SURVIVAL ANALYSIS IN TELEMETRY STUDIES: THE STAGGERED ENTRY DESIGN

KENNETH H. POLLOCK, Southeastern Cooperative Wildlife and Fisheries Statistics Project, Institute of Statistics, North Carolina State University, Box 8203, Raleigh, NC 27695-8203
SCOTT R. WINTERSTEIN, Department of Fisheries and Wildlife, Michigan State University, East Lansing, MI 48824
CHRISTINE M. BUNCK, Patuxent Wildlife Research Center, U.S. Fish and Wildlife Service, Laurel, MD 20708
PAUL D. CURTIS, Department of Zoology, North Carolina State University, Box 7617, Raleigh, NC 27695-7617

Abstract: The estimation of survival distributions for radio-tagged animals is important to wildlife ecologists. Allowance must be made for animals being lost (or censored) due to radio failure, radio loss, or emigration of the animal from the study area. The Kaplan-Meier procedure (Kaplan and Meier 1958), widely used in medical studies subject to censoring, can be applied to this problem. We developed a simple modification of the Kaplan-Meier procedure that allows for new animals to be added after the study has begun. We present 2 examples using telemetry data collected from northern bobwhite quail (*Colinus virginianus*) to show the simplicity and utility of the Kaplan-Meier procedure and its modifications. The log rank test used to compare 2 survival distributions can also be modified to allow for additions during the study. Simple computer programs that can be run on a personal computer are available from the authors.

J. WILDL. MANAGE. 53(1):7–15

Radio-tagged animals are used to study survival. Present techniques for analyzing data from these studies assume that each survival event (typically an animal surviving a day) is independent and has a constant probability over all animals and all periods (Trent and Rongstad 1974, Bart and Robson 1982, Heisey and Fuller 1985). We believe these assumptions are often unrealistic and restrictive. White (1983) generalized discrete approaches using the same framework as that of band return models (Brownie et al. 1985) and he developed a flexible computer program (SURVIV) for use with his approach. Heisey and Fuller (1985) generalized the Trent and Rongstad (1974) approach to allow mortality from different causes (e.g., predation, starvation) and developed a microcomputer program called MICROMORT.

Typically an animal's exact survival time (at least to within 1–2 days) is known unless that survival time is right censored (i.e., only known to be greater than some value). Pollock (1984) and Pollock et al. (1989) suggested a useful approach based on continuous survival models allowing right censoring that is widely used in medicine and engineering (Kalbfleisch and Prentice 1980, Cox and Oakes 1984) and provided examples of the Kaplan-Meier procedure. The Kaplan-Meier procedure does not require specification of a particular parametric continuous distribution; e.g., the exponential or Weibull. Related ecological papers using survival methods include Muenchow (1986), Pyke and Thompson (1986), Kurzejeski et al. (1987), and White et al. (1987).

We present a simple description of the Kaplan-Meier procedure with an example using northern bobwhite quail survival data collected by PDC. We then generalize the Kaplan-Meier procedure to allow gradual (or staggered) entry of animals into the study. The calculations are illustrated with an example from the quail data. Finally, we present the log-rank test for comparison of survival distributions (modified for staggered entry of animals) with an example. We also present a discussion of model assumptions and directions for future research.

We thank J. D. Nichols and W. L. Link for helpful comments on an earlier draft of this paper. We acknowledge G. C. White and D. M. Heisey for their helpful reviews that improved the final version.

THE KAPLAN-MEIER OR PRODUCT LIMIT PROCEDURE

The Kaplan-Meier or product limit estimator was developed by Kaplan and Meier (1958) and is discussed by Cox and Oakes (1984:48) and Kalbfleisch and Prentice (1980:13). The survival function ($S[t]$) is the probability of an arbitrary animal in a population surviving t units of time from the beginning of the study. A nonparametric estimator of the survival function can be obtained by restricting ourselves to the discrete time points when deaths occur $a_1, a_2, \ldots, a_g$. We define $r_1, \ldots, r_g$ to be the numbers of an-

Table 1. Kaplan-Meier survival estimates for northern bobwhite quail radiotagged in North Carolina, spring 1985.

