favor of the signs-of-activities estimator (i.e., $\alpha_2 > \alpha_1$) in all simulations. The percentage frequency of the occurrence of bimodality in the unweighted approach also is reported.

Table 2 includes coverage percentages of nominal 95% confidence intervals for N as described earlier using all four estimators with equal capture probabilities $p_M = p_F = p$. Coverage percentages are uniformly close to the nominal level. In addition, there is no difference in coverage between the weighted and nonweighted approach, because bimodality occurred in only a small percentage of the most extreme case ($p = 0.1$, $\sigma = 1$), as shown in the last column of Table 2. In general, two trends seem to be apparent in Table 2:

Table 2.—Equal probability of capture results of 1,000 simulations for each level of p and σ with $N = 228$, $M = 3$ and $b = 1$; for columns 3, 4, 5, and 6, percent coverage of nominal 95% confidence intervals are given; numbers in parentheses indicate percent of excessively large ($N_{upper} \geq 10,000$) intervals (the last column shows the percentage of combined profile log-likelihoods that were bimodal)

p	σ	Removal estimator	Signs-of-activity estimator	Combined estimator	Weighted combined estimator	Percent bimodality
0.5	0.4	94.8	96.0	95.7	95.7	0.0
		(0)	(0.1)	(0)	(0)	
	0.7	94.0	94.2	94.5	94.5	0.0
		(0)	(0)	(0)	(0)	
	1.0	94.4	94.5	95.3	95.3	0.0
		(0)	(0)	(0)	(0)	
0.25	0.4	96.3	94.1	94.2	94.2	0.0
		(19.0)	(0)	(0)	(0)	
	0.7	95.8	94.7	94.5	94.5	0.0
		(20.9)	(0)	(0.2)	(0.2)	
	1.0	96.7	96.4	96.3	96.3	0.0
		(19.5)	(5.1)	(1.2)	(1.2)	
0.1	0.4	93.4	95.4	95.5	95.5	0.0
		(80.5)	(6.0)	(2.3)	(2.3)	
	0.7	95.7	94.4	94.6	94.6	0.0
		(80.9)	(31.6)	(22.0)	(22.0)	
	1.0	93.9	95.7	95.7	95.7	0.3
		(80.5)	(55.0)	(45.3)	(45.4)	

1. The frequency of excessively large confidence intervals (percentage in parenthesis indicates occurrence of intervals with upper limit $\geq 10,000$) increases with increasing σ at any fixed level of p for all estimators but the removal estimator (since it has no reliance on σ).

2. The frequency of excessively large confidence intervals increases with decreasing p for any fixed level of σ.

The last result is particularly striking for the removal method. The coverage looks good for all values of p until one looks at the frequency of excessively large confidence intervals for $p = 0.1$. In this case, most of the intervals that caught the population size were, for all intents and purposes, infinite. Figure 5 shows the removal profile log-likelihood for a typical sample at $p = 0.1$ with infinite upper confidence interval endpoint. Note that the log-likelihood is flat; the lower confidence interval endpoint is well defined at $N_0 = 62.2$, but the upper endpoint is nonexistent. The other three estimators show similar tendencies for small p but not with the frequency shown by the simple removal method.

Results for unequal capture probabilities ($p_M \neq p_F$) are presented in Table 3. The only apparent trends analogous to the equal probability of capture scenario are:

1. The frequency of bimodal profile log-likelihoods produced tend to increase as the capture probabilities become more disparate.

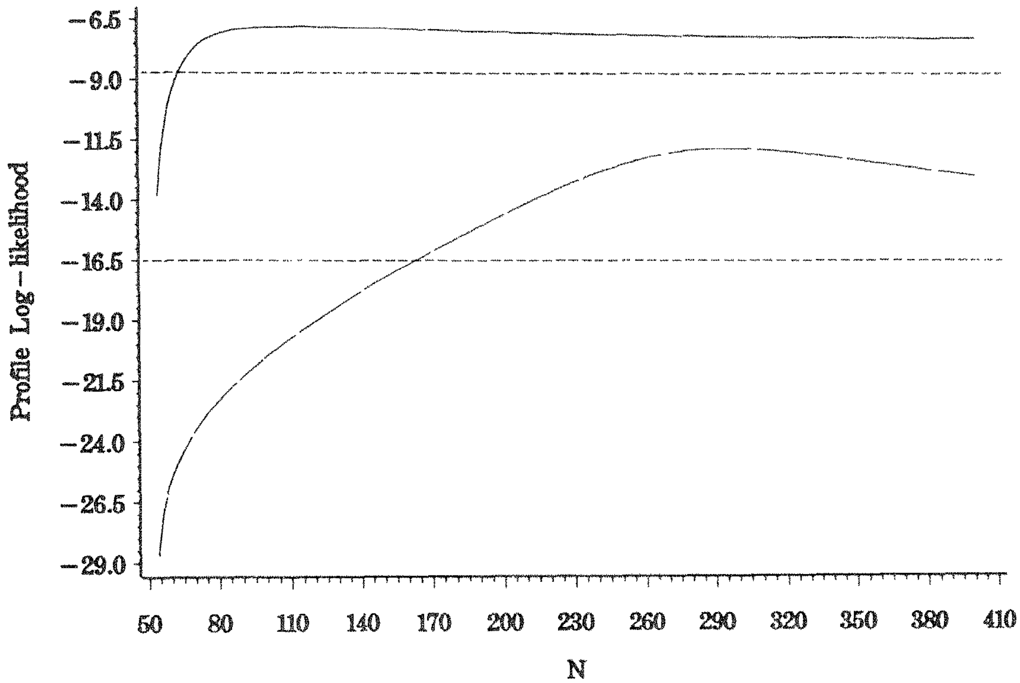


Figure 5.—Removal profile log-likelihood (solid) and signs-of-activities profile log-likelihood (dashed) with associated "infinite" 95% confidence intervals (horizontal dashed lines) for a simulated population with $N = 228$, $M = 3$, $b = 1$, $p = 0.1$, and $\sigma = 0.4$.

2. The frequency of bimodal profile log-likelihoods decreases with increasing σ for the most extreme (p_M, p_F) scenario.

3. There is no difference between the combined and weighted combined coverage until more than 5% bimodal profile log-likelihoods were encountered.

For the most part, those estimators containing signs-of-activities information produced results similar to those in Table 2 except in the most extreme cases where $(p_M, p_F) = (0.7, 0.1)$. In these cases, the signs-of-activities estimator still performed well but the two combined estimators deteriorated with increasing σ. This deterioration is not occurring because of bimodality since the frequency of bimodal profile log-likelihoods decreases with increasing σ; rather, it appears to be more closely linked to the absolute failure of the removal method in these extreme cases. Table 4 shows the reason for this. The removal estimator was successful only twice in 3,000 simulation attempts; most failed with a constant average confidence interval width of approximately 23 individuals. This implies a very peaked profile log-likelihood for the removal method with rapidly decreasing tail, biased on the low side of the true population size. However, the signs-of-activities log-likelihood shows much larger increasing intervals that are unencombered by the downward bias of the removal log-likelihood. When these two log-likelihoods are combined, the characteristics of the removal log-likelihood tend to produce a bimodal combined log-likelihood, and the average width of the confidence intervals for the two combined methods tend to increase only slightly with σ compared to the signs-of-activities intervals. Thus, the combination of the biased removal log-likelihood with the signs-of-activities log-likelihood produces intervals that are too conservative, and, therefore, miss the true population size far too often. Weighting the combined estimator helped, but since the signs-of-activities estimator outperforms both combined estimators in these extreme cases (and performs equally as well in all other cases for for both equal and unequal probability of capture scenarios),

Table 3.—Unequal probability of capture results of 1,000 simulations for each level of p_M, p_F and σ with $N_M = 114$, $N_F = 114$, $M = 3$ and $b = 1$; for columns 3, 4, 5, and 6, percent coverage of nominal 95% confidence intervals are given; numbers in parentheses indicate percent of excessively large ($N_{upper} \geq 10,000$) intervals (the last column shows the percentage of combined profile log-likelihoods that were bimodal)

(p_M, p_F)	σ	Removal estimator	Signs-of-activity estimator	Combined estimator	Weighted combined estimator	Percent bimodality
(0.5, 0.25)	0.4	82.8	95.9	95.7	95.7	0.1
		(0)	(0)	(0)	(0)	
	0.7	83.1	94.8	95.1	95.1	0.2
		(0)	(0)	(0)	(0)	
	1.0	81.3	96.5	96.2	96.2	0.2
		(0)	(0.2)	(0)	(0)	
(0.25, 0.1)	0.4	91.6	94.0	94.3	94.3	0.0
		(49.4)	(0.2)	(0)	(0)	
	0.7	90.7	94.6	94.7	94.7	0.1
		(45.2)	(6.2)	(1.0)	(1.0)	
	1.0	91.8	94.1	94.1	94.1	0.0
		(47.7)	(24.6)	(8.4)	(8.4)	
(0.5, 0.1)	0.4	30.9	95.1	94.7	94.7	5.1
		(0)	(0)	(0)	(0)	
	0.7	30.4	94.0	94.4	94.4	4.5
		(0.3)	(0.4)	(0.1)	(0.1)	
	1.0	34.5	93.5	93.7	93.6	5.3
		(0.1)	(4.5)	(0)	(0)	
(0.7, 0.1)	0.4	0.2	94.8	80.2	92.2	77.6
		(0)	(0)	(0)	(2.4)	
	0.7	0.0	95.6	67.9	78.3	60.1
		(0)	(0)	(0.5)	(2.2)	
	1.0	0.0	95.0	52.9	63.3	38.6
		(0)	(2.0)	(0.1)	(0.9)	

one must question the wisdom of investing the extra effort in either of the combined log-likelihood approaches.

