

Analyzing Interval-Censored Data with the ICLIFETEST Procedure

Changbin Guo, Ying So, and Gordon Johnston, SAS Institute Inc.

ABSTRACT

SAS/STAT® 13.1 includes the new ICLIFETEST procedure, which is specifically designed for analyzing interval-censored data. This type of data is frequently found in studies where the event time of interest is known to have occurred not at a specific time but only within a certain time period. PROC ICLIFETEST performs nonparametric survival analysis of interval-censored data and is a counterpart to PROC LIFETEST, which handles right-censored data. With similar syntax, you use PROC ICLIFETEST to estimate the survival function and to compare the survival functions of different populations. This paper introduces you to the ICLIFETEST procedure and presents examples that illustrate how you can use it to perform analyses of interval-censored data.

INTRODUCTION

Time-to-event data are frequently found in clinical trials and medical studies. Usually, patients are followed up until a predefined event occurs. A patient is *right-censored* if observation of the time to event is prevented because the patient drops out of the study or the study ends. In certain situations, a complete follow-up is not possible because of the nature of the event or for budgetary reasons. For example, in a study of cosmetic deterioration after breast cancer surgery (Finkelstein and Wolfe 1985), patients are scheduled to visit the clinic every four to six months to have their status examined. The deterioration time is unlikely to be precisely observed unless it coincides with a clinic visit, and thus at best it can be known only to fall within an interval between two consecutive visits. This phenomenon is referred to as *interval-censoring*. In a sense, interval-censoring generalizes the typical right-censoring by treating right-censored observations as special intervals whose right boundary is infinity and treating exact event times as intervals of zero length.

The past several decades have seen scores of specialized statistical methods developed for analyzing interval-censored data; many of them have been generalizations from their right-censoring counterparts. For example, Turnbull (1976) developed a self-consistent algorithm to obtain the survival function estimates that is equivalent to the Kaplan-Meier estimator under right-censoring; Finkelstein (1986) generalized the log-rank test to interval-censored data. The common thread among all the new developments is that conventional theories, such as the counting processes and martingales that have been successfully applied to right-censored data, do not play well in the new arena, making it difficult to justify the methods theoretically. Many practical solutions have been proposed that usually rely on computationally intensive approaches such as the multiple imputation, permutation, or bootstrap method.

SAS/STAT offers a comprehensive set of procedures for analyzing right-censored data, including the LIFEREG, LIFETEST, PHREG, and QUANTLIFE procedures. Each specializes in a different aspect of statistical analysis. For example, PROC LIFETEST computes the nonparametric survival function estimates and makes comparisons between them by using formal statistical tests, whereas the other three procedures fit different models to the survival distribution. PROC LIFEREG is the only procedure that performs parametric analysis by fitting the accelerated failure time model for various failure time distributions, and it can handle interval-censored data.

PROC ICLIFETEST offers a set of nonparametric statistical methods of estimating survival functions and statistical testing for interval-censored data. From this perspective, it generalizes the LIFETEST procedure to handle interval-censored data. The rest of this paper introduces this new procedure, illustrates its usage through several examples, and compares its functionalities with those of PROC LIFETEST and several SAS macros.

FEATURES OF PROC ICLIFETEST

The main features of the ICLIFETEST procedure are as follows:

- **Nonparametric estimation**

- uses the efficient EMICM algorithm (Wellner and Zhan 1997) to estimate survival functions by default
- supports Turnbull's algorithm (Turnbull 1976) and the iterative convex minorant (ICM) algorithm (Groeneboom and Wellner 1992)
- computes standard errors of the survival estimates by using a multiple imputation method or a bootstrap method
- supports several transformation-based confidence intervals
- produces survival plots

- **Comparison of survival distributions**

- provides the weighted generalized log-rank test
- supports a variety of weight functions for testing early or late differences
- supports a stratified test for survival differences within predefined populations
- supports a trend test for ordered alternatives
- supports multiple-comparison functionalities

EXAMPLE: ANALYZING BREAST COSMESIS DATA

This example demonstrates how to use the ICLIFETEST procedure to conduct a nonparametric analysis of interval-censored data. The breast cancer data from Finkelstein and Wolfe (1985) are a typical example of interval-censored data. The data were collected from 94 patients in a retrospective study that compared the risks of breast cosmetic deterioration after tumorectomy. The participants received one of two treatments: radiation therapy or radiation plus chemotherapy. They visited the clinic about every four to six months, so the exact deterioration time could not be precisely observed. Also, because certain scheduled visits were missed, the intervals between observations were much longer. Thirty-eight patients had never experienced the event of cosmetic deterioration by the end of the study period, so these observations are right-censored.