Week (t)	Dates	No. at risk (r_j)	No. deaths (d_j)	No. censored	Survival ($\hat{S}(t)$)	95% CI
1	3 Mar–9 Mar	18	0	0	1.0000	1.0000–1.0000
2	10 Mar–16 Mar	18	0	0	1.0000	1.0000–1.0000
3	17 Mar–23 Mar	18	2	0	0.8889	0.7520–1.0258
4	24 Mar–30 Mar	16	0	0	0.8889	0.7437–1.0341
5	31 Mar–6 Apr	16	0	0	0.8889	0.7437–1.0341
6	7 Apr–13 Apr	16	1	0	0.8333	0.6667–1.0000
7	14 Apr–20 Apr	15	0	0	0.8333	0.6612–1.0055
8	21 Apr–27 Apr	15	1	1	0.7778	0.5922–0.9633
9	28 Apr–4 May	13	1	2	0.7179	0.5107–0.9252
10	5 May–11 May	10	1	1	0.6462	0.4079–0.8844
11	12 May–18 May	8	0	0	0.6462	0.3798–0.9125
12	19 May–25 May	8	0	1	0.6462	0.3798–0.9125
13	26 May–1 Jun	7	0	0	0.6462	0.3614–0.9309

imals at risk at these points and $d_1, d_2, \ldots, d_g$ to be the number of deaths at the same points. The probability of surviving from zero to a_1 is estimated by $\hat{S}(a_1) = 1 - d_1/r_1$, because d_1/r_1 is the estimated proportion dying in that interval. The probability of surviving from a_1 to a_2 is similarly given by $1 - d_2/r_2$ and $\hat{S}(a_2)$ is then given by the product: $(1 - d_1/r_1)(1 - d_2/r_2)$. Therefore the estimated survivor function for any arbitrary time t is given by:

$$\hat{S}(t) = \prod_{j \mid a_j < t,} (1 - d_j/r_j) \qquad (1)$$

which is the mathematical way of stating we are considering the product of all j terms for which $a_j <$ the time t.

Consider r_j, which is the number at risk at time a_j. In this situation we would start with a fixed sample (n). The number at risk at a particular death time a_j will then be n − the number of deaths before a_j − the number of animals censored before time a_j.

As an example of the use of this model, we present results from a spring season of a radiotagging study on northern bobwhite conducted by PDC at Fort Bragg, North Carolina. In this section we consider the data collected in spring 1985. The pertinent data on each of the 18 radiotagged quail is included in Table 1. Six birds died and 5 birds disappeared (were censored) during the study, leaving 7 birds that survived for the 13 weeks of the study.

We estimated $S(t)$ (Table 1, Fig. 1) in the following manner. The computations involve the 5 weeks where deaths were recorded; therefore $a_1 = 3$, $a_2 = 6$, $a_3 = 8$, $a_4 = 9$, and $a_5 = 10$. We estimate $\hat{S}(a_1)$ as:

$$\hat{S}(a_1) = \hat{S}(3) = 1 - d_1/r_1 = 1 - 2/18 = 0.8889,$$

because there are 2 deaths at time 3 and there are 18 animals still at risk (r_1). The next death time a_2 is at 6 weeks ($a_2 = 6$) and at that time there is 1 death ($d_2 = 1$) and 16 animals at risk ($r_2 = 16$). There are 16 at risk because 2 were lost to death at time 1. Therefore $\hat{S}(a_2)$ is given by:

$$\hat{S}(a_2) = \hat{S}(6) = (1 - d_1/r_1)(1 - d_2/r_2) = (1 - 2/18)(1 - 1/16) = 0.8333.$$

Similarly $\hat{S}(a_3)$ is given by:

$$\hat{S}(a_3) = \hat{S}(8) = (1 - d_1/r_1)\cdot(1 - d_2/r_2)(1 - d_3/r_3) = (1 - 2/18)(1 - 1/16)\cdot(1 - 1/15) = 0.7778,$$

$$\hat{S}(a_4) = \hat{S}(9) = (1 - d_1/r_1)(1 - d_2/r_2)\cdot(1 - d_3/r_3)(1 - d_4/r_4) = (1 - 2/18)(1 - 1/16)(1 - 1/15)\cdot(1 - 1/13) = 0.7179,$$

and

$$\hat{S}(a_5) = \hat{S}(10) = (1 - d_1/r_1)(1 - d_2/r_2)\cdot(1 - d_3/r_3)(1 - d_4/r_4)\cdot(1 - d_5/r_5) = (1 - 2/18)(1 - 1/16)(1 - 1/15)\cdot(1 - 1/13)(1 - 1/10) = 0.6462.$$

The censored observation at $t = 8$ is still considered at risk until the instant after that time so that $r_3 = 15$, not 14, but then $r_4 = 13$. The estimate of the survivor function ($\hat{S}[t]$) is presented for each week in Table 1 but it only changes at the death times. Thus, $\hat{S}(t)$ stays at 1.00 until $t = 3$ where it becomes 0.8889 and

stays there until $t = 6$ (the next death time) and so on.

