REVISED

ESTIMATION AND ANALYSIS OF SURVIVAL DISTRIBUTIONS FOR RADIO-TAGGED ANIMALS

Kenneth H. Pollock[1], Scott R. Winterstein[1], and Michael J. Conroy[2]

[1]Department of Statistics, North Carolina State University
Box 8203, Raleigh, North Carolina 27695-8203, U.S.A.

[2]Patuxent Wildlife Research Center, U.S. Fish and Wildlife Service
Laurel, Maryland 20708, U.S.A.

SUMMARY

We present results on the estimation of survival distributions for an important problem in animal ecology. The problem involves estimation of survival distributions using radio-tagged animals. It requires allowance for censored observations due to radio failure, emigration from the study area and animals surviving past the end of the study period. We show that techniques already used in medical and engineering studies may be applied to this problem. Emphasis is placed on the model assumptions and the need for further research. An example to illustrate the strengths and weaknesses of this approach is presented.

Key words: Survival analysis; Right censoring; Left truncation; Radio telemetry; Kaplan-Meier Estimates; Exponential distribution; Cox regression model.

1. Introduction

Radio-telemetry is becoming an increasingly popular methodology for studying wild animal populations. An animal, captured by trap, dart gun or some other method, is fitted with a small radio transmitter and released. From release, the animal's unique radio signal can be monitored until the animal dies or is censored (see below).

The most common application of radio-telemetry technology has been to the study of animal movements in relation to daily activity patterns, seasonal changes, habitat types, and interaction with other animals. Time series approaches will become very important to the thorough analyses of these data (See Dunn and Gipson, 1977; Pantula and Pollock, 1984).

Biologists have also begun to use radio-tagged animals to study survival. Present techniques for analyzing the 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 (See Trent and Rongstad, 1974; Bart and Robson, 1982). These assumptions are often believed to be unrealistic and restrictive. White (1983) has generalized their discrete approaches in the framework of band return models (Brownie et al. (1978)).

Typically an animal's exact survival time (at least to within one or two days) is known unless that survival time is right censored. We suggest an approach based on the continuous survival models allowing right censoring which are widely used in medical and engineering applications (Kalbfleisch and Prentice, 1980; Cox and Oakes, 1984). Emphasis is placed on the important assumptions of these models in the radio-telemetry framework. A numerical example is also examined. A shorter version of our approach appeared in Pollock (1984).

2. Model Structure

We assume that a random sample of n animals has been radio-tagged. Further all animals are monitored regularly (usually daily) so that for practical purposes exact times of death are known. Also we assume that there is a fixed study area to cover and if an animal with a functional radio is present it is found (with probability one).

In terms of modelling we consider two sets of conceptual random variables:

(i) $T_1, T_2, \ldots, T_n$ form a set of survival times from tagging to death which would be observable if there were no censoring. We assume these constitute a random sample from some probability distribution with density $f(t;\underline{\theta})$ and survivor function $F(t;\underline{\theta}) = P\ (T>t)$.

(ii) $C_1, C_2, \ldots, C_n$ form a set of censoring times which would be observable if there were no deaths. We assume these constitute a random sample from some probability distribution with density $g(c;\underline{\gamma})$ and survivor function $G(c;\underline{\gamma}) = P\ (C>c)$.

The censoring could be due to any one of three possible causes:

(i) An animal has a transmitter which fails before the animal dies;

(ii) An animal emigrates out of the study area; and

(iii) An animal survives past the end of the study period. In some studies it may be possible to ignore emigration for biological reasons.

Assuming that one has a random censoring mechanism, which implies that the two sets of conceptual random variables are independent, the likelihood for the survival times is given by the following equation (Kalbfleisch and Prentice, 1980, p. 40):

$$L(\underline{\theta}) = \prod_{i=1}^{n} (f(t_i;\underline{\theta}))^{\delta_i} (F(c_i;\underline{\theta}))^{1-\delta_i}$$

where δ_i is the censoring indicator. If $\delta_i = 1$, the observation is uncensored $(0<t_i<c_i)$ while if $\delta_i = 0$ the observation is right censored $(t_i>c_i)$. Note that it is also possible to write down a similar likelihood function for the censoring times $(L(\underline{\gamma}))$. This will be explored further in the next section.