The following statements create the SAS data set RT for the group that received radiation alone. The variable **lTime** provides the last follow-up time at which cosmetic deterioration had not occurred for the patient, and the variable **rTime** provides the last follow-up time immediately after the event. Note that for the ICLIFETEST procedure to recognize the observations as right-censored, their right bounds must be represented by missing values in the input data set.

```
data RT;
  input lTime rTime @@;
  trt = 'RT    ';
  datalines;
45 . 25 37 37 .
 6 10 46 . 0 5
 0 7 26 40 18 .
46 . 46 . 24 .
46 . 27 34 36 .
 7 16 36 44 5 11
17 . 46 . 19 35
 7 14 36 48 17 25
37 44 37 . 24 .
 0 8 40 . 32 .
 4 11 17 25 33 .
15 . 46 . 19 26
11 15 11 18 37 .
```

```

22 . 38 . 34 .
46 . 5 12 36 .
46 .
;

```

The following statements create the SAS data set RCT for patients who received both radiation and chemotherapy:

```

data RCT;
  input lTime rTime @@;
  trt = 'RT+RCT';
  datalines;
  8 12 0 5 30 34
  0 22 5 8 13 .
  24 31 12 20 10 17
  17 27 11 . 8 21
  17 23 33 40 4 9
  24 30 31 . 11 .
  16 24 13 39 14 19
  13 . 19 32 4 8
  11 13 34 . 34 .
  16 20 13 . 30 36
  18 25 16 24 18 24
  17 26 35 . 16 60
  32 . 15 22 35 39
  23 . 11 17 21 .
  44 48 22 32 11 20
  14 17 10 35 48 .
;

```

The following statements combine the data sets RT and RCT into a single data set named BCS that is to be analyzed by PROC ICLIFETEST:

```

data BCS;
  set RT RCT;
run;

```

The goal is to determine how the time to deterioration is distributed. The following statements invoke the ICLIFETEST procedure to estimate the survival functions for both treatment groups:

```

ods graphics on;
proc iclifetest data=BCS impute(seed=1234);
  strata trt;
  time (lTime, rTime);
run;

```

In the TIME statement, the variables that represent the interval boundaries, **lTime** and **rTime**, are enclosed in parentheses and separated by a comma. Because the treatment indicator variable, **Trt**, is specified in the STRATA statement, PROC ICLIFETEST conducts the analysis separately for each treatment group. By default, the procedure plots the estimated survival functions; this is equivalent to specifying PLOT=SURVIVAL in the PROC ICLIFETEST statement. You can specify an integer seed for the random number generator that is used to create imputed data sets for calculating standard errors of the survival estimates. If you do not specify the SEED= option, a random seed is obtained from the computer's clock.

The results of this analysis are shown in [Figure 1](#), [Figure 2](#), [Figure 3](#), and [Figure 4](#).

Figure 1 Data and Methods Information

The ICLIFETEST Procedure

Data and Methods Information	
Data Set	WORK.BCS
Left Boundary	lTime
Right Boundary	rTime
Estimation Method	EMICM
Number of Observations Read	94
Number of Observations Used	94
Number of Imputations for Standard Errors	1000
Imputation Seed	1234

As displayed in Figure 1, PROC ICLIFETEST recognizes **lTime** and **rTime** as the left and right boundaries, respectively, of the input intervals. The default estimation method is EMICM (Wellner and Zhan 1997). All the observations in the data set can be used. By default, PROC ICLIFETEST performs 1,000 imputations to estimate the standard errors of the survival estimates.

Figure 2 displays the nonparametric survival estimates for the radiation group (TRT=RT).