It is interesting to note that the signs-of-activities method itself is actually a combined estimator of sorts because it incorporates both removal and signs-of-activities data into its log-likelihood—the removal data being incorporated through the T_i and $\bar{T}$ terms. However, the signs-of-activities data appear to dominate the removal data in this log-likelihood because the signs-of-activities estimates are particularly insensitive to departures from assumptions in the removal data. Departure from normality for the signs-of-activities data were not studied here, but on the basis of the results of the simulations in Tables 2 and 3, such a study could prove useful and enlightening. In his quest for an estimator that incorporates auxilliary information, Routledge may have unwittingly gone one step too far—his signs-of-activities estimator may fulfill his original goals with the most parsimony.

Table 4.—Mean confidence interval widths for intervals that caught the true population size, and that did not (in parenthesis) for unequal probability of capture results of 1,000 simulations at each level of σ with $(p_M = 0.7, p_F = 0.1)$, $N = 228$, $M = 3$ and $b = 1$; (the last column shows the percentage of bimodal combined profile log-likelihoods that produced disjoint confidence intervals)

σ	Removal width	Signs-of-activity width	Combined width	Weighted combined width	Percent disjoint intervals
0.4	80.8 (22.7)	53.1 (16.2)	78.5 (37.5)	65.5 (24.8)	22.3
0.7	-- (22.6)	110.2 (26.4)	93.4 (50.0)	89.7 (47.0)	12.3
1.0	-- (22.6)	246.5 (42.7)	98.4 (55.2)	98.2 (54.3)	9.3

Both Tables 2 and 3 suggest that when the total probability of capture is small for all individuals, a higher percentage of excessively large confidence intervals is to be expected for all estimators. The removal estimator is particulary sensitive to this; viz., in the equal cases where $p = 0.25$ and 0.1, and the unequal cases with $(p_M, p_F) = (0.25, 0.1)$. However, intervals tend to be excessively large only in the most extreme cases $(p = 0.1)$ in all signs-of-activities-based estimators; though, if σ is small even in these extreme cases, these estimators produce relatively few large intervals. Figure 5 shows that profile log-likelihoods produced from such samples tend to have heavy tails and thus will produce large or "infinite" confidence intervals; in both instances, infinite intervals were encountered. However, as the simulations in Table 2 would suggest, infinite removal and signs-of-activities confidence intervals do not necessarily imply infinite combined intervals. Indeed, in the case of the sample data for Figure 5, the two combined intervals were $[158, 2304]$—both the removal and signs-of-activities methods produced an upper endpoint in excess of 10,000, yet both of the combined intervals were substantially less. This is not always the case, however, as other simulations yielded upper combined endpoints that also were in excess of 10,000. If both the removal and signs-of-activities estimators produce excessively large confidence intervals, then it is more than likely that there is a problem with the data. As a result, the combined methods should not be relied on as a possible "fix" to the situation.

Conclusions

The simulations presented in Tables 2 and 3 suggest that bimodality of the combined log-likelihood is a result of the violation of the assumption of equal capture probability for all individuals. Routledge was aware of this violation in his cockroach experiment and discussed some simulation results (Routledge 1989, p. 118) that are similar to ours. We have extended his observations by pointing out the possibility of bimodality in the combined estimator and excessively large confidence intervals in all estimators under certain scenarios.

The removal method is quite sensitive to the assumption of equal probability of capture and breaks down with increasing frequency as this assumption is increasingly violated. The removal method's contribution to the combined log-likelihood is split equally with the signs-of-activities log-likelihood and thus carries its poor qualities along with it, producing poor results in the combined estimator under the worst scenarios studied. Weighting the combined log-likelihood in favor of the more reliable information found in the signs-of-activities data does not help except in the most extreme cases. The simple signs-of-activities estimator is immune to the effects that violation of the equal catchability assumption has on the removal estimator because it combines the removal data directly through the definition of its likelihood. Thus, the signs-of-activities estimator appears to be not only the most robust with respect to the scenarios developed here, but is also the simplest of the combined information estimators to implement. Finally, all estimators suffer from the undesireable characteristic of producing heavy tails with excessively large confidence intervals when the total probability of capture for all individuals is low; in such cases, the combined estimators perform somewhat better than the signs-of-activities estimator, but the size of the confidence intervals may still be quite large.

As a guide to the interpretation of experiments conducted using the removal and auxiliary information, plotting the profile log-likelihood for reasonable range of N_0 is invaluable for determining possible bimodality and heavy tails associated with the two major problems discussed previously. Spreadsheet programs that facilitate this in an interactive manner are presented in Appendix IV.

Acknowledgment

We thank Dr. Stanley Zarnoch and Dr. Rick Routledge for their very helpful and insightful reviews of this manuscript.

Literature Cited

Fisher, R. A. 1956. **Statistical Methods and Scientific Inference**. Edinburgh: Oliver and Boyd.

Kendall, M. G.; Stuart, A. 1967. **The Advanced Theory of Statistics, Volume 2: Inference and Relationship**, 2nd ed. New York: Hafner.

Lasdon, L. S.; Waren, A. D. 1979. **Generalized reduced gradient software for linearly and non-linearly constrained problems**. In: Greenberg, H. J., ed. Design and Implementation of Optimization Software. Rockville, MD: Sijthoff and Noordhoff International.

Lasdon, L. S.; Waren, A. D. 1986. **GRG2 user's guide**. Austin, TX: University of Texas, Department of General Business, School of Business Administration.

Moran, P. A. P. 1951. **A mathematical theory of animal trapping**. Biometrika. **38**:307–311.

Otis, D. L.; Burnham, K. P.; White, G. C.; Anderson, D. R. 1978. **Statistical inference from capture data on closed animal populations**. Wildlife Monographs. No. 62.

Routledge, R. D. 1989. **The removal method for estimating natural populations: incorporating auxilliary information**. Biometrics. **45**:111–121.

SAS Institute, Inc. 1990. **SAS© Language: Reference, Version 6**, 1st ed. Cary NC: SAS Institute, Inc.

Schnute, J. 1983. **A new approach to estimating populations by the removal method**. Journal of Fisheries and Aquatic Sciences. **40**:2153–2169.

White, G. C.; Anderson, D. R.; Burnham, K. P.; Otis, D. L. 1982. **Capture-recapture and removal methods for sampling closed populations**. Publ. LA-8787-NERP. Los Alamos, NM: Los Alamos National Laboratory.

Wismer, D. A.; Chattergy, R. 1978. **Introduction to Nonlinear Optimization: a Problem Solving Approach**. New York: North-Holland.

Zippin, C. 1956. **An evaluation of the removal method for estimating animal populations**. Biometrics. **12**:163–189.

Appendix I: Likelihood Derivations

This appendix presents a derivation of the removal and signs-of-activities log-likelihood functions presented in Equations 1 and 2, respectively. To our knowledge, a *detailed* derivation of the removal log-likelihood has not been published, though others have outlined it (Moran 1951; Otis et al., 1978; Schnute 1983; Zippin 1956). Nor did Routledge (1989) include a derivation for his signs-of-activities log-likelihood. It seems reasonable to have these derivations available, presented in a step-by-step manner for the interested reader, or as an aid to teaching these methods.

The Removal Likelihood

The removal likelihood comes from the binomial model, but in a conditional sense. To illustrate, on every removal attempt, an individual is either captured or not with constant probability of capture p. Thus, we have a basic binomial model for the first attempt: we remove R_1 individuals out of N total individuals with probability p. However, we have assumed that the population is closed both geographically and demographically (White et al. 1982, p. 3); therefore, on each subsequent removal attempt, the number of individuals *remaining* decreases as illustrated in the following:

Removal Attempt	Model	Probability mass function
1	$R_1 \sim \text{Binomial}(N, p)$	$f(R_1) = \binom{N}{R_1} p^{R_1}(1-p)^{N-R_1}$
2	$R_2 \mid R_1 \sim \text{Binomial}(N - R_1, p)$	$f(R_2) = \binom{N-R_1}{R_2} p^{R_2}(1-p)^{N-R_1-R_2}$
3	$R_3 \mid R_1 + R_2 \sim \text{Binomial}(N - R_1 - R_2, p)$	$f(R_3) = \binom{N-R_1-R_2}{R_3} p^{R_3}(1-p)^{N-R_1-R_2-R_3}$

Note that the quantity $N - R_1 - R_2 - \cdots - R_i$ in the above table can be written as $N - \sum_{j=1}^{i} R_j = N - T_i$ so that, in general, $R_i \sim \text{Binomial}(N - T_{i-1}, p)$ and $f(R_i) = \binom{N-T_{i-1}}{R_i} p^{R_i}(1-p)^{N-T_i}$.

The likelihood for the removal method is given as $\mathcal{L}(\underline{r}; N, p) = f(R_1) \times f(R_2) \times f(R_3) \times \cdots \times f(R_M)$, which becomes

$$\mathcal{L}(\underline{r}; N, p) = \overbrace{\frac{N!}{(N-T_M)! \prod_{i=1}^{M} R_i!}}^{(i)} \overbrace{p^{T_M}}^{(ii)} \overbrace{(1-p)^{(S-T_M)}}^{(iii)}. \tag{10}$$

This may be readily derived by decomposing it into three pieces and working with each piece separately, keeping in mind the conditional nature of the removals given above.