Cox and Oakes (1984:51) also discussed how to estimate the variance (var) of the estimate at an arbitrary time point using Greenwood's (Greenwood 1926) formula:

$$\text{var}[\hat{S}(t)] = [\hat{S}(t)]^2 \sum_{1}^{j|a_j<t} \frac{d_j}{r_j(r_j - d_j)}, \qquad (2)$$

where the summation is for all death times $a_j < t$. Cox and Oakes (1984) also discuss an alternate estimate that is better in the tails of the distribution:

$$\text{var}(\hat{S}[t]) = \frac{[\hat{S}(t)][1 - \hat{S}(t)]}{r(t)}. \qquad (3)$$

Approximate confidence intervals can be obtained based on equation (2) or (3). For example, a 95% confidence interval at $t = t_0$ would be:

$$\hat{S}(t_0) \pm 1.96\,[\text{var}\,\hat{S}(t_0)]^{\frac{1}{2}} \qquad (4)$$

because of the asymptotic (large sample) normality of the estimates $\hat{S}(t)$. In Table 1 approximate 95% confidence intervals are given at all points using the simpler second variance equation (3). The confidence intervals get wider and wider as time increases (Fig. 1).

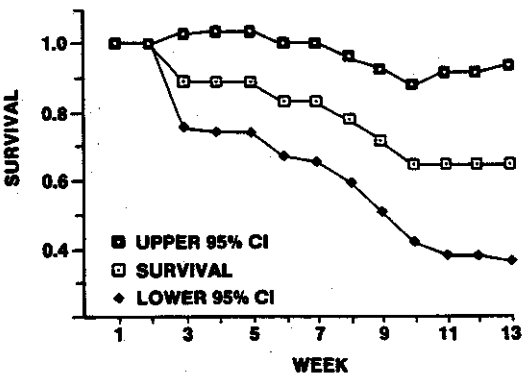

Fig. 1. The Kaplan-Meier survival function for northern bobwhite quail radiotagged in North Carolina, spring 1985.

EXTENSION OF THE KAPLAN-MEIER PROCEDURE TO STAGGERED ENTRY OF ANIMALS

We extend the concept of the Kaplan-Meier estimates to allow animals to enter at different times and for the time variable to be measured from the point where the first group of animals is tagged. Previously we presented an example of the Kaplan-Meier estimator and showed it is based on equation (1) for the survival function where r_j is the number at risk and d_j is the number of deaths. Typically we assume r_j is decreasing due to deaths and censoring but there is no reason it has to be. New animals will only be considered in those product terms where they are at risk. The formula for the variance of $\hat{S}(t)$ also allows for new animals to enter during the study. Any newly tagged animals are assumed to have the same survival function as the previously tagged animals.

Using telemetry data collected from northern bobwhite in the winter of 1985–86, we illustrate the extension of the Kaplan-Meier estimator to staggered entry of new animals (Table 2). In week 1 there were 20 animals radiotagged ($r_1 = 20$), no deaths, no censors, and 1 animal was added so that the number of animals radiotagged in week 2 is 21 and the survival estimate stays at 1.0. In week 2 the only change is that another animal is added so that at week 3 the number at risk is $r_3 = 22$. In week 3 the survival estimate is: $\hat{S}(3) = 1 - d_3/r_3 = 1 - 2/22 = 0.9091$. In week 4 the number at risk is 19 (2 deaths and 1 censor in week 3 gives $22 - 2 - 1 = 19$) and there were 5 deaths so:

$$\begin{aligned}\hat{S}(4) &= (1 - d_3/r_3)(1 - d_4/r_4)\\ &= (1 - 2/22)\,(1 - 5/19)\\ &= 0.6699.\end{aligned}$$

The survival estimates for later times can be obtained in a similar manner. The approximate 95% confidence limits are also given based on equations (3) and (4). Here the confidence limits do not necessarily get wider with time because new animals may be added (Fig. 2). At $t = 13$ there is a marked decrease in the confidence interval width because the number at risk increases from 10 to 16 and at $t = 14$ it increases to 22 due to the large number of new animals added at those times.