Figure 2 Nonparametric Survival Estimates

The ICLIFETEST Procedure

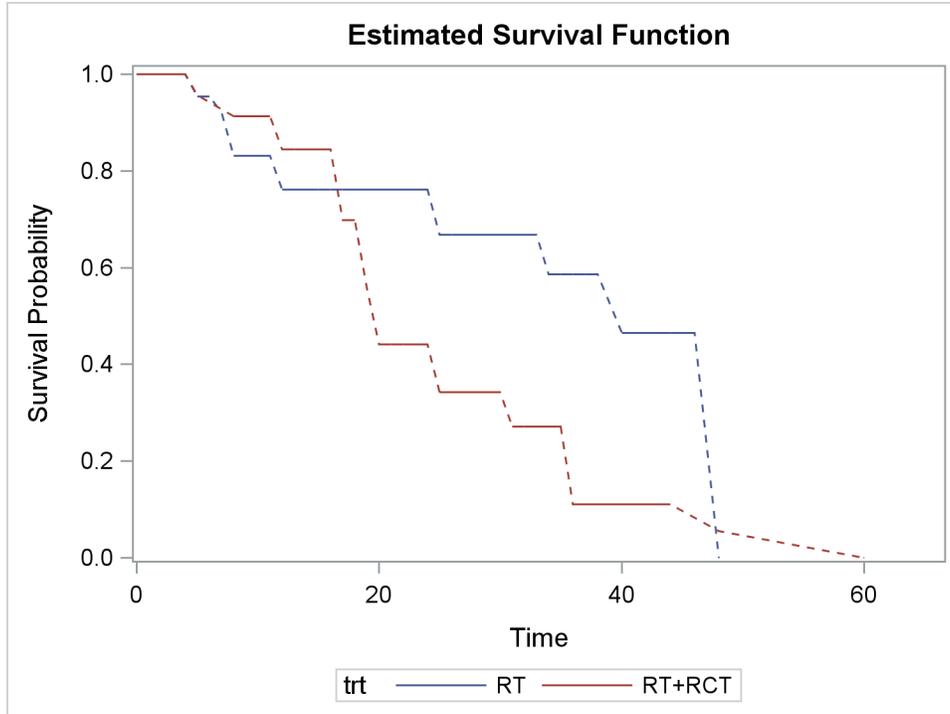
Stratum 1: trt = RT

Nonparametric Survival Estimates				
Time Interval		Probability Estimate		Imputation Standard Error
Start	End	Failure	Survival	
0	4	0.0000	1.0000	0.0000
5	6	0.0463	0.9537	0.0354
7	7	0.0797	0.9203	0.0458
8	11	0.1684	0.8316	0.0580
12	24	0.2391	0.7609	0.0629
25	33	0.3318	0.6682	0.0706
34	38	0.4136	0.5864	0.0739
40	46	0.5344	0.4656	0.0758
48	Inf	1.0000	0.0000	0.0000

Note that the nonparametric estimates of failure probability and survival probability are available only for a set of nonoverlapping intervals, and they are constant within each interval. For example, consider the time interval (8, 11]. The estimated failure probability between 8 and 11 is 0.1684. The interval (11, 12], which is not displayed, is an interval for which the survival estimate cannot be uniquely determined; such intervals are referred to as *Turnbull intervals*. The failure probability increases to 0.2391 at 12 and remains constant up to 24. Figure 2 also shows standard errors for the estimates.

Figure 3 displays the estimated survival functions for the two treatments.

Figure 3 Nonparametric Survival Estimates by Treatment



Clearly, the group that received radiation alone tends to go longer before experiencing cosmetic deterioration than the group that received both radiation and chemotherapy. As shown in Figure 2, the estimated survival probabilities are undetermined within the Turnbull intervals. For ease of visualization, dashed lines are plotted across the Turnbull intervals for which the estimates are not defined.

The ICLIFETEST procedure also produces a summary of the frequencies of various types of censoring (Figure 4). The categories are exact times, left-censoring, right-censoring, and interval-censoring. The percentage of each category within a stratum is also calculated.