(i) Notice that $\binom{N}{R_1} = \frac{N!}{(N-T_1)!R_1!}$, $\binom{N-T_1}{R_2} = \frac{(N-T_1)!}{(N-T_2)!R_2!}$, $\binom{N-T_2}{R_3} = \frac{(N-T_2)!}{(N-T_3)!R_3!}$, ... This implies that

$$\binom{N}{R_1} \times \binom{N-T_1}{R_2} \times \binom{N-T_2}{R_3} \times \cdots \times \binom{N-T_{M-1}}{R_M} = \frac{N!}{(N-T_1)!R_1!} \times \frac{(N-T_1)!}{(N-T_2)!R_2!} \times$$

$$\frac{(N-T_2)!}{(N-T_3)!R_3!} \times \cdots \times \frac{(N-T_{M-1})!}{(N-T_M)!R_M!} = \frac{N!}{(N-T_M)! \prod_{i=1}^{M} R_i!}.$$

Note that the adjoining numerator and denominator terms $(N - T_i)!$ cancel except in the last term leaving the final result.

(ii) Multiplying the individual values in this component of the likelihood yields: $p^{R_1} \times p^{R_2} \times p^{R_3} \times \cdots \times p^{R_M} = p^{\sum_{i=1}^{M} R_i} = p^{T_M}$.

(iii) Similarly: $(1-p)^{(N-T_1)} \times (1-p)^{(N-T_2)} \times (1-p)^{(N-T_3)} \times \cdots \times (1-p)^{(N-T_M)} = (1-p)^{\sum_{i=1}^{M}(N-T_i)} = (1-p)^{(S-T_M)}$.

Multiplying (i) through (iii) yields the completed likelihood given in (10). Finally, taking the logarithm to the base e of (10) yields the log-likelihood λ_1 in (1). Part of this log-likelihood ($\sum_{i=1}^{M} \log(R_i!)$) is a constant term and need not be included in the maximization process.

The ML estimator for the parameter p may be found in closed-form by taking the partial derivative of λ_1 with respect to p, setting this result equal to zero, and solving for p

$$\frac{\partial \lambda_1}{\partial p} = \frac{T_M}{p} - \left(\frac{S - T_m}{1-p}\right) = 0,$$

or, after a little algebra,

$$\hat{p} = \frac{T_M}{\hat{S}}.$$

The Signs-of-Activities Likelihood

It was noted earlier that the signs-of-activities model derives from the normal distribution with mean $\mathrm{E}[Y_i] = b(N - T_i)$ and variance $\mathrm{Var}(Y_i) = \sigma^2(N - T_i)$; that is, $Y_i \sim \mathrm{Normal}(\mathrm{E}[Y_i], \mathrm{Var}(Y_i))$. Given this probability model, the likelihood is

$$\begin{aligned} \mathcal{L}(\underline{y}; b, \sigma^2, N) &= \prod_{i=0}^{M} \frac{1}{\sqrt{2\pi \mathrm{Var}(Y_i)}} \exp\left(\frac{-(Y_i - \mathrm{E}[Y_i])^2}{2\mathrm{Var}(Y_i)}\right) \\ &= (2\pi)^{-\frac{(M+1)}{2}} \left[\prod_{i=0}^{M} \mathrm{Var}(Y_i)^{-\frac{1}{2}}\right] \exp\left(-\sum_{i=0}^{M} \frac{(Y_i - \mathrm{E}[Y_i])^2}{2\mathrm{Var}(Y_i)}\right). \end{aligned} \tag{11}$$

Taking the logarithm to the base e of this expression yields

$$\lambda_2 = \overbrace{-\frac{(M+1)}{2}\log(2\pi)}^{(i)} \overbrace{- \frac{1}{2}\sum_{i=0}^{M} \log \mathrm{Var}(Y_i)}^{(ii)} \overbrace{- \frac{1}{2}\sum_{i=0}^{M} \frac{(Y_i - \mathrm{E}[Y_i])^2}{\mathrm{Var}(Y_i)}}^{(iii)}.$$

Again, as in the removal likelihood, some algebra must be done on each of the three components of the signs-of-activities log-likelihood given above to arrive at (2).

(ii) Substituting $(N - T_i)\sigma^2$ for $\mathrm{Var}(Y_i)$

$$= -\frac{1}{2}\sum_{i=0}^{M} \log\left((N - T_i)\sigma^2\right)$$

$$= -\frac{1}{2}\sum_{i=0}^{M} \log(N - T_i) - \frac{(M+1)}{2}\log(\sigma^2).$$

(*iii*) Again, substituting for $\mathrm{Var}(Y_i)$ as in (*ii*), and substituting $b(N - T_i)$ for $\mathrm{E}[Y_i]$ yields

$$= -\frac{1}{2\sigma^2}\sum_{i=0}^{M} \frac{(Y_i - b(N - T_i))^2}{(N - T_i)}.$$

Now combining (*i*) with the second term in (*ii*) yields $-\frac{(M+1)}{2}\log(2\pi\sigma^2)$, and adding this expression to what is left in (*ii*) and (*iii*) gives the expression for λ_2 in (2). As in the removal log-likelihood, the last term in (2) is a constant term and may be disregarded for maximization.

The ML estimators of b and σ^2 can be found in closed-form by taking the partial derivatives of λ_2 with respect to each of these parameters, setting these equal to zero, and solving for each parameter. First, find the MLE for b

$$\frac{\partial\lambda_2}{\partial b} = \sum_{i=0}^{M} (Y_i - b(N - T_i)) = 0,$$

or

$$\sum_{i=0}^{M} Y_i = b(M+1)(N - \overline{T}),$$

and recalling the definition of $\bar{Y}$ we get

$$\hat{b} = \frac{\overline{Y}}{(\hat{N} - \overline{T})}.$$

Finally, the MLE of σ^2 is found by noting that $\log(2\pi\sigma^2) = \log(2\pi) + \log(\sigma^2)$ in λ_2

$$\frac{\partial\lambda_2}{\partial\sigma^2} = -\frac{(M+1)}{\sigma} + \frac{1}{\sigma^3}\sum_{i=0}^{M} \frac{(Y_i - b(N - T_i))^2}{N - T_i} = 0,$$

or

$$\hat{\sigma}^2 = \frac{1}{M+1}\sum_{i=0}^{M} \frac{(Y_i - \hat{b}(\hat{N}_S - T_i))^2}{\hat{N}_S - T_i}.$$

Appendix II: Asymptotic F Extension for the Weighted Signs-of-Activities LRT

This appendix outlines the justification of the use of Routledge's asymptotic transformed F-distribution for the weighted signs-of-activities likelihood ratio test through derivation of the likelihood ratio test statistic. Consider the weighted signs-of-activities likelihood, which is simply the signs-of-activities likelihood given in (11) raised to the α_2 power, this can be written as

$$\mathcal{L}(\underline{y}; N, b, \sigma^2)^W = \left[(2\pi)^{-\frac{(M+1)}{2}}\left[\prod_{i=0}^{M}\mathrm{Var}(Y_i)^{-\frac{1}{2}}\right]\exp\left(-\sum_{i=0}^{M}\frac{(Y_i - \mathrm{E}[Y_i])^2}{2\mathrm{Var}(Y_i)}\right)\right]^{\alpha_2}$$
$$= (2\pi)^{-\frac{\alpha_2(M+1)}{2}}\left[\prod_{i=0}^{M}\mathrm{Var}(Y_i)^{-\frac{\alpha_2}{2}}\right]\exp\left(-\frac{\alpha_2}{2}\sum_{i=0}^{M}\frac{(Y_i - \mathrm{E}[Y_i])^2}{\mathrm{Var}(Y_i)}\right).$$

Now, using the last expression for the weighted signs-of-activities likelihood, and substituting $b(N - T_i)$ for $\mathrm{E}[Y_i]$ and $(N - T_i)\sigma^2$ for $\mathrm{Var}(Y_i)$, we can write the likelihood ratio for the weighted signs-of-activities model as

$$\Lambda_S^W = \frac{\mathcal{L}(\underline{y}; N_0, b, \sigma^2)^W}{\mathcal{L}(\underline{y}; \hat{N}_S, b, \sigma^2)^W} = \frac{(2\pi)^{-\frac{\alpha_2(M+1)}{2}}\left[\prod_{i=0}^{M}\left((N_0 - T_i)\sigma_0^2\right)^{-\frac{\alpha_2}{2}}\right]\exp\left(-\frac{\alpha_2}{2}\sum_{i=0}^{M}\frac{\left(Y_i - b_0(N_0 - T_i)\right)^2}{(N_0 - T_i)\sigma_0^2}\right)}{(2\pi)^{-\frac{\alpha_2(M+1)}{2}}\left[\prod_{i=0}^{M}\left((\hat{N}_S - T_i)\hat{\sigma}^2\right)^{-\frac{\alpha_2}{2}}\right]\exp\left(-\frac{\alpha_2}{2}\sum_{i=0}^{M}\frac{\left(Y_i - \hat{b}(\hat{N}_S - T_i)\right)^2}{(\hat{N}_S - T_i)\hat{\sigma}^2}\right)}.$$