If one assumes a parametric form for $f(t;\underline{\theta})$ (e.g., Exponential distribution, Weibull distribution, Gamma distribution), standard maximum likelihood inference could be carried out. Alternately, one could obtain the "distribution free" product limit estimate of $F(t;\underline{\theta})$ originally derived by Kaplan and Meier in 1958 (Kalbfleisch and Prentice, 1980, p. 13). If important covariates, such as weight at tagging, were available they could be incorporated into the proportional hazards model (Cox, 1972; Kalbfleisch and Prentice, 1980, p. 32).

3. Model Assumptions

Here we briefly discuss the assumptions of the model as they apply to radio-telemetry data. As this is a new approach, further research on the validity of these assumptions in practice and on model robustness is required.

We have assumed that a random sample of animals of a particular age and sex class has been obtained. Take, for example, a study on winter survival of mallards. If lighter adult males tend to be captured and these have lower survival rates, a negative bias to the survival estimates will result. Of course this assumption is also crucial to survival estimates obtained from capture-recapture and band return studies (Jolly, 1965; Seber, 1965; Pollock, 1981; Brownie et al., 1978).

This model requires the assumption that survival times are independent for the different animals. Again, this assumption is also required of capture-recapture and band return models. Geese which form tight family groups would be an example where this assumption could fail. As another example, the death of a female mammal, such as a Black Bear, still nursing her young would not be independent of the fate of those young.

In most studies time of death will be known to the nearest day. Therefore use of continuous time survival distributions seems justified.

The assumption that the censoring mechanism is random is extremely important and requires more attention than can be given here. Possible violations could result from a predator killing an animal and also destroying the radio or an animal emigrating because it is more (or less) healthy than its companions. We wish to point out that medical studies often suffer a similar "emigration" problem; patients doing poorly (or well) may decide to leave the study.

One of the most important considerations in application of survival analysis to radio-telemetry data is the definition of a time origin. In medical studies the natural time origin is time of treatment. In radio-telemetry there is no such 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. It should be kept in mind that survival from the time origin could be vastly affected by seasonal effects so that for example survival for one month from a summer time origin could be quite different from survival for one month from a winter time origin. Seasonal effects are probably much less important in medical studies.

In some studies animals may be introduced into the study gradually over a long period of time. This could be due to practical problems of capturing animals all at one time or because the biologist deliberately wants to introduce more animals into the study to increase precision after a lot of his animals have died. This is a situation where some animals will be subject to left truncation because they will only contribute to the likelihood if the truncation point (the point where they enter the study) is exceeded. Cox and Oakes (1984, p. 178) discuss how the likelihood needs to be modified to allow for left truncation.

Cox and Oakes (1984, p. 178) also point out the Kaplan-Meier product limit estimator mentioned earlier can easily be generalized to allow for left truncation. At any particular point in time we just need to consider how many animals are at risk. The animals at risk will now consist of animals present from the time origin who have not died or been censored plus any new animals which are now at risk.

4. Example

As an example of the use of this model, we present results from the first year of a multi-year study on movements and overwintering survival of black ducks being conducted by the U. S. Fish and Wildlife Service under Conroy's direction. Fifty female black ducks from two locations in New Jersey were captured and fitted with radios. The ducks were captured over a period of about four weeks from 8 November 1983 to 14 December 1983 and included 31 hatch-year birds

(birds born during the last breeding season) and 19 after-hatch-year birds (all birds at least one year of age). A condition index, body weight (g) divided by the wing length (mm), was calculated for each duck. The location and status (alive, missing, or dead) of each bird were recorded daily from the date of release until 15 February 1984, when the study was terminated. Diligent effort was made to locate each bird using roof mounted antennas on trucks, strut-mounted antennas on fixed-wing aircraft, and hand-held antennas on foot and by boat. The pertinent data on each of the fifty radio tagged ducks are given in Table 1. Eighteen birds died, seventeen survived to the end of the study, and the remaining 15 disappeared.

(Table 1 to appear here)

Three computer programs written specifically for the analysis of survival data were used in our analysis. PHGLM (Harrell, 1983) and SURVTEST (Therneau, 1983) are SAS procedures, while SURVREG (Preston and Clarkson, 1983) is an interactive FORTRAN program. Briefly, SURVTEST tests for differences between survival curves; PHGLM computes Kaplan-Meier (1958) or proportional hazard model estimates, fits Cox (1972) proportional hazards linear models, and can be used to test for differences between survival distributions; and SURVREG is similar to PHGLM in the estimates it calculates and its model building capabilities, but also provides four parametric distributions for use in estimation and model building.