Figure 4 Summary of Various Censoring Types

		Number of Censored and Uncensored Values				
		Type of Censoring				
Stratum		Total	Left	Interval	Right	Uncensored
ID	trt					
1	RT	46	3 (6.5%)	18 (39.1%)	25 (54.3%)	0 (0.0%)
2	RT+RCT	48	2 (4.2%)	33 (68.8%)	13 (27.1%)	0 (0.0%)
Total		94	5 (5.3%)	51 (54.3%)	38 (40.4%)	0 (0.0%)

The following statements use a generalized log-rank test to formally evaluate whether patients in the two treatment groups have the same survival rate:

```
proc iclifetest data=BCS impute(seed=1234);
  time (lTime, rTime);
  test trt;
run;
```

You can specify an integer seed for the random number generator that is used to create imputed data sets for obtaining standard errors of the survival estimates and for performing the generalized log-rank test. If you do not specify the SEED= option, a random seed is obtained from the computer's clock.

The results of this analysis are shown in Figure 5, Figure 6, Figure 7, Figure 8, and Figure 9. The "Data and Methods Information" table in Figure 5 contains two more rows than the table in Figure 1. PROC

ICLIFETEST uses 1,000 imputations to estimate the covariance matrix of the generalized log-rank statistics. It also applies the default weights (WEIGHT=SUN) to the test. These weights are constant over time so that all time points are treated equally.

Figure 5 Data and Methods Information

The ICLIFETEST Procedure

Data and Methods Information	
Data Set	WORK.BCS
Left Boundary	lTime
Right Boundary	rTime
Estimation Method	EMICM
Number of Observations Read	94
Number of Observations Used	94
Number of Imputations for Standard Errors	1000
Number of Imputations for Testing	1000
Imputation Seed	1234
Weight Function for Generalized Log-Rank Test	SUN

Figure 6 presents the generalized log-rank statistics for the two treatment groups. Figure 7 shows the covariance estimate of the generalized log-rank statistics.

Figure 6 Generalized Log-Rank Statistics

The ICLIFETEST Procedure

Generalized Log-Rank Statistics	
trt	Log-Rank
RT	-9.14185
RT+RCT	9.141845

Figure 7 Covariance of Generalized Log-Rank Statistics

Covariance Matrix for the Generalized Log-Rank Statistics		
trt	RT	RT+RCT
RT	11.3939	-11.3939
RT+RCT	-11.3939	11.3939

Figure 8 presents the chi-square statistic and the corresponding p -value that are obtained by using the chi-square distribution with one degree of freedom.

Figure 8 Generalized Log-Rank Test

Test of Equality over Group			
Weight	Chi-Square	DF	Pr > Chi-Square
SUN	7.3349	1	0.0068

Because there are only two treatment groups, the generalized log-rank statistics have the same magnitude but opposite signs. You can reproduce the chi-square statistic by using one generalized log-rank statistic and the corresponding variance as $(-9.14185)^2/11.3939 = 7.3349$.

You can use a different weight function to perform the generalized log-rank test by specifying the WEIGHT= option. Different weights emphasize different parts of the survival curve. The following statements perform the test by using weights derived by Fay (1996), which give more weight to the early times:

```
proc iclifetest data=BCS impute(seed=1234);
  time (lTime, rTime);
  test trt / weight=fay;
run;
```

As Figure 9 shows, the p -value of 0.0294 is larger than the previous result of 0.0068 under constant weights. This difference is expected, because the new weight function $\hat{S}(t)$ gives more weight to the earlier times, when the two survival curves are closer together.

Figure 9 Generalized Log-Rank Test

The ICLIFETEST Procedure

Test of Equality over Group			
			Pr >
Weight	Chi-Square	DF	Chi-Square
FAY	4.7454	1	0.0294