However, we see in the exponential expressions above that

$$\sigma_0^2 = \left(\frac{1}{M+1}\right)\sum_{i=0}^{M}\frac{\left(Y_i - b_0(N_0 - T_i)\right)^2}{(N_0 - T_i)} \quad \text{and} \quad \hat{\sigma}^2 = \left(\frac{1}{M+1}\right)\sum_{i=0}^{M}\frac{\left(Y_i - \hat{b}(\hat{N}_S - T_i)\right)^2}{(\hat{N}_S - T_i)},$$

which, when substituted into the last expression for Λ_S^W above, yields

$$\frac{(2\pi)^{-\frac{\alpha_2(M+1)}{2}}\left[\prod_{i=0}^{M}\left((N_0 - T_i)\sigma_0^2\right)^{-\frac{\alpha_2}{2}}\right]\exp\left(-\frac{\alpha_2}{2}(M+1)\right)}{(2\pi)^{-\frac{\alpha_2(M+1)}{2}}\left[\prod_{i=0}^{M}\left((\hat{N}_S - T_i)\hat{\sigma}^2\right)^{-\frac{\alpha_2}{2}}\right]\exp\left(-\frac{\alpha_2}{2}(M+1)\right)}$$

after cancellations. Now everything in this expression except for the quantities enclosed in brackets $[\cdots]$ cancels. After some algebra, this leaves

$$\Lambda_S^W = \left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right)^{\frac{\alpha_2}{2}(M+1)}\prod_{i=0}^{M}\left(\frac{\hat{N}_S - T_i}{N_0 - T_i}\right)^{\frac{\alpha_2}{2}}.$$

Finally, taking the logarithm of the above quantity yields

$$\log \Lambda_S^W = \frac{\alpha_2}{2}\left[(M+1)\log\left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right) + \sum_{i=0}^{M}\log\left(\frac{\hat{N}_S - T_i}{N_0 - T_i}\right)\right],$$

or

$$W^W = -2\log\Lambda_s^W/\alpha_2 = -(M+1)\log\left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right) - \sum_{i=0}^{M}\log\left(\frac{\hat{N}_S - T_i}{N_0 - T_i}\right).$$

But the right-hand side is the same as (A.1) in Routledge (1989, p. 120). Therefore, the asymptotic argument given by Routledge for his W in (A.1) was simply applied to W^W for our confidence intervals in the weighted case.

Appendix III: Simulation Programs and Related Files

This appendix discusses the use of the computer programs *RemovSim* and *CiSim* for simulating removal and signs-of-activities experiments. *RemovSim* is used to generate as many simulated combined removal and signs-of-activities experiments as are desired by the user, or to process existing data from such combined experiments. *RemovSim* produces an output file with statistics from these experiments, including estimates of population size from the removal, signs-of-activities, combined, and weighted combined estimators, and prints detailed reports on individual simulations.

CiSim reads an input file created by *RemovSim* and determines the confidence intervals for each of the population estimates produced by *RemovSim*. *CiSim* also produces a detailed output file of the confidence interval results. These results can be analyzed for statistics such as the percent of times the confidence intervals have included the true population size by estimator, with a set of SAS[1] (1990) programs.

RemovSim and *CiSim* are written in FORTRAN for use on personal computers and have been compiled and linked with Microsoft FORTRAN. Both programs use a set of proprietary nonlinear optimization subroutines known as GRG2 (Lasdon and Waren 1986) to maximize the log-likelihoods and search for confidence interval endpoints.

RemovSim

Fiure 6 shows a sample run of *RemovSim* in which 10 simulated removal and signs-of-activities experiments are generated. *RemovSim* asks the user several questions; answers given to some questions determine further questions that are asked of the user and the amount of output that is generated. All default answers in *RemovSim* are shown enclosed in angle brackets (<>). The following discusses the questions and answers for the sample run in Figure 6.

- The title of the run is only used when extra output is requested; if no extra output is desired (see below), type a carriage return.
- The output file will hold the simulated removal and signs-of-activities experiments and is used as an input file to *CiSim*.

[1] The use of trade, firm, or corporation names in this report is for the information and convenience of the user. Such use does not constitute an official endorsement or approval by the U.S. Department of Agriculture or Forest Service of any product or service to the exclusion of others that may be suitable.

```
D:\REMOVAL\SIMULAT>removsim

RemovSim:  simulates removal & signs-of-activities experiments
           for closed populations...

Enter a title...
Example of equal capture probabilities

Enter the output file name...
equal.ind

GRG2 reports are only generated if you choose no weighting and
the # of simulations is <=10...
    >>Print all final GRG2 reports (Y/<N>)?

Do you want to weight by Removal, Signs or <N>one?

Run <s>imulations or (o)ne experiment?

***>Two types of populations may be simulated...
     (1) For ONE population with all individuals sharing the same p(capture),
         enter (p,N) for males ONLY, and enter p=0, N=0 for females.
     (2) For two subpopulations (m,f) with differing p(capture)s, enter
         different values for (p,N) in the next two questions.  For equal sex
         ratios, make N for males and N for females both equal to half of the
         population size...

Enter p and N for MALES: .4 100
Enter p and N for FEMALES: 0 0

You have chosen to draw samples from ONE large population of size:  100
with p(capture) = .4000

Enter # of removals (2<=M<=9) <3>:
Enter sigma and b:  .4 1
Do you wish to fix SIGMA <n>?
Do you wish to fix b <n>?

Enter the number of simulations to run [1,1000] < 10>:
Enter the seeds for binomial and normal variates as follows...
        if SEED < 0 then SEED is used as is;
        if SEED = 0 then <Default> is used;
        if SEED > 0 then clock time is used.
Enter the binomial seed <-12345>:
Enter the normal seed <-12345>:
Sim #:  5
Sim #:  10

(* Execution time = 4.83000 seconds *)
Stop - Now run CiSim for CIs.
```

Figure 6.—Sample run of the program *RemovSim* with no weighting of bimodal likelihoods, and all individuals in the population assumed to have the same capture probability.

- GRG2 reports are technical, covering the results of the nonlinear optimizations on estimating the population size; their description is beyond the scope of this report. Lasdon and Waren (1986) should be consulted for this information along with the source code for *RemovSim*. Because this option generates much output, you are limited to up to 10 experiments. No GRG2 reports are generated if weighting is desired because weighting is itself an iterative algorithm that calls GRG2 (also an iterative algorithm) many times to determine the optimal weights; thus, huge amounts of output are possible without this restriction. If reports are printed, they are written to the file "`REPORT.DMP`." The default answer to this question (no) was taken by the user in this case.

- The user may choose to employ the weighting algorithm described in the text, or do no weighting of the combined log-likelihood. If weighting is desired, it will be implemented toward either the removal peak or the signs-of-activities peak for bimodal log-likelihoods. This condition applies to all of the bimodal likelihoods generated in an entire simulation run of *RemovSim*. The default of not weighting the bimodal log-likelihoods was chosen.

- Here you may choose to simulate combined removal and signs-of-activities experiments, or to process the data from a removal and signs-of-activities experiment that already has been conducted. The user took the default answer here because simulations were desired.

- The comments after the "***>" explain that the following questions allow the user to generate samples from one homogeneous population, or from two groups with differing capture probabilities or numbers of individuals within the population. This option is useful for looking at the behavior of the estimators in the case of differing capture probabilities for male and female members of the population as discussed in the section on simulation results, or when there is an unequal sex ratio in the population, or both.

- This example simulation run generated experiments from one population where all of the members are assumed to have equal probability of capture. The next two questions determine this: there are 100 individuals in the population that are assigned a probability of capture of 0.4 through the first question pertaining to males, and the population characteristics for females are entered as zeros. In *RemovSim*, if population information is entered only for males, the simulations will be drawn from one population. If information is entered for both males and females, two separate but coexisting groups of individuals are assumed within the population. In this latter case, removals and signs-of-activities are generated for M removals from both groups independently, and are subsequently added together for the total population values of R_i, Y_i, T_i, etc., which are then used to evaluate the log-likelihoods for parameter estimation. *RemovSim* echos back what you have chosen for capture probabilities and population size after you have answered the two questions cited.

- M is the number of removals desired in any one experiment as discussed earlier. *RemovSim* allows up to nine removals (implying 10 signs-of-activities observations); the default of three was chosen by the user.

- Here the user entered $\sigma = 0.4$ and $b = 1$. These quantities are used as parameters in the generation of normal random variates for the simulation of signs-of-activities experiments. Recalling that $Y_i \sim \text{Normal}(b(N - T_i), \sigma^2(N - T_i))$, we see that σ and b are used to transform standard normal random variates to normal random variates with mean $b(N - T_i)$ and variance $\sigma^2(N - T_i)$ in the generation of the Y_i.

- The formulas for σ and b found in the body of this report show that these quantities depend on the removal and signs-of-activities data through the T_i and the Y_i, respectively. However, due to the stochastic nature of removal and signs-of-activities experiment generation, σ and b will not necessarily equal the values entered in the previous question for each individual experiment. The following two questions allow the user to override the calculation of the quantities σ and b from the sample data so that they are fixed at the quantities entered previously for all simulations in the run. These fixed quantities are used instead of the data-based values for σ and b in the evaluation of the log-likelihood and subsequent estimation of N. By choosing the defaults here, the user elects not to fix these quantities but to take the values of σ and b as calculated from the individual simulations.

- The number of removal and signs-of-activities experiments to be generated for Monte Carlo evaluation of the estimators is entered here. The default of 10 is unrealistic for any evaluation of estimator precision and accuracy; at least 500 should be used. It was set low here as a default so that users could look at small output files to get a feel for the results before committing their personal computers to *RemovSim* for longer runs. A run of 500 simulations may take from 10 minutes to well over an hour depending on the type of computer used, the number of bimodal likelihoods, and whether or not weighting of bimodal likelihoods is desired.