The Kaplan-Meier estimates listed in Table 1 are plotted in Figure 1. This survival distribution is for the entire population of fifty black ducks, lumped across both locations and age classes. In obtaining these estimates we have also limited the interval of interest to the 63-day period from 15 December 1983 to 15 February 1984. This is in accordance with our suggestion of possibly measuring survival time from the release of the last animal captured. In this particular case, interest was only in overwintering survival and all fifty ducks were known to be alive on 15 December.

(Figure 1 to appear here)

Present techniques for analyzing data from radio-telemetry studies assume that each survival event is independent and has constant probability (Trent and Rongstad, 1974; Bart and Robson, 1982). Using this assumption in the continuous time framework the exponential survival distribution is obtained. Its fit using SURVREG is indicated by the solid line in Figure 1. While both curves yield an estimate of about 58 percent of the population surviving the 63-day period, the exponential underestimates in the early half of the period and overestimates in the latter half.

Using PHGLM and SURVTEST, we tested for differences between locations and between age classes. We could find no reason to reject the null hypothesis of no difference in either case. Small samples in some cells precluded testing for location-age class interaction.

Condition index (body weight/wing length) is used to gauge the general physical condition of a bird. Birds with a higher condition index are assumed to be in better physical condition than those with a lesser one. An a priori hypothesis to be tested in this study was that there is a significant relation between condition index and survival. When the population was separated on the basis of being above or below the median condition index, a significant difference in the survival distributions of the two groups was detected ($p<0.05$).

(Figure 2 to appear here)

A more detailed analysis of the effect of condition index on survival yielded the surprising result seen in Figure 2. Within each age class, the birds were separated into two groups using the median condition index value for that age class. Tests for differences in the survival distributions of these four groups suggested that (1) survival of hatch-year birds was not affected by their condition index, (2) survival of after-hatch-year birds was significantly affected by their condition index, (3) after-hatch-year birds with condition indices below the median had significantly lower probability of survival than did hatch-year birds and after-hatch-year birds with condition indices above the median, and (4) after-hatch-year birds with condition indices above the median had significantly higher survival probability than did hatch-year birds.

It can be seen in Figure 2 that no after-hatch-year birds with above median condition indices died during the study; all ten birds in the group represent censored observations. Clearly black ducks are not immortal, regardless of their physical condition. This suggests that censoring may not be random here; older birds with higher condition indices may be more prone to emigration than the other birds. As suggested above, if we assume no radios failed, an emigration distribution can be determined for these birds.

We point out that in all three programs, hypothesis tests breakdown if there are no deaths in a particular group. When all the observations for a group are censored observations, the asymptotic standard errors needed for testing are not appropriate. Figure 2, however, indicates that the differences we suggest are real.

The results presented in Figure 2 strongly suggest that condition index has a significant effect on overwintering survival in black ducks. The median is a convenient measure on which to divide a population, but may have no biological meaning. Perhaps the thirty-third percentile or two-thirds of the mean represent more biologically important values. As a further test of the importance of condition index we entered it as a covariate in a proportional hazards model (Cox, 1972), along with weight, wing length, and condition index squared. Using the procedure outlined in PHGLM, only condition index contributed significantly to a model for the entire population (χ_1^2 = 4.76: p = 0.029). The probability of survival increased as condition index increased or alternately the hazard function declined as the condition index increased ($\hat{\beta}$ = -1.678, SE ($\hat{\beta}$) = 0.800). When a model for only after-hatch-year birds was fit, the same result was obtained (χ_1^2 = 4.92: p = 0.027 with $\hat{\beta}$ = -2.628, SE ($\hat{\beta}$) = 1.375). As expected, condition index did not contribute significantly to a model for only hatch-year birds (χ_1^2 = 1.06: p = 0.303) but the regression coefficient had the same sign as before ($\hat{\beta}$ = -1.140, SE ($\hat{\beta}$) = 1.113).