LOG-RANK TEST EXTENSION TO STAGGERED ENTRY OF ANIMALS

Often it is important to compare 2 estimated survival functions to see if they could have come from the same underlying true survival curve. For example, in Table 3 we present some bobwhite survival estimates for fall 1985 and fall 1986. We would like to know if the survival patterns are the same for the 2 years. Graphical comparison would be possible by plotting survival functions on the same graph; however, a

Table 2. Kaplan-Meier survival estimates for northern bobwhite quail radiotagged in North Carolina, winter 1985–86, modified to allow for the staggered entry of new animals into the study.

Week (t)	Dates	No. at risk (r_j)	No. deaths (d_j)	No. censored	No. new added	Survival ($\hat{S}(t)$)	95% CI
1	17 Nov–23 Nov	20	0	0	1	1.0000	1.0000–1.0000
2	24 Nov–30 Nov	21	0	0	1	1.0000	1.0000–1.0000
3	1 Dec–7 Dec	22	2	1	0	0.9091	0.7946–1.0236
4	8 Dec–14 Dec	19	5	0	0	0.6699	0.4968–0.8429
5	15 Dec–21 Dec	14	3	0	0	0.5263	0.3366–0.7161
6	22 Dec–28 Dec	11	0	0	0	0.5263	0.3122–0.7404
7	29 Dec–4 Jan	11	0	0	0	0.5263	0.3122–0.7404
8	5 Jan–11 Jan	11	2	0	0	0.4306	0.2386–0.6226
9	12 Jan–18 Jan	9	1	0	0	0.3828	0.1863–0.5792
10	19 Jan–25 Jan	8	0	1	0	0.3828	0.1744–0.5912
11	26 Jan–1 Feb	7	0	0	3	0.3828	0.1600–0.6056
12	2 Feb–8 Feb	10	0	0	6	0.3828	0.1964–0.5692
13	9 Feb–15 Feb	16	4	0	10	0.2871	0.1683–0.4059
14	16 Feb–22 Feb	22	4	0	5	0.2349	0.1490–0.3207
15	23 Feb–1 Mar	23	4	1	6	0.1940	0.1228–0.2652
16	2 Mar–8 Mar	24	4	0	0	0.1617	0.1025–0.2209
17	9 Mar–15 Mar	20	2	0	0	0.1455	0.0866–0.2045

formal hypothesis testing procedure is also needed.

There are many possible tests available (Lee 1980:122) but we concentrate attention on the log-rank test (Savage 1956, Kalbfleisch and Prentice 1980:17, Cox and Oakes 1984:104) because of the test's simplicity and easy generalization to when animals have staggered entry into the study.

To compare 2 survivor functions we generalize the formula we used for defining the Kaplan-Meier estimates. Let $a_1, a_2, \ldots, a_k$ denote the death times for the sample formed by combining the 2 samples. Suppose there are d_j deaths and r_j animals at risk at a_j, with d_{0j} and d_{1j} being from samples 1 and 2, respectively. Similarly there are r_{0j} and r_{1j} animals at risk from the 2 samples.

For each of the k points the data can be represented as a 2 × 2 contingency table. For the jth contingency table we have formula for the expected value (E) and variance of d_{1j} given by:

$$\mathrm{E}(d_{1j}) = d_j r_{1j}/r_j$$

and

$$\mathrm{var}_1(d_{1j}) = d_j r_{0j} r_{1j}(r_j - d_j)/r^2_j(r_j - 1).$$

An approximate Chi-square test statistic with 1 degree of freedom can be obtained by combining the results from all the contingency tables (assuming conditional independence and asymptotic normality of the d's) in the following way:

$$\chi^2 = \frac{\left[\sum_{j=1}^{k} d_{1j} - \sum_{j=1}^{k} \mathrm{E}(d_{1j})\right]^2}{\sum_{j=1}^{k} \mathrm{var}(d_{1j})}.$$

Cox and Oakes (1984:105) consider 2 modifications to the var of d_{1j} that give rise to χ^2_2 and χ^2_3, respectively. The first modification is:

$$\mathrm{var}_2(d_{1j}) = \frac{r_{0j} r_{1j} d_j}{r_j^2},$$

which is a slightly more conservative test. The second modification is

$$\sum_{j=1}^{k} \mathrm{var}(d_{1j}) = \left[\frac{1}{\sum_{j=1}^{k} (d_j r_{0j}/r_j)} + \frac{1}{\sum_{j=1}^{k} (d_j r_{1j}/r_j)}\right]^{-1},$$

which is an even more conservative test. This test in any of its 3 versions easily generalizes to the case of staggered entry because r_{0j} and r_{1j}, the number at risk at each time point, can be redefined to include newly tagged animals.