- There are three options available for the choice of the initial random number seeds for both the binomial and normal random variate generators (both of these rely on two distinct uniform random number generators). If a negative number is entered, its absolute value is used as the seed; if zero is entered, -12345 is used as the seed; finally, if a positive number is entered, the computer's clock time is used as the initial seed. Both seeds always are written to the output file so that even if the clock time is used, you always will be able to duplicate simulation runs. In this run, the default random number seeds were chosen for each distribution.

- *RemovSim* simply prints a message every 5th simulation to let you know that something is happening. When all of the simulations have been completed, the run time in seconds is printed and a reminder is given that confidence intervals still need to be constructed with the program *CiSim*.

If it is desirable to weight the bimodal log-likelihoods in favor of either the removal peak or the signs-of-activities peak, the following question would be answered by responding with either an `r` or `s` for removal or signs-of-activities, respectively.

```
Do you want to weight by Removal, Signs or <N>one?  r
```

When weighting is desired, two other questions will be posed to the user:

```
***>Next 2 questions pertain to binary search algorithm...
    Enter a tolerance value < .10D-03>:
    Enter the number of iterations [1,50] <20>
    (negative # to print details):
```

The first question asks for a tolerance value to be used to terminate the binary search algorithm applied to the weighting of any bimodal log-likelihoods. After numerous runs of *RemovSim* by the authors, it appears that the default value given is a reasonable value, assuring that all bimodal log-likelihoods checked become unimodel in the fewest number of iterations. The second question asks for the maximum number of iterations that *RemovSim* will use in the weighting algorithm to turn a bimodal log-likelihood into a unimodal one. Again, after much experimentation, a value of 20 is seldom exceeded in the binary search algorithm. We recommend that the defaults be chosen for these two questions until you are comfortable with *RemovSim*. If the last question is answered as a negative number, the absolute value is taken as the maximum number of iterations, and the output file produced has the output presented in more readable form. However, output produced with this option *must not* be used as input into *CiSim* because it is formatted incorrectly.

Finally, *RemovSim* may be run not as a simulation program but as a processing program for combined removal and signs-of-activities experiments that already have been conducted. To process one experiment with *RemovSim*, the data must be entered into an input file in pairs of (R_i, Y_i)—one per record in "free" format. For example, the following illustrates an example input file called **mjg29.fil**:

```
0, 219
96 132
32 104
18 85
```

In this experiment, $M = 3$ removals were done with $(R_0, Y_0) = (0, 219)$ and $(R_M, Y_M) = (18, 85)$. Processing these data with *RemovSim* is simple; rather than accepting the default answer when asked whether simulations or one experiment is desired, answer as follows:

```
Run <s>imulations or (o)ne experiment?  o

What is the file name for the experimental data?
mjg29.fil
```

With these responses, *RemovSim* will process the data from the single experiment, use weighting if desired, and write the results to the output file as entered earlier in the run (in the case of Figure 6, the results will be stored in equal.ind). This one experiment can then be processed for confidence intervals with *CiSim*.

CiSim

Figure 7 shows a sample run of *CiSim* using the results of the 10 simulated removal experiments from the run of *RemovSim* in Figure 6. The sample run is discussed in the following.

- The note at the beginning of the run informs the user that *CiSim* requires a file by the name of "**IWANT.KEY**" to be present in the current user directory. **IWANT.KEY** is a simple file consisting of removal experiment numbers, one per line, that act as a *key* into the output file from *RemovSim*. As an example, if the user wanted to calculate confidence intervals for experiments 2 and 7 in the file **EQUAL.IND** generated in *RemovSim*, then **IWANT.KEY** would have two records and would look like the left-hand box in the illustration that follows. However, if the user later decided to process all 10 of the experiments in **EQUAL.IND**, then **IWANT.KEY** might look like the rightmost box, or might contain only one record, with zero in the leftmost column of that record. Thus, if a zero is found in the first record, the entire file is processed regardless of what might be in subsequent records.

2	...1st line...	0
7	...2nd line...	2
	...3rd line...	7

- As in *RemovSim*, the title is asked for first and is only used if extra output is requested (see below). A carriage return will default to a blank title.

```
D:\REMOVAL\SIMULAT\RUNS>cisim

CiSim:  Confidence interval calculation for removal & signs-of-activities
        experiments simulated in RemovSim...

***>Please note that the file named IWANT.KEY may be used to selectively
    process individual populations from the desired input file.  Just put
    the population numbers you want to process, one per record, in this
    file.  If you want to process all populations, place a zero in the
    first record of the file.

Enter a title...
Example of equal capture probability confidence intervals

Enter the input file name (from RemovSim)...
equal.ind

Enter the output file name...
equal.ici

Print all final GRG2 reports (Y/<N>)?  y
Enter the following for likelihood ratio CIs...
        alpha level < .05>:
        F value <18.51281883>:
        Chi-square value <3.841455221>:

***>GRG2 output will not be printed; to print details, choose
    the individual simulations desired with IWANT.KEY!

Simulation #:  5
Simulation #:  10

(* Execution time = 6.53000 seconds *)
Clean completion!
```

Figure 7.—Sample run of the program *CiSim* using the output file generated in *RemovSim* (see Figure 6).

- The input file name requested must be a file that was created by *RemovSim* as explained in the description of the program *RemovSim*.

- The output file holds the results of the confidence interval estimates for each experiment in the input file that matches an experiment number listed in `IWANT.KEY` (or all experiments). The output file is in a format that can be used directly by the SAS summary programs (discussed later).

- You may print the detailed GRG2 optimization reports only if you choose individual simulations using `IWANT.KEY`. There is far too much output that will be generated with this option for anything but a small number of experiments. Even if you answer in the affirmative ("y") to this question, the program checks to see if you have requested processing of all simulations. If you have, it overrides your answer and does not generate a detailed report. This should be satisfactory as details normally are desired for only a small number of simulation experiments. When generated, the detailed report is output to a file by the name of "`CIREPORT.DMP`."

- *CiSim* does not have built-in functions for calculating critical values of the F and chi-square distributions used in the calculations of the likelihood ratio confidence interval; therefore, these must be entered by the user from tables or other sources. *CiSim* queries the user for 1) the α level pertaining to the F and χ^2 critical values that will be entered; 2) the critical F value (the $1-\alpha$ quantile of the $F(1, M-1)$ distribution) used in the calculation of $\tilde{F}_\alpha$ in Equation 4; 3) the chi-square value (the $1-\alpha$ quantile of the $\chi^2(1)$ distribution) used for finding the confidence intervals for the removal estimates using the likelihood ratio method described earlier. In this run the user has chosen the default answers to calculation of confidence intervals. The default answers always produce 95% confidence intervals for each of the simulation experiments in the input file created by *RemovSim*, if defaults have been chosen there as well.

- The note ("***>...") given by *CiSim* simply catches the user's request for detailed output when all simulations have been elected to be processed in `IWANT.KEY`. Although there are only 10 experiments in the input file, this is still flagged as illegal.

- *CiSim* notifies the user as every 5th simulation experiment is processed and then prints the total execution time. Execution time for *CiSim* normally is greater than that for *RemovSim* because at least two confidence interval endpoints must be found with GRG2 for every likelihood estimate found in *RemovSim*. If bimodal likelihoods are common in the experiments, then up to four endpoints are estimated using GRG2.

Output Files

RemovSim generates an output file with eight records per simulation experiment. Table 5 details the individual records and fields within the records for all of the variables printed by *RemovSim* with the exception of records three through seven. These records have the same fields as record two but with analogous information for different estimators than the removal estimator. The differences are presented in the following list.

- Record three contains analogous information for the signs-of-activities estimator. The field differences are: **3**[1] The constant term in the signs-of-activities log-likelihood is

$$\sigma^{-2}\sum_{i=0}^{M}\frac{(Y_i - b(N-T_i))^2}{N-T_i},$$

  which reduces to $1/2(M+1)$; **3**[4] the code "s" for signs-of-activities; and, **3**[7] & **3**[8] are 0.0 and 1.0 for α_1 and α_2, respectively.

- Record four is for the combined estimator, with starting value equal to $\hat{N}_R$. The field differences are: **4**[1] both constant terms in **2**[1] and **3**[1] are excluded; **4**[4] the code "c" for combined; **4**[6] the constant code "r" signifying that the starting value was $\hat{N}_R$; **4**[7] and **4**[8] are both equal to one.

- Record five is the same as record four except that the starting value is $\hat{N}_S$ rather than $\hat{N}_R$. Consequently, the only field that changes is **5**[6], with code "s."

- Record six is the same as records four and five except that the starting value is the true, unknown population size; thus, **6**[6] now has code "p."

- Record seven holds the weighted combined estimate. In this case, because of the workings of the bisection search algorithm, both $\hat{N}_R$ and $\hat{N}_S$ are used as starting values. Therefore, field **7**[6] simply contains "w," signifying the weighted estimate. Fields **7**[7] and **7**[8] will contain the actual optimal weights for α_1 and α_2 in the case of a bimodal log-likelihood. If the combined log-likelihood was not found to be bimodal, all fields except for **7**[4] and **7**[6] will be zero.