We do not claim to have presented an exhaustive analysis of these data here. Our purpose is purely for illustration of the methodology. We do plan to do further analyses on hunting and non hunting deaths and obtain marginal survival curves for each type of death. It should be emphasized that competing risk theory (Kalbfleisch and Prentice 1980, p. 163) tells us that a model which allows dependency between the two different sources of mortality is nonidentifiable so that only marginal survival curves can be obtained.

5. General Discussion

The radio-tagged survival analysis procedure presented above provides a general framework for analyses of these studies. Radio-telemetry is likely to become an even more common technique as the technology improves and costs are reduced. The large body of statistical research into survival analysis in medicine and engineering should prove valuable and the necessary computer packages already exist.

We believe that the techniques we've outlined above, within the constraints of the listed assumptions, provide the researcher with a more realistic and sophisticated analysis than has heretofore been possible.

While many biologists will be most interested in the distribution free survival estimates themselves, most want the hypothesis testing and model fitting capabilities available. In particular we believe the testing of ecological hypotheses regarding the influence of individual animal covariates (such as condition index) on survival using the proportional hazards model is extremely important.

In the analysis discussed in this paper we have put most emphasis on the Kaplan-Meier product limit estimator because of its simplicity and generality. An important question is when should one use parametric modelling as compared to non parametric (Kaplan-Meier)? Previous approaches to analysis of radio-telemetry data (Trent and Rongstadt 1974, Bart and Robson, 1982) could be viewed as very special cases of parametric modelling. Although discrete their approaches are very similar to fitting an exponential distribution. Miller (1983) has done a comparison of maximum likelihood estimation and the Kaplan-Meier procedure when the underlying distribution is exponential and there is right censoring. As Miller (1983) points out this comparison is biased against the Kaplan-Meier estimator and its efficiency can be low. This is especially troublesome when t is large and Miller (1983) states "Parametric modelling should be considered as a means of increasing the precision in the estimation of small tail probabilities". He further states that it is surprising that so little work has been done on this question considering the importance of survival analysis in many disciplines (medicine, engineering).

Lagakos (1979) in a review paper on right censoring and survival analysis discusses informative censoring. Again it is surprising how little has been done on this problem. One practical approach is to calculate extreme bounds for the estimated survival curve by considering each censored observation to be either a death or a survivor until the end of the study. Of course if there is a lot of censoring early in the study these bounds can be very wide.

Finally we emphasize the importance of definition of the time origin in this application of survival analysis. In our example the survival functions only apply to the black duck population between early winter (December 15) and early spring (February 15). Also the extension to left truncation with right censoring which could be used if animals gradually enter the study is very important, and should be studied further.

ACKNOWLEDGEMENTS

We wish to thank Chris Bunck, Tom Gerig, Gary Hensler, Jim Nichols and Ken Williams for their assistance during the preparation of this article. Walter Kremers and a referee suggested the extension of the Kaplan-Meier estimator to left truncation. We also wish to thank Patuxent Wildlife Research Center for their support of Scott R. Winterstein on a research assistantship. The research assistantship came through the Cooperative Fisheries Unit at North Carolina State University and we wish to thank Mel Huish, Unit Leader, for his valuable assistance.

REFERENCES

Bart, J., and Robson, D. S. (1982). Estimating survivorship when the subjects are visted periodically. Ecology 63, 1078-1090.

Brownie, C., Anderson, D. R., Burnham, K. P., and Robson, D. S. (1978). Statistical inference from band recovery data: a handbook. U. S. Department of Interior, Fish and Wildlife Service Resource Publication No. 131. Washington, D. C.

Cox, D. R. (1972). Regression models and life tables (with discussion). Journal of the Royal Statistical Society Series B 34, 187-220.

Cox, D. R., and Oakes, D. (1984). Analysis of Survival Data. New York: Chapman and Hall.

Dunn, J. E., and Gipson, P. S. (1977). Analysis of radio-telemetry data in studies of home range. Biometrics 33, 85-101.

Harrell, F. E. (1983). The PHGLM procedure. SUGI Supplemental Library User's Guide. 1983 Edition. SAS Institute Inc., Cary, North Carolina.

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 Prentice, R. L. (1980). The Statistical Analysis of Failure Data. New York: Wiley.

Kaplan, E. L., and Meier, P. (1958). Nonparametric estimation from incomplete observations. Journal of the American Statistical Association 53, 457-481.