In Table 4 we present the calculations of the log-rank test for data given in Table 3 that compares bobwhite survival distributions for fall 1985 and fall 1986. We present the number at risk

and the number of deaths for every week although deaths do not occur every week. In those weeks where there are no deaths there are no contributions to the test statistics.

We now present the calculations for the 3 approximate Chi-square test statistics with 1 degree of freedom:

$$\chi^2_1 = \frac{(6 - 3.317)^2}{1.681} = 4.28,$$

$$\chi^2_2 = \frac{(6 - 3.317)^2}{1.729} = 4.16,$$

and

$$\chi^2_3 = \frac{(6 - 3.317)^2}{\left(\frac{1}{3.683} + \frac{1}{3.317}\right)^{-1}} = 4.13,$$

where

$$\sum_{j=1}^{k} d_{1j} = 6,$$

$$\sum_{j=1}^{k} E(d_{1j}) = 3.317,$$

and hence

$$\sum_{j=1}^{k} E(d_{0j}) = 7 - E(d_{1j}) = \sum_{j=1}^{k} d_j r_{0j}/r_j = 3.683,$$

$$\sum_{j=1}^{k} d_j r_{0j} r_{1j}/r_j^2 = 1.729,$$

and

$$\sum_{j=1}^{k} d_j r_{0j} r_{1j}(r_j - d_j)/r_j^2(r_j - 1) = 1.681.$$

All the Chi-square tests are similar. The approximate P value is 0.04 that indicates there is a significant difference between the 2-year survival curves at the 5% level.

Biologists might compare 2 survival curves at a particular time. In this case, a simple approximately normal test statistic can be based on the equation:

$$Z = \frac{\hat{S}_1(t_1) - \hat{S}_2(t_1)}{\sqrt{\widehat{\text{var}}\,\hat{S}_1(t_1) + \widehat{\text{var}}\,\hat{S}_2(t_1)}},$$

where $\hat{S}_1(t_1)$ is the estimate of the first survival curve at time t_1 based on equation (1) and $\hat{S}_1(t_1)$ is defined similarly. The variances are calculated by using equation (3). In some cases transformation to log $S(t_1)$ may improve the normality of the test statistic.

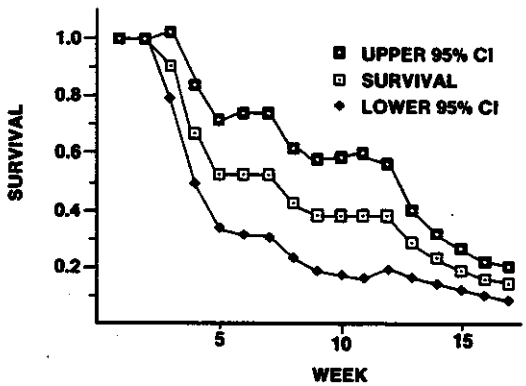

Fig. 2. The Kaplan-Meier survival function, modified for staggered entry of animals, for northern bobwhite quail radiotagged in North Carolina, winter 1985–86.

ASSUMPTIONS

We assume that animals of a particular sex and age class have been sampled randomly. For example, in a study on male winter survival of a species, if lighter adult males tend to be captured and these have lower survival rates, the survival estimates will be negatively biased. This assumption is also required of survival estimates obtained from capture–recapture and band-return studies (Jolly 1965, Seber 1965, Pollock 1981, Brownie et al. 1985).

We make the assumption that survival times are independent for the different animals. This assumption is also required of capture–recapture and band-return models as before. For Canada geese (*Branta canadensis*) that form tight family groups this assumption would probably fail. Additionally, the death of a female mammal still nursing her young would not be independent of the fate of those young. Violation of this assumption will not cause bias but it will make our estimates appear to have smaller variances than they actually do.