Different starting values are used in records four through six for each *RemovSim* simulation experiment to facilitate the identification of bimodal log-likelihoods. The individual modes in a bimodal combined log-likelihood are closely associated with the two independent peaks of the removal and signs-of-activities log-

Table 5.—Description of the records and fields associated with each simulated removal and signs-of-activities experiment generated by *RemovSim*

Record [Field][a]	Columns	Contents of field
1 [1][a]	1	Contains either `r`, `s`, or `n`, depending on whether *RemovSim* has weighted any bimodal combined log-likelihoods by the removal, signs-of-activities, or no weighting, respectively.
[2][a]	2–4	Probability of capture for males (note that if 0 is entered in **1**[4], this is the probability of capture for *all* individuals in the population).
[3][a]	5–9	Total number of males in the population (note that if 0 is entered in **1**[5], this is the total number of *all* individuals in the population).
[4][a]	10–12	Probability of capture for females. This may be 0 if the user has selected a homogeneous population (see **1**[2]).
[5][a]	13–17	Total number of females in the population. This may be 0 if the user has selected a homogeneous population (see **1**[3]).
[6][a]	18–21	The value of σ used for the signs-of-activities normal random variate generator.
[7][a]	22–25	The value of b used in the signs-of-activities normal random variate generator.
[8][a]	26–29	The number of experiments simulated by the run of *RemovSim* in the file.
[9][a]	30–31	The number of removal attempts (M).
[10][a]	32–37	The *initial* random number seed for the binomial random number generator.
[11][a]	38–43	The *initial* random number seed for the normal random number generator.
[12]	44–49	The value of T_M for this simulation.
[13][a]	50–52	The maximum number of iterations that may be used for the weighting algorithm.
[14]	53–55	The actual number of iterations used in the weighting algorithm for the current simulation experiment.
[15][a]	56–61	The convergence tolerance used in the weighting algorithm to judge convergence to unimodality.
[16][a]	62	`True` if σ is fixed; `False` otherwise.
[17][a]	63	`True` if b is fixed; `False` otherwise.
[18][a]	64	`True` if a homogeneous population was chosen (i.e., **1**[4] & **1**[5] are both 0); `False` otherwise.
[19]	65	This field relates to the weighting algorithm. If, after the weighting algorithm has finished this is still `True`, then there were not enough iterations specified in **1**[13] for convergence. If **1**[21]is `False`, then this also should be `False`.
[20]	66	This field will be `True` if a very large signs-of-activities estimate was generated. This implies that the signs-of-activities log-likelihood is very *flat*, and, therefore, also implies that the combined log-likelihood should be unimodal.
[21]	67	`True` if a bimodal combined log-likelihood; `False` otherwise.
[22]	68	In very rare cases there can actually be a very small second mode in the signs-of-activities log-likelihood very close to the T_M boundary. If this is the case, this field will be `True`; otherwise it will be `False`.
[23][a]	71–76	This is the true (but unknown!) population size.
[24]	77–80	The simulation experiment number.
2 [1]	1–22	The value of the log-likelihood for the removal estimator without the constant $\sum_{i=1}^{M} \log(R_i!)$ included.
[2]	23–44	The value of the complete removal log-likelihood with constant included.
[3]	45–57	The removal estimate ($\hat{N}_R$) for N.
[4][a]	58	A constant code "(`r`)" signifying the removal estimate.
[5]	59	This field holds the termination message from GRG2 for the final solution in the maximization. Values of 0–2 imply that GRG2 has found a good solution with 0 being the highest level of confidence through meeting the Kuhn-Tucker Conditions. Any value 3–5 should not be trusted as an estimate of N that has met the convergence criteria (Lasdon and Waren 1986).

Table 5.—*Continued*

Record [Field][a]	Columns	Contents of field
[6][a]	60	A constant code ("p") signifying that the initial starting value for maximization of the removal log-likelihood is the true, unknown, population size (1[23]).
[7][a]	61–70	The value for $\alpha_1 \equiv 1.0$ used in the maximization of the log-likelihood. All maximizations actually use the combined log-likelihood in the form of equation set (9) so that using 1 here and 0 in 2[8] yields the removal log-likelihood.
[8][a]	71–80	The value for $\alpha_2 \equiv 0.0$ (see 2[7] for details).
3–7		These records are the same format as record 2, the differences are explained in the text.
8 [1][a]	1–4	$R_0 \equiv 0.0$ always.
[2]	5–9	Y_0.
	10–90	(R_i, Y_i) pairs in FORTRAN (`f4.0, f5.0`) format for $i = 1$ to M.

[a]These fields remain constant for every simulation experiment generated in the run of *RemovSim*.

likelihoods. Figure 8 shows this correspondence. The smaller peak in the combined log-likelihood corresponds to the peak in the removal log-likelihood; similarly, the global peak in the combined log-likelihood corresponds to the signs-of-activities log-likelihood peak. Therefore, using $\hat{N}_R$ and $\hat{N}_S$ as starting values in the search for maxima in the combined log-likelihood is one method of determining bi- or unimodality of that log-likelihood. If the two starting values converge to different maxima while satisfying the convergence conditions in GRG2, then the log-likelihood is bimodal; similarly, if they converge to the same maxima, then the likelihood is unimodal. Beginning the search from N is another check that may be made on bimodality. This procedure is used in *RemovSim* to determine bimodality, it has been tested against numerous profile log-likelihoods as a check and appears to be a very reliable method for determining bimodality among the methods tested. These different estimates also are used in confidence interval construction in *CiSim*.

The output from *CiSim* is in the form of a file with the first record being a header record, with sets of 10 records per simulation experiment thereafter. The details of this record structure are in Table 6. Similarly to the *RemovSim* record structure, the first record in each group contains general information about the experiment; each subsequent record or pair of records contains confidence interval endpoint information and log-likelihood values for each of the different estimates. For example, records 6 and 7 contain up to four confidence interval endpoints for the combined log-likelihood estimate of N with starting value of $\hat{N}_S$. This corresponds to the estimate found in record 5 of the *RemovSim* output file. Notice that this direct correspondence of order between records in the *RemovSim* file and the *CiSim* file is consistent for the entire experiment. Also, because a bimodal log-likelihood may have up to four confidence interval endpoints, all of the *CiSim* records that correspond to one of the unweighted combined estimates in the *RemovSim* file use two records to store this information. If the confidence intervals are not disjoint for a bimodal log-likelihood, then the first and the fourth fields contain the endpoints; e.g., for records 6 and 7, this information is in 6[1] and 7[3]. In the case of disjoint intervals, all four fields are used.

As with the different estimates in *RemovSim*, it may not be immediately obvious why endpoints need to be calculated multiple times for the same unweighted combined log-likelihood. The answer lies in fixing attention on the unweighted combined log-likelihood in Figure 8. Notice that the two peaks are not of equal value in λ— the right-most peak corresponding to the estimate arrived at by using $\hat{N}_S$ as the starting value, has a higher value of λ than the other peak. Therefore, one can calculate the difference in Equation 6 from each peak, and arrive at *different* confidence interval endpoints each time, providing that the two peaks have different maximum values of λ as in Figure 8. In this example, the left-most peak has value $\lambda(\hat{N}_{C(\hat{N}_R)}) = -30.19615$ with corresponding likelihood ratio endpoints at $\lambda(N_{0(\hat{N}_R)}) = -34.85196$, yielding only two confidence interval endpoints [148, 257] (the lower horizontal line in Figure 8). However, the right-most peak has value $\lambda(\hat{N}_{C(\hat{N}_S)}) = -27.11697$ with corresponding likelihood ratio endpoints at $\lambda(N_{0(\hat{N}_S)}) = -31.77278$,

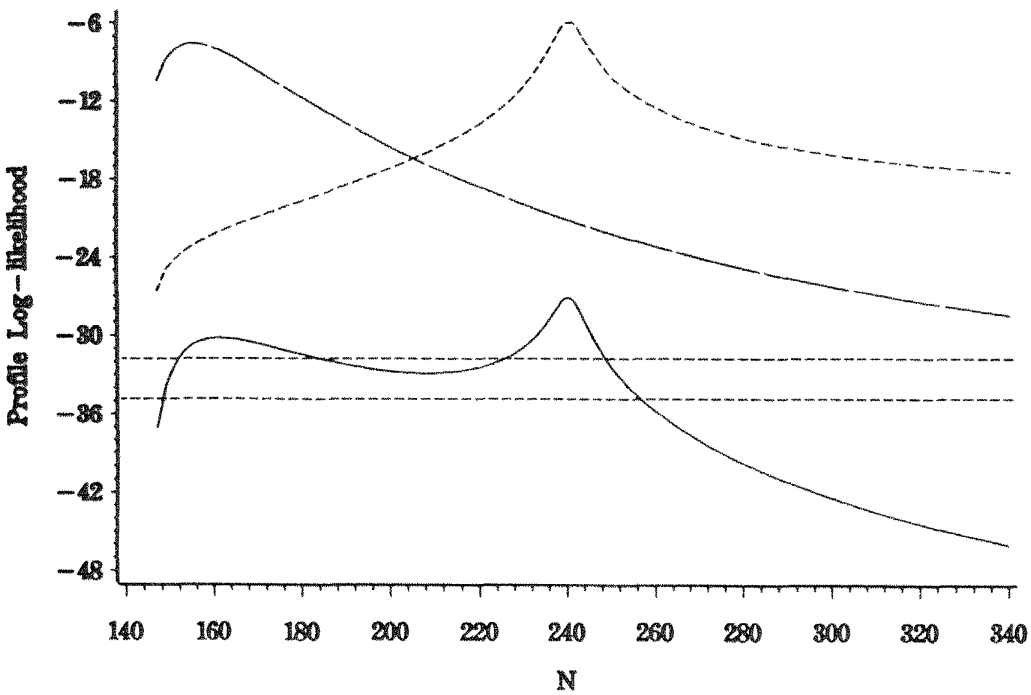


Figure 8.—Correspondence between modes of removal (large dashes) and signs-of-activities (small dashes) profile log-likelihood modes with combined (solid) profile log-likelihood modes. The two horizontal dashed lines show the 95% confidence level with the LRT statistic at equality for the combined model; the top line is associated with the "signs-of-activities" peak and the bottom with the "removal" peak.

yielding the disjoint interval $[152, 184] \cup [226, 249]$ (the upper horizontal line in Figure 8). The latter of these would actually be used as the true confidence intervals for the global maximum at $\hat{N} = 240$ (the other mode produced an estimate for N of 161) in this simulation experiment; the other information is provided to allow detailed comparison against plots of the profile log-likelihoods as was done here.