Lagakos, S. W. (1979). General right censoring and its impact on the analysis of survival data. Biometrics 35, 139-156.

Miller, R. G. (1983). What Price Kaplan-Meier? Biometrics 39, 1077-1081.

Pantula, S. G., and Pollock, K. H. (1985). Nested analysis of variance with autocorrelated errors. Biometrics (accepted).

Pollock, K. H. (1984). Estimation of survival distribution in ecology. Proceedings of 12th International Biometrics Conference. Tokyo, Japan.

Pollock, K. H. (1981). Capture-recapture models allowing for age-dependent survival and capture rates. Biometrics 37, 521-529.

Preston, D. L., and Clarkson, D. B. (1983). A User's Guide to SURVREG: Survival Analysis with Regression. American Statistician 37, 174. (Review).

Seber, G. A. F. (1965). A note on the multiple-recapture census. Biometrika 52, 249-259.

Therneau, T. M. (1983). The SURVTEST procedure. SUGI Supplemental Library User's Guide. 1983 Edition. SAS Institute Inc., Cary, North Carolina.

Trent, T. T., and Rongstad, O. J. (1974). Home range and survival of cottontail rabbits in Southwestern Wisconsin. Journal of Wildlife Management 38, 459-472.

White, G. C. (1983). Numerical estimation of survival rates from band-recovery and biotelemetry data. Journal of Wildlife Management 47, 716-728.

Table 1. Female black duck survival distribution (days).

Animal (i)	Age	Survival (t_i)	Indicator (δ_i)	Condition index*	$\widehat{F}(t_i)$+
1	AHY	2	1	4.188	0.9800
2	AHY	6	0	4.500	0.9800
3	HY	6	0	4.286	0.9800
4	HY	7	1	4.394	0.9591
5	AHY	13	1	4.045	0.9383
6	HY	14	0	4.275	0.9383
7	AHY	16	0	4.240	0.9170
8	AHY	16	1	4.115	0.9170
9	AHY	17	0	5.259	0.8951
10	AHY	17	1	4.167	0.8951
11	AHY	20	0	4.118	0.8951
12	AHY	21	1	4.096	0.8722
13	HY	22	1	3.992	0.8492
14	HY	26	1	4.576	0.8033
15	HY	26	1	3.730	0.8033
16	HY	27	1	4.226	0.7804
17	AHY	28	0	4.873	0.7804
18	HY	29	1	3.713	0.7567
19	AHY	32	0	4.529	0.7331
20	HY	32	1	3.852	0.7331
21	HY	34	1	4.741	0.6842
22	HY	34	1	4.348	0.6842
23	HY	37	1	4.596	0.6598
24	HY	40	1	3.964	0.6353
25	AHY	41	1	3.818	0.6109
26	HY	44	1	4.078	0.5865
27	HY	49	0	4.216	0.5865
28	AHY	54	0	4.632	0.5865
29	HY	56	0	4.007	0.5865
30	HY	56	0	4.556	0.5865
31	HY	57	0	4.601	0.5865
32	AHY	57	0	4.684	0.5865
33	HY	58	0	4.154	0.5865
34	HY	63	0	4.088	0.5865
35	HY	63	0	4.222	0.5865
36	HY	63	0	4.351	0.5865
37	HY	63	0	4.552	0.5865
38	HY	63	0	4.604	0.5865
39	AHY	63	0	4.982	0.5865
40	AHY	63	0	4.704	0.5865
41	HY	63	0	4.373	0.5865
42	HY	63	0	4.361	0.5865
43	AHY	63	0	3.818	0.5865
44	AHY	63	0	4.555	0.5865
45	HY	63	0	3.874	0.5865
46	HY	63	0	4.487	0.5865
47	HY	63	0	4.218	0.5865
48	HY	63	0	3.887	0.5865
49	HY	63	0	4.243	0.5865
50	AHY	63	0	4.111	0.5865

*Weight (g)/wing length (mm).

+Kaplan-Meier estimates obtained from SAS procedure PHGLM See text for details.

FIGURE LEGENDS

Figure 1. Survival distributions for female black ducks. X's are the Kaplan-Meier estimates obtained from SAS procedure PHGLM (Harrell 1983) and are the values listed in Table 1. The line represents the fit of the exponential equation

$$\text{Survival Probability} = e^{-(\text{Survival time}/114.722)}.$$

This is the curve obtained when the methodology suggested by Trent and Rongstad (1974) and Bart and Robson (1982) is employed. See text for further discussion.