Another assumption that is common to any method involving marked animals is that capturing the animal or having it carry a radio tag does not influence its future survival. Clearly, failure of this assumption will negatively bias the survival estimates. As radio tags are becoming smaller this is less of a problem. Short-term effects could be eliminated by having a condi-

Table 3. Comparison of survival distributions of radiotagged northern bobwhite quail in North Carolina, fall 1985 and 1986.

	Fall 1985					Fall 1986				
Week	No. at risk	No. deaths	No. censored	No. added	$\hat{S}$ (t)	No. at risk	No. deaths	No. censored	No. added	$\hat{S}$ (t)
1	7	1	0	0	0.8571	7	0	1	0	1.0000
2	6	0	0	2	0.8571	6	0	0	5	1.0000
3	8	0	0	5	0.8571	11	1	0	0	0.9091
4	13	0	1	6	0.8571	10	0	0	6	0.9091
5	18	0	0	0	0.8571	16	1	0	0	0.8523
6	18	0	0	0	0.8571	15	0	0	0	0.8523
7	18	0	0	0	0.8571	15	1	0	0	0.7955
8	18	0	0	0	0.8571	14	0	0	0	0.7955
9	18	0	0	1	0.8571	14	3	0	0	0.6250

tioning period (e.g., 1 week) after tagging when an animal's survival time is not considered until it has survived that period.

The assumption that the censoring mechanism is random (i.e., is not related to an animal's fate) is important. Possible violations could result from a predator killing an animal and at the same time destroying the radio or an animal emigrating because it is stronger than its companions. Medical studies may suffer a similar emigration problem; patients doing well (or badly) may decide to leave the study. A review of the literature on survival analysis shows that very little has been done regarding alternative, so called, informative censoring models. We believe this is mainly due to the difficulty of the problem rather than lack of research. However, bounds can be generated for the survival curve by allowing censoring to take 2 very extreme forms. A lower bound can be obtained by assuming that every censored observation was really a death and an upper bound by assuming that every censored observation was not a death and that the animal survived to the end of the study.

Sometimes it may be reasonable to assume that emigration or radio failure is zero. Then the likelihood of the censoring time could provide very useful information. For example, in a study of winter survival of waterfowl with reliable radios, the censoring times would primarily reflect emigration. Estimation of this emigration time distribution could be informative to the biologist, especially if it could be related to covariates such as those reflecting weather severity.

The definition of a time origin is crucial. In medical studies the natural time origin is the time treatment begins. In radio telemetry there is no natural time origin. In studies where all the animals are captured at or near the same time the obvious time origin might be the date when the last animal was captured. Survival from the origin could be seriously influenced by seasonal effects, with survival for 1 week from a summer time origin quite different than survival for 1 week from a winter time origin.

Animals may need to be introduced into the study over a long period of time. This could result from difficulties in capturing animals at 1 time or from the introduction of additional animals into the study to replace animals that have died. We have shown that the Kaplan-Meier estimator of the survivor function and the log-rank test for comparing survival curves can be generalized to allow for staggered entry. In this case the time origin will be when the first group of animals is tagged.

A special assumption of the staggered entry design is that newly tagged animals have the same survival function as previously tagged animals. If there were enough animals in both groups contingency table tests of this assumption could be made. In practice, however, the animals will often be added in very small groups thereby prohibiting a quantitative assessment of this assumption.

DISCUSSION

The Kaplan-Meier procedure is simple, flexible, and allows staggered entry of newly tagged animals. Many biologists will find the simple log-rank test for comparing survival distributions valuable. It is also easily adapted to the staggered entry case. However, the log-rank test is not powerful when the survival curves have rad-

Table 4. Log-rank test calculations comparing survival distributions of radiotagged northern bobwhite quail in North Carolina, fall 1985 and 1986, modified to allow for the staggered entry of new animals into the study.

Week	Fall 1985		Fall 1986		Total				
	No. at risk (r_{0j})	No. deaths (d_{0j})	No. at risk (r_{1j})	No. deaths (d_{1j})	No. at risk (r_j)	No. deaths (d_j)	$E(d_{1j})^a$	var $(d_{1j})^b$	var $(d_{1j})^c$
1	7	1	7	0	14	1	0.500	0.250	0.250
2	6	0	6	0	12	0	0	0	0
3	8	0	11	1	19	1	0.579	0.244	0.244
4	13	0	10	0	23	0	0	0	0
5	18	0	16	1	34	1	0.471	0.249	0.249
6	18	0	15	0	33	0	0	0	0
7	18	0	15	1	33	1	0.455	0.248	0.248
8	18	0	14	0	32	0	0	0	0
9	18	0	14	3	32	3	1.313	0.738	0.691
Total		1		6		7	3.317	1.729	1.681

[a] $E(d_{1j}) = d_j r_{1j}/r_j$.
[b] Var $(d_{1j}) = d_j r_{1j} r_{0j}/r_j^2$.
[c] Var $(d_{1j}) = d_j r_{1j} r_{0j}(r_j - d_j)/r_j^2(r_j^{-1})$.

ically different shapes (Cox and Oakes 1984: 107).