SAS Summary Programs

There are three SAS (1990) programs that are used to summarize the results of the simulation experiments from paired runs of *RemovSim* and *CiSim*. The first, *MakeRmvl*, is used to create a SAS dataset from an output file from *RemovSim*. The second, *MakeCI*, does the same for an output file from *CiSim*. Directions for use are given in the program code and are not duplicated here.

The final program, *CiSum*, is used to summarize all of the simulation results in the output file produce by *RemovSim* and *CiSim*. *CiSum* produces summary reports for each estimator with statistics on intervals sizes and catch frequency. It also provides summary statistics on the form of the log-likelihoods. These statistics relate to the boolean variables in record 1 of both output files.

Appendix IV: Spreadsheet Programs

Two spreadsheet programs were developed to analyze individual removal, or removal with signs-of-activities experiments. These programs were developed for Microsoft Excel for Windows (version 3.1) running on personal computers. With a small amount of work (basically changing file names for macro calls), they also can be run on the Apple Macintosh version of Excel (the programs were originally developed on the

Table 6.—Description of the records and fields associated with confidence interval estimation from *CiSim* for each simulated removal and signs-of-activities experiment generated by *RemovSim*

Record [Field][a]	Columns	Contents of field
h		The very first record of the file is a header record, its two fields are...
[1]	1–30	The input file name from *RemovSim* containing the estimates to which the confidence intervals apply.
[2]	31–45	The $\chi^2(1)$ critical value for the likelihood ratio confidence intervals on the removal estimator.
		The remaining records in the file are in groups of 10 for each of the simulation experiments in the input file (**h**[1]).
1 [1][a]	1	Contains either `r`, `s`, or `n`, depending on whether *RemovSim* has weighted any bimodal combined log-likelihoods by the removal, signs-of-activities, or no weighting, respectively.
[2][a]	2–5	The α level chosen for confidence interval construction.
[3][a]	6–17	The critical F value chosen.
[4][a]	18–29	The critical value for $\tilde{F}_\alpha$ as calculated from (4).
[5]	31–48	The GRG2 convergence criteria in FORTRAN (`a1`) format corresponding, in order of occurence, to each of the different confidence interval endpoints given in records **2** through **10**.
[6]	49	`True` if the combined log-likelihood is unimodal; `False` otherwise.
[7]	50	`True` if the removal peak is higher in a bimodal log-likelihood; `False` otherwise.
[8]	51–61	The value of N_0 at the bottom of the "valley" in a bimodal combined log-likelihood.
[9]	62–72	The value of the full (constants included) log-likelihood at N_0 in **1**[8].
[10]	73–76	The number of golden section search iterations used to find the value of N_0 in **1**[8].
[11]	77–80	The simulation experiment number; this number corresponds directly to the simulation number in **1**[24] of the *RemovSim* input file in **h**[1].
2 [1]	1–20	Lower endpoint for the removal estimate confidence interval.
[2]	21–40	Full removal log-likelihood value for **2**[1].
[3]	41–60	Upper endpoint for the removal estimate confidence interval.
[4]	61–80	Full removal log-likelihood value for **2**[3].
3 [1]	1–20	Lower endpoint for the signs-of-activities estimate confidence interval.
[2]	21–40	Full signs-of-activities log-likelihood value for **3**[1].
[3]	41–60	Upper endpoint for the signs-of-activities estimate confidence interval.
[4]	61–80	Full signs-of-activities log-likelihood value for **3**[3].
4 [1]	1–20	Lower confidence interval endpoint for the combined estimate with starting value $\hat{N}_R$.
[2]	21–40	Full combined log-likelihood value for **4**[1].
[3]	41–60	Lower middle confidence interval endpoint for the combined estimate with starting value $\hat{N}_R$ in the case of a bimodal log-likelihood that produces disjoint intervals; zero otherwise.
[4]	61–80	Full combined log-likelihood value for **4**[3].
5 [1]	1–20	Upper middle confidence interval endpoint for the combined estimate with starting value $\hat{N}_R$ in the case of a bimodal log-likelihood that produces disjoint intervals; zero otherwise.
[2]	21–40	Full combined log-likelihood value for **5**[1].
[3]	41–60	Upper confidence interval endpoint for the combined estimate with starting value $\hat{N}_R$.
[4]	61–80	Full combined log-likelihood value for **5**[3].
6 [1]	1–20	Lower confidence interval endpoint for the combined estimate with starting value $\hat{N}_S$.
[2]	21–40	Full combined log-likelihood value for **6**[1].

Table 6.—*Continued*

Record [Field][a]	Columns	Contents of field
[3]	41–60	Lower middle confidence interval endpoint for the combined estimate with starting value $\hat{N}_S$ in the case of a bimodal log-likelihood that produces disjoint intervals; zero otherwise.
[4]	61–80	Full combined log-likelihood value for **6**[3].
7 [1]	1–20	Upper middle confidence interval endpoint for the combined estimate with starting value $\hat{N}_S$ in the case of a bimodal log-likelihood that produces disjoint intervals; zero otherwise.
[2]	21–40	Full combined log-likelihood value for **7**[1].
[3]	41–60	Upper confidence interval endpoint for the combined estimate with starting value $\hat{N}_S$.
[4]	61–80	Full combined log-likelihood value for **7**[3].
8 [1]	1–20	Lower confidence interval endpoint for the combined estimate with starting value N.
[2]	21–40	Full combined log-likelihood value for **8**[1].
[3]	41–60	Lower middle confidence interval endpoint for the combined estimate with starting value N in the case of a bimodal log-likelihood that produces disjoint intervals; zero otherwise.
[4]	61–80	Full combined log-likelihood value for **8**[3].
9 [1]	1–20	Upper middle confidence interval endpoint for the combined estimate with starting value N in the case of a bimodal log-likelihood that produces disjoint intervals; zero otherwise.
[2]	21–40	Full combined log-likelihood value for **9**[1].
[3]	41–60	Upper confidence interval endpoint for the combined estimate with starting value N.
[4]	61–80	Full combined log-likelihood value for **9**[3].
10[1]	1–20	Lower endpoint for the weighted combined estimate confidence interval.
[2]	21–40	Full weighted combined log-likelihood value for **10**[1].
[3]	41–60	Upper endpoint for the weighted combined estimate confidence interval.
[4]	61–80	Full weighted combined log-likelihood value for **10**[3].

[a]These fields remain constant for every simulation experiment generated in the run of *RemovSim*.

Macintosh as a teaching aid and subsequently converted to the Windows environment and enhanced). These programs use the macro and nonlinear solution capabilities of Excel and produce results comparable in numerical accuracy to those of *RemovSim* and *CiSim* that both use GRG2 for numerical solutions.

Before either spreadsheet program can be used to analyze the experimental data, the macro sheet *RemvMacs* must be loaded into Excel. *RemvMacs* contains two macros that are used to calculate the removal and signs-of-activities log-likelihood functions. Each macro has variable names with detailed comments and cell notes accompanying the macro source code, so they are not described in detail here. An understanding of the macro code is not essential to using the spreadsheet programs that rely on them.

The main part of the *RemvSoa* spreadsheet program is presented in Figure 9. This page of the spreadsheet is described below using spreadsheet column letters and row numbers to designate the blocks of cells being described.

B3 This is the cell to enter a title for the experiment.

B5:F10 This area provides a place to enter the R_i and Y_i values collected as data in the experiment. Currently, the spreadsheet is dimensioned for $M = 3$ removals; however, this can be changed

Removal Estimation with Signs-of-Activities

J.H. Gove

Simulated bimodal population

Removal Attempt	Ri	Yi	Ti	LogRFact
0	0	219	0	
1	96	132	96	345.3794071
2	32	104	128	81.55795946
3	18	85	146	36.39544521

Intermediate quantities...

m =	3
Tm =	146
SumLogRFact =	463.332812
MinN =	147

Weights...

w_1 =	1
w_2 =	1
	2

Optimal Solution

Mathematical programming Model...

Maximize: lambda(N) = w_1 *lambda_1(N) + w_2*lambda_2(N)

Subject to: N-Tm > 0

Objective Function Components...

lambda_1...	p_1...		(Removal)
-21.1590658	0.29436232		
lambda_2...			(SOA)
-5.95791411			

Optimal values...

	lambda...	p...
	-27.1169799	0.29436232
N - Tm =	93.9957978	
N =	239.995798	

Figure 9.—Portion of the *RemvSoa* spreadsheet showing the main data entry and optimization sections with simulated data from Appendix III.

by adding more rows to the spreadsheet and simply changing a few formulas in the body of the sheet. The T_i are calculated automatically, as is the quantity $\log(R_i!)$ in column *F*. We recommend that you turn off the automatic calculation when entering the R_i and Y_i values so that the entire spreadsheet (which contains many calls to the macro functions) is not updated automatically each time a new entry is added. Automatic calculation can then be turned on when entry of data into these cells has been completed.