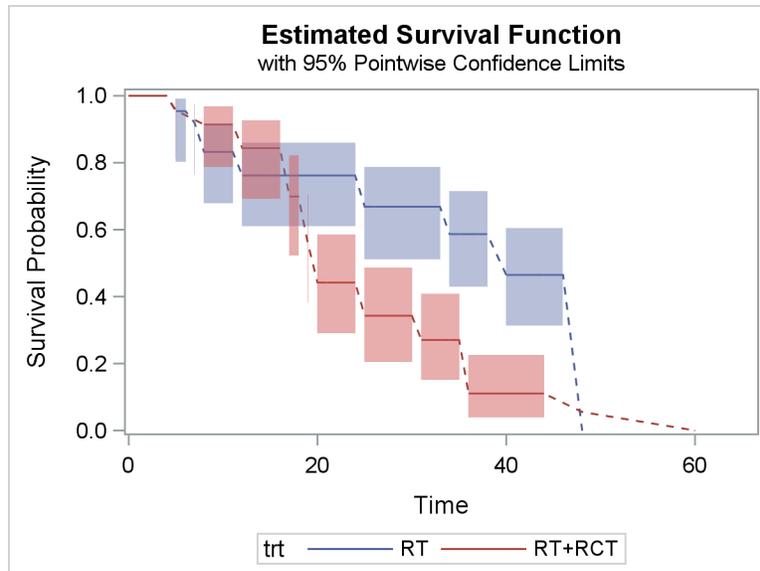
CUSTOMIZING SURVIVAL PLOTS

The PLOT= option in PROC ICLIFETEST is very similar to the PLOT= option in PROC LIFETEST. You can request a plot of the pointwise confidence limits by using the CL option:

```
proc iclifetest data=BCS plot=s(cl) impute(seed=1234);
  strata trt;
  time (lTime, rTime);
run;
```

Figure 10 displays the new plot.

Figure 10 Survival Curves with 95% Confidence Limits

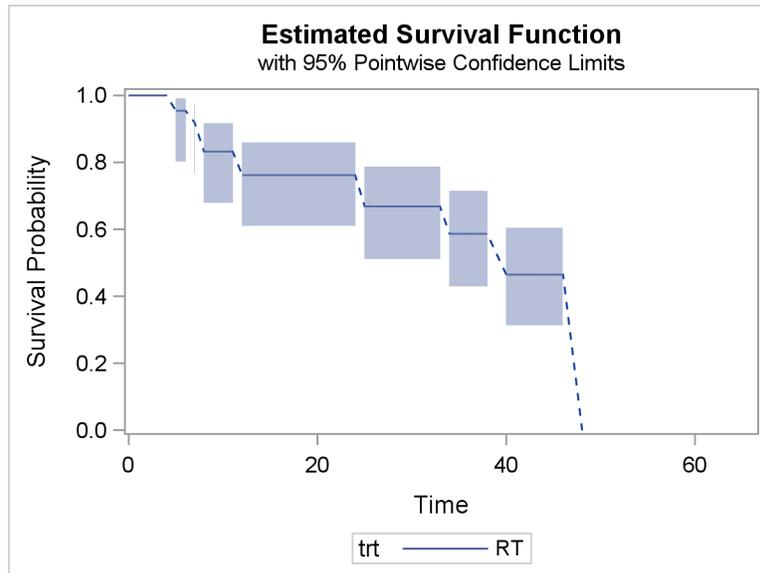


By default, the STRATA=OVERLAY option is used to produce overlaid plot for multiple survival curves. You can request a plot of individual survival curves by using the STRATA=INDIVIDUAL option:

```
proc iclifetest data=BCS plot=s(cl strata=individual) impute(seed=1234);
  strata trt;
  time (lTime, rTime);
run;
```

Figure 11 displays the new plot.