We believe testing of ecological hypotheses regarding the influence of individual animal covariates on survival using the proportional hazards model (Cox 1972) is important. In Pollock et al. (1989), this model was illustrated by showing how winter survival of female black ducks (*Anas rubripes*) is related to their condition index at the start of the winter. The Cox proportional hazards model is described clearly by Cox and Oakes (1984:91). This model can also generalize to the case of staggered entry of animals.

The efficiency of the Kaplan-Meier estimator when there is staggered entry of animals should be studied thoroughly but some preliminary statements about sample size can be made based on our analyses. Table 2 and Figure 2 illustrate that precision is poor unless the number of animals tagged at a particular time is >20. To get good precision, however, a minimum of 40–50 animals would need to be tagged at all times. Usually we would recommend periodic insertion of animals to keep the number tagged approximately constant. If there is a period of interest when mortality is likely to be high the biologist should be prepared to introduce a large number of newly radio-tagged animals at that time.

We have emphasized the Kaplan-Meier product limit estimator because of its simplicity and generality. However, it is not uniquely the best for all circumstances. For example, if the survival curves follow a simple parametric form such as the exponential or Weibull then use of Kaplan-Meier violates the principle of parsimony (keeping the number of parameters as small as possible).

Miller (1983) compared maximum likelihood estimators and the Kaplan-Meier procedure when the underlying distribution was exponential and there was right censoring. He stated that this comparison is biased against the Kaplan-Meier estimator, but its efficiency can be low. This is especially troublesome for long times and Miller (1983) stated "Parametric modeling should be considered as a means of increasing the precision in the estimation of small tail probabilities." He also mentioned that this question has been studied little which is surprising considering the importance of survival analysis in engineering and medicine.

This Heisey and Fuller (1985) approach can be viewed as a piecewise exponential model (D. M. Heisey, Minn. Dep. Nat. Resour., pers. commun.; Whittemore 1985). D. M. Heisey suggested that if one lets the intervals in the Heisey and Fuller (1985) approach be individual days, that one obtains exactly the staggered-entry Kaplan-Meier method.

Lagakos (1979) discussed informative censoring (i.e., where the censoring is related to the fate of the animal) and the lack of research on the subject. One practical approach discussed is to calculate extreme bounds for the estimated survival curve by considering each censored observation to be a death or a survivor until the end of the study. If there is a lot of censoring

early in the study, these bounds can be so wide as to be impractical.

Finally, we discuss the situation where cause of death can be classified into several categories. For example, the biologist may want to separate hunting deaths from nonhunting deaths. Marginal survival curves can be obtained by treating deaths from any other cause as censored observations. For example, if one were considering the survival curve just related to hunting, then all animals that died of nonhunting causes would be viewed as censored observations. Unfortunately, this approach does not consider that different causes of death may not be independent. Competing risk models when there are several possible causes of deaths have been studied in depth but are not useful for estimating the dependency. These models are what statisticians refer to as "nonidentifiable." Therefore, biologists are forced to use marginal or crude survival curves. They should be aware that these crude curves have limitations and do not consider the possibility of dependency between sources of mortality that is a very important question. Kalbfleish and Prentice (1980:163), Cox and Oakes (1984:142), and Heisey and Fuller (1985) provide information on competing risk models and their problems.

COMPUTER PROGRAMS

Two computer programs written specifically for the analysis of survival data when all animals enter at the same time are PHGLM (Harrell 1983) and LIFETEST (SAS 1985). Other programs include SURVREG (Preston and Clarkson 1983) and LIFEREG (SAS 1985) for use with parametric models when all animals enter at the same time.

A simple computer program that calculates the Kaplan-Meier estimator and the log-rank test when there is potentially staggered entry of animals is available free from the authors. It will run on IBM compatible personal computers in conjunction with the Lotus Spread Sheet software (Kapor and Sachs 1983). The program is also available through SESAME, a bulletin board, maintained by the Southeastern Cooperative Wildlife and Fisheries Statistics Project at North Carolina State University.