B12:C16 M, T_M, and $\sum_{i=1}^{M} \log(R_i!)$ are all quantities whose formulas may need to be changed if more removals are added to the spreadsheet in *B5:F10*. *MinN* is just $T_M + 1$ and is calculated automatically. M must be entered manually.

B18:C20 These weights are assigned to the objective function in equation set (9) to determine which log-likelihood function will be used in the maximization problem. In this spreadsheet, the w_i correspond exactly to the α_i used as weights in the body of this report (cf. Equation 7). These weights ***must*** add to either one or two. Setting w_1 to one and w_2 to zero gives the removal log-likelihood; setting the weights to the reverse (i.e, $w_1 = 0$, $w_2 = 1$) gives the signs-of-activities log-likelihood; setting $w_1 = w_2 = 1$ yields the combined log-likelihood; setting them to any other values that sum to one produces a weighted combined log-likelihood. These weights must be changed manually to the desired values.

C31:D31 These are the values of $\lambda_1(\hat{N})$ and $\hat{p}$ for the removal log-likelihood using the R_i values from the experiment. The values for these fields are calculated by the removal log-likelihood macro in *RemvMacs* during the maximization process; these are the return values from this macro and are treated as an array of cells in Excel.

C33 This cell holds the value of $\lambda_2(\hat{N})$ calculated using the R_i and Y_i and the signs-of-activities log-likelihood macro in *RemvMacs* during the maximization process.

C37 Cell *C26* presents the formula used to calculate λ in cell *C37* and is the same as (7). As stated previously, it is a combination of the removal and signs-of-activities log-likelihoods that depends on the weights.

C38 The constraint from equation set (9) assuring that $\hat{N} > T_M$ is calculated here for the optimal solution.

C39 This is the maximum likelihood estimate for the total population size (N). To have Excel solve for this value for any data entered into the spreadsheet, you must invoke Excel's "Solver." Solver is set up for the optimization in equation set (9) based on the user's data and cells *C37:C39*.

All of the calculations necessary for estimating population size using data collected from a removal or combined removal and signs-of-activities experiment are performed in the section of *RemvSoa* presented in Figure 9.

The estimation of confidence interval endpoints using the likelihood ratio method is similar to population size estimation in that it can be treated as either an optimization, or search problem. As such, confidence interval estimation may be automated as it is in *CiSim*, where it is treated as an optimization problem; however, it may also be visualized as a simple search problem—this approach lends itself neatly to spreadsheet applications. Simply, values of $\lambda(N_0)$ are calculated in a list of spreadsheet rows using some desired increment in N_0, usually from $T_M + 1$ to an upper bound. These log-likelihood values are then searched by eye to find the values of $\lambda(N_0)$ such that

$$\lambda(N_0) \geq \lambda(\hat{N}) - \frac{C_\alpha}{2}$$

in the list; where, C_α is the appropriate critical value from the χ^2 or $\tilde{F}$ distributions and $\hat{N}$ is the maximum likelihood estimate for N, both depending on which log-likelihood is being maximized. If enough values are selected in the range of N_0, then profile log-likelihoods also can be plotted on the spreadsheet.

RemvSoa provides a portion of the spreadsheet area beginning in row 44 for just such calculations. Figure 10 presents a portion of such a listing in cells *A47:E77*. These cells contain enough values of the combined log-likelihood (λ) in column E to determine the four confidence interval endpoints for the profile log-likelihood presented in Figure 8. (Normally, enough rows would be used to calculate log-likelihood values by increments of one or two to produce a complete profile log-likelihood as in Figure 8 on the spreadsheet; the disjoint subsets of values were used here to facilitate the illustration of confidence interval construction.) To determine the confidence interval endpoints for the combined estimate of N using Figure 10, we first calculate the value that the log-likelihood will take at these four points. For $M = 3$, $F(1,2) = 18.5128$ (from an F table at $\alpha = .05$), and we see from Equation 6 that

$$\lambda(N_0) \geq \lambda(\hat{N}_C) - \frac{\tilde{F}_{.05}}{2},$$

and $\tilde{F}_{.05} = 9.3116$ from (4) so that we calculate $\lambda(N_0) \geq -31.7728$. Finally, we peruse the list in Figure 10 to find all values of N_0 such that $\lambda(N_0) \geq -31.7728$. Because only whole values of N are shown in Figure 10, we settle for the closest values of N_0 satisfying this inequality. These values, shown as shaded rows in Figure 10, produce the confidence interval $[152, 183] \cup [226, 248]$. This is almost exactly the same interval obtained from *CiSim* in Appendix III; the difference is that *CiSim* calculated the intervals using an optimization method that gave a more accurate estimate in terms of significant digits (i.e., noninteger values of N_0) for two of the endpoints. We could very easily have arrived at the exact same endpoints as *CiSim* simply by adding several more spreadsheet rows around our shaded endpoints where we increment N by a little finer value, for example, 0.1 in these added rows. We could then round the 2nd and 4th endpoints to 184 and 249, respectively, arriving at the same values as *CiSim*.

As mentioned earlier, the listed values of N with the respective log-likelihood values can be used to plot the profile log-likelihood for any set of R_i and Y_i entered into the spreadsheet. There are two other sections of the *RemvSoa* spreadsheet that plot the removal and signs-of-activities profile log-likelihoods on one graph and the combined log-likelihood on another. These, however, are not shown here as they are very similar to Figure 8, which was produced by SAS. Thus, *RemvSoa* provides all of the tools needed (except the bisection search for optimal weights procedure) to analyze a combined removal and signs-of-activities experiment as outlined by Routledge (1989) and this report.

Another spreadsheet program that will not be described here is available for analyzing simple constant capture probability model removal experiments covered in White et al. (1982, p. 101–108). The spreadsheet *CCPRemvl* is laid out in a similar fashion to *RemvSoa* and includes estimation of N and graphical solution section complete with profile log-likelihood. It also includes the chi-square goodness-of-fit test and estimated standard error that can be used to calculate approximate confidence intervals based on the normal distribution. These then can be compared with the likelihood ratio confidence interval endpoints calculated from the log-likelihood values. In this latter case, the correct χ^2 value is used rather than the modified $\tilde{F}_\alpha$ in evaluating the log-likelihood endpoint values for $\lambda_1(N_0)$.

Both of these spreadsheet programs have been adapted for $M > 3$ and have been used as a teaching aid in a graduate level sampling course at the University of New Hampshire.

	A	B	C	D	E
44		**Graphical Solution...**			
45					
46	N...	lambda_1...	p...	lambda_2...	lambda...
47	147	-10.50538743	0.67281106	-26.5791207	-37.0845082
48	148	-9.511111188	0.66363636	-25.5422548	-35.053366
49	149	-8.834821423	0.65470852	-24.9290054	-33.7638269
50	150	-8.362971999	0.6460177	-24.4857535	-32.8487255
51	151	-8.03491938	0.63755459	-24.1335998	-32.1685192
52	152	-7.81343271	0.62931034	-23.8377413	-31.651174
53	153	-7.673661107	0.6212766	-23.579825	-31.2534861
54	154	-7.598062374	0.61344538	-23.3490323	-30.9470946
55	182	-12.19573724	0.45341615	-19.4437307	-31.6394679
56	183	-12.39787947	0.44923077	-19.3237178	-31.7215972
57	184	-12.59865136	0.44512195	-19.2034397	-31.802091
58	185	-12.79796529	0.44108761	-19.0828097	-31.880775
59	186	-12.9957461	0.43712575	-18.9617424	-31.9574885
60	187	-13.19192968	0.43323442	-18.8401524	-32.0320821
61	188	-13.3864617	0.42941176	-18.7179547	-32.1044164
62	224	-19.21245443	0.32589286	-12.7903976	-32.002852
63	225	-19.34407126	0.32372506	-12.5170756	-31.8611468
64	226	-19.4742794	0.3215859	-12.2290854	-31.7033648
65	227	-19.60309773	0.31947484	-11.9246015	-31.5276993
66	228	-19.73054498	0.3173913	-11.6014822	-31.3320272
67	229	-19.85663972	0.31533477	-11.257208	-31.1138477
68	230	-19.98140034	0.31330472	-10.8888165	-30.8702168
69	231	-20.10484504	0.31130064	-10.4928457	-30.5976908
70	246	-21.81035682	0.28404669	-8.5384138	-30.3487706
71	247	-21.91505269	0.28239845	-9.00938551	-30.9244382
72	248	-22.01870965	0.28076923	-9.43826574	-31.4569754
73	249	-22.121342	0.2791587	-9.82903755	-31.9503795
74	250	-22.2229638	0.27756654	-10.18615	-32.4091138
75	251	-22.32358895	0.27599244	-10.5137919	-32.8373808
76	252	-22.42323107	0.27443609	-10.8156652	-33.2388962
77	253	-22.52190362	0.2728972	-11.0949563	-33.6168599

ure 10.—Portion of the *RemvSoa* spreadsheet corresponding to Figure 9 showing a list of N_0, $\hat{p}$, and log-lihood values ($\lambda_1(N_0)$, $\lambda_2(N_0)$, and $\lambda(N_0)$) used for determining the LRT confidence interval endpoints aded rows).

☆U.S. GOVERNMENT PRINTING OFFICE: 1995 - 652-061/00937