Figure 11 Plot of One Survival Curve

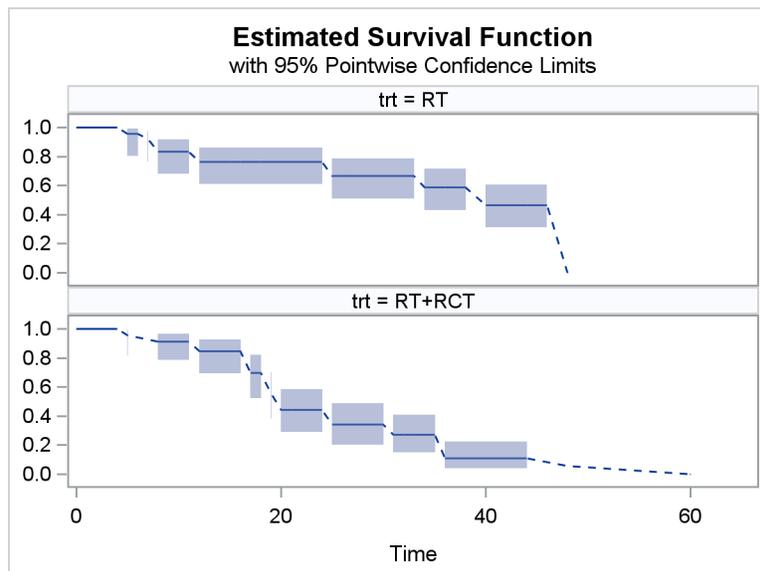


If there are many curves, you can use the STRATA= PANEL option to display the curves in panels:

```
proc iclifetest data=BCS plot=s(cl strata=panel) impute(seed=1234);
  strata trt;
  time (lTime, rTime);
run;
```

Figure 12 displays the new plot.

Figure 12 Paneled Survival Plot

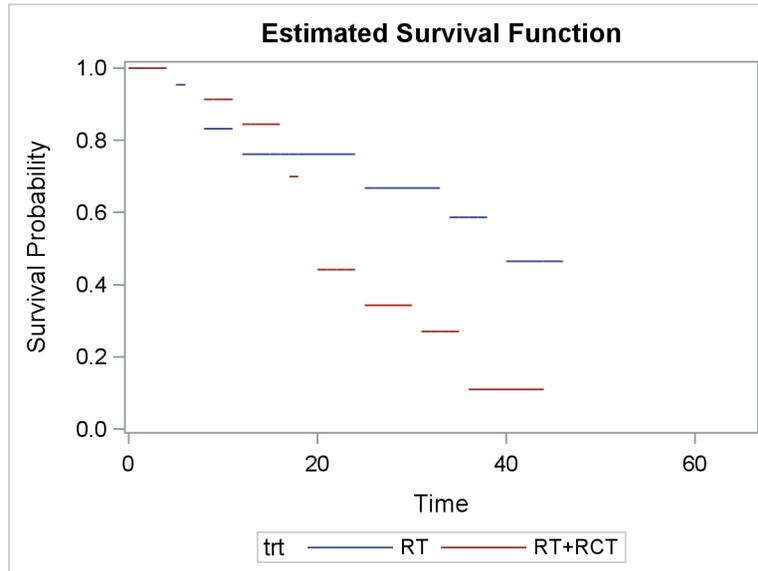


You can suppress plotting of the dashed lines for the Turnbull intervals by using the NODASH option:

```
proc iclifetest data=BCS plot=s(nodash) impute(seed=1234);
  strata trt;
  time (lTime, rTime);
run;
```

Figure 13 displays the new plot.

Figure 13 Disjoint Survival Estimates



DESCRIPTION OF STATISTICAL METHODS

This section provides some background information about the statistical methods that are implemented in PROC ICLIFETEST. It discusses two types of analysis that PROC ICLIFETEST performs: estimating survival functions and comparing survival distributions.

Estimation of Survival Functions

Usually, the first step in analyzing interval-censored data is to estimate the survival function. By definition, the survival function is a function of time that measures the probability that a subject will survive beyond a given time. Although it is easy to apply the Kaplan-Meier method to obtain the survival estimates for right-censored data, this method does not extend directly to interval-censored data.

Suppose $T_i, i = 1, \dots, n$, are the survival times for n subjects based on survival function $S(t)$. Further suppose that the subjects are being followed up periodically, so for subject i , the survival time is known to be within the interval (L_i, R_i) .

Peto (1973) and Turnbull (1976) prove that the nonparametric survival estimate can jump over only the so-called Turnbull intervals $(q_1, p_1), \dots, (q_m, p_m)$. These intervals are determined by the data as follows: first, all the boundary values L_i, R_i with labels of L and R attached are ordered, and then the intervals that have L as the left boundary and R as the right boundary are picked up. For example, suppose the data set contains only three intervals, $(1,3)$, $(2,4)$, and $(5,6)$. The ordered values are 1(L), 2(L), 3(R), 4(R), 5(L), and 6(R), yielding two intervals, $(2,3)$ and $(5,6)$.