LITERATURE CITED

BART, J., AND D. S. ROBSON. 1982. Estimating survivorship when the subjects are visited periodically. Ecology 63:1078–1090.

BROWNIE, C., D. R. ANDERSON, K. P. BURNHAM, AND D. S. ROBSON. 1985. Statistical inference from band recovery data: a handbook. U.S. Fish and Wildl. Serv. Resour. Publ. 131. Washington, D.C. 305pp.

COX, D. R. 1972. Regression models and life tables (with discussion). R. Stat. Soc. Ser. B 34:187–220.

———, AND D. OAKES. 1984. Analysis of survival data. Chapman and Hall, New York, N.Y. 201pp.

GREENWOOD, M. 1926. The errors of sampling of the survivorship tables. Appendix 1 *in* Rep. on Public Health and Stat. Subjects 33. London, U.K.

HARRELL, F. E. 1983. The PHGLM procedure. SAS User's Group Int. Suppl. Library User's Guide. SAS Inst. Inc., Cary, N.C.

HEISEY, D. M., AND T. K. FULLER. 1985. Evaluation of survival and cause specific mortality rates using telemetry data. J. Wildl. Manage. 49:668–674.

JOLLY, G. M. 1965. Explicit estimates from capture–recapture data with both death and immigration—stochastic models. Biometrika 52: 225–247.

KALBFLEISCH, J. D., AND L. PRENTICE. 1980. The statistical analysis of failure data. John Wiley and Sons, New York, N.Y. 321pp.

KAPLAN, E. L., AND P. MEIER. 1958. Nonparametric estimation from incomplete observations. J. Am. Stat. Assoc. 53:457–481.

KAPOR, M., AND J. SACHS. 1983. Lotus 1-2-3 User's Manual. Lotus Dev. Corp., Cambridge, Mass. 362pp.

KURZEJESKI, E. W., L. D. VANGILDER, AND J. B. LEWIS. 1987. Survival of wild turkey hens in North Missouri. J. Wildl. Manage. 511:188–193.

LAGAKOS, S. W. 1979. General right censoring and its impact on the analysis of survival data. Biometrics 35:139–156.

LEE, E. T. 1980. Statistical methods for survival data analysis. Lifetime Learning Publ., Belmont, Calif. 557pp.

MILLER, R. G. 1983. What price Kaplan-Meier? Biometrics 39:1077–1081.

MUENCHOW, G. 1986. Ecological use of failure time analysis. Ecology 67:246–250.

POLLOCK, K. H. 1981. Capture–recapture models allowing for age-dependent survival and capture rates. Biometrics 37:521–529.

———. 1984. Estimation of survival distributions in ecology. Int. Biometrics Conf. 12:187–195.

———, S. R. WINTERSTEIN, AND M. J. CONROY. 1989. Estimation and analysis of survival distributions for radio-tagged animals. Biometrics. In Press.

PRESTON, D. L., AND D. B. CLARKSON. 1983. A user's guide to SURVREG: survival analysis with regression. Am. Stat. 37:174.

PYKE, D. A., AND J. N. THOMPSON. 1986. Statistical analysis of survival and removal rate experiments. Ecology 67:240–245.

SAS. 1985. User's Guide: Statistics Version 5 ed. SAS Inst. Inc., Cary, N.C.

SAVAGE, I. R. 1956. Contributions to the theory of order statistics—the two sample case. Ann. Math. Stat. 27:590–615.

SEBER, G. A. F. 1965. A note on the multiple-recapture census. Biometrika 52:249–259.

TRENT, T. T., AND O. J. RONGSTAD. 1974. Home range and survival of cottontail rabbits in southwestern Wisconsin. J. Wildl. Manage. 38:459–472.

WHITE, G. C. 1983. Numerical estimation of survival rates from band-recovery and biotelemetry data. J. Wildl. Manage. 47:716–728.

———, R. A. GARROTT, R. M. BARTMANN, L. H. CARPENTER, AND A. W. ALLDREDGE. 1987. Survival of mule deer in northwest Colorado. J. Wildl. Manage. 51:852–859.

WHITTEMORE, A. A. 1985. Analyzing cohort mortality data. Am. Stat. 39:437–441.

Received 27 September 1987.
Accepted 27 July 1987.