Figure 2. Kaplan-Meier estimates for black ducks by age-class and condition index (body weight/wing length). Triangles are after-hatch-year birds with a condition index greater than the median for their age-class. X's are after-hatch-year birds with a condition index less than the median for their age-class. Squares are hatch-year birds. See text for details.

FIGURE 1

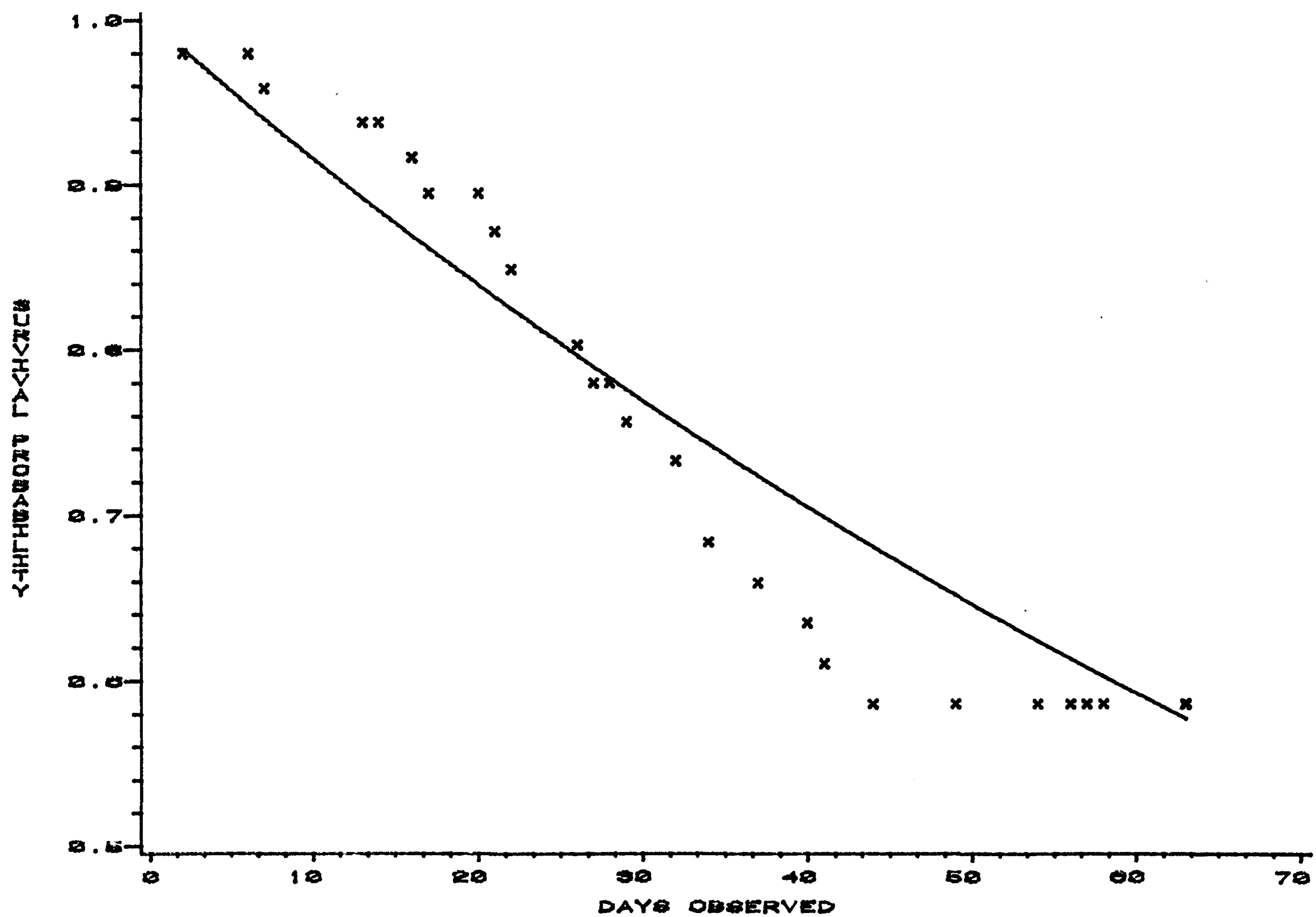
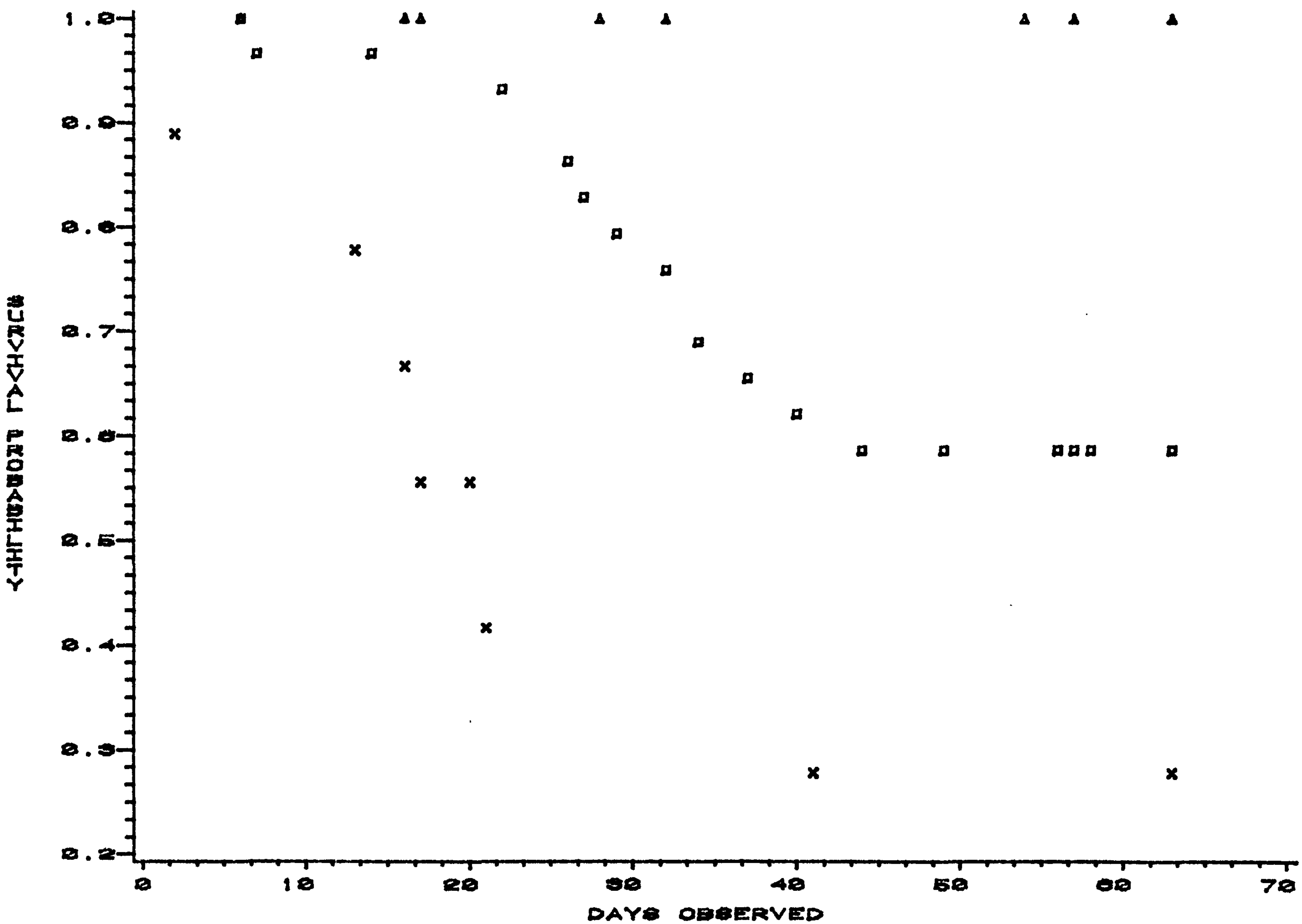

SURVIVAL PROBABILITY
1.0
0.9
0.8
0.7
0.6
0.5
0
10
20
30
40
50
60
70
DAYS OBSERVED

FIGURE 2


1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
SURVIVAL PROBABILITY
0
10
20
30
40
50
60
70
DAYS OBSERVED