Let $\theta_j = P(t \in (q_j, p_j)), j = 1, \dots, m$. The likelihood function for $\theta = \{\theta_j, j = 1, \dots, m\}$ is then

$$L(\theta) = \prod_{i=1}^n \left(\sum_{j=1}^m \alpha_{ij} \theta_j \right)$$

where α_{ij} is 1 if $I_j = (q_j, p_j)$ is contained in $(L_i, R_i]$ and 0 otherwise.

Recognizing that the optimization becomes more difficult when the number of parameters is large, Turnbull (1976) developed a self-consistent algorithm to obtain the estimates,

$$\theta_j = \frac{1}{n} \sum_{i=1}^n \mu_{ij}(\boldsymbol{\theta}) = \frac{1}{n} \sum_{i=1}^n \frac{\alpha_{ij} \theta_j}{\sum_{j=1}^m \alpha_{ij} \theta_j}$$

where $\mu_{ij}(\boldsymbol{\theta})$ is the expected probability that the event T_i occurs within I_j for the i th subject, given the observed data.

Suppose your data generate m Turnbull intervals as $\{I_j = (q_j, p_j], j = 1, \dots, m\}$. Denote the probability estimate for the j th interval as $\hat{\theta}_j$, and denote the nonparametric survival estimate as $\hat{S}(t) = \sum_{k: p_k > t} \hat{\theta}_k$ for $t \notin$ any I_j .

The algorithm is in fact an expectation-maximization (EM) algorithm (Dempster, Laird, and Rubin 1977), in the sense that it iteratively updates the parameters $\boldsymbol{\theta}$ and the expected quantities $\mu_{ij}(\boldsymbol{\theta})$. Convergence is declared if, for a chosen number $\epsilon > 0$,

$$\sum_{j=1}^m |\hat{\theta}_j^{(l)} - \hat{\theta}_j^{(l-1)}| < \epsilon$$

where $\hat{\theta}_j^{(l)}$ denotes the updated value for θ_j after the l th iteration.

Unfortunately, it is not guaranteed that the converged values are the global maximum and thus constitute a maximum likelihood estimate (MLE). Gentleman and Geyer (1994) introduced the Kuhn-Tucker conditions based on constrained programming to check whether the algorithm converges to a legitimate MLE. These conditions state that a necessary and sufficient condition for the estimate to be an MLE is that the Lagrange multipliers $\gamma_j = n - c_j$ are nonnegative for all the θ_j 's that are estimated to be zero, where c_j is the derivative of the log-likelihood function with respect to θ_j :

$$c_j = \frac{\partial \log(L)}{\partial \theta_j} = \sum_{i=1}^n \frac{\alpha_{ij}}{\sum_{j=1}^m \alpha_{ij} \theta_j}$$

Two alternative methods to Turnbull's algorithm have been developed: the iterative convex minorant (ICM) algorithm (Groeneboom and Wellner 1992) and the hybrid EMICM algorithm (Wellner and Zhan 1997). The latter algorithm is particularly appealing because it converges to the MLE if it exists and is unique. PROC ICLIFETEST uses the EMICM algorithm as the default estimation method.

Standard approaches to obtaining the standard errors of the survival estimates based on the likelihood function are no longer justified, because the number of unknown parameters might increase as the sample size increases. Simulations (Goodall, Dunn, and Babiker 2004) show that estimates obtained this way tend to be inflated. Several alternative ad hoc methods have been proposed. The ICLIFETEST procedure implements two resampling approaches proposed by Sun (2001) that are built on the multiple imputation and bootstrap methods, respectively.

Let M denote the number of resampling data sets. The undefined Turnbull intervals of the nonparametric survival estimates present a special challenge to using the bootstrap method. Because each bootstrap sample could generate a different set of Turnbull intervals, some time points for evaluating the variances based on the original Turnbull intervals might be located within the intervals in a bootstrap sample, with the result that their survival probabilities become unknown. A simple ad hoc solution is to shrink the Turnbull interval to its right boundary and convert the survival estimates to a right continuous function:

$$\hat{S}_m(t) = \sum_{j: p_j > t} \hat{\theta}_j$$

Let A_1^k, \dots, A_n^k denote the n independent samples from the original data with replacement, $k = 1, \dots, M$. Let $\hat{S}_m^k(t)$ be the modified estimate of the survival function computed from the k th resampling data set. Then you can estimate the variance of $\hat{S}(t)$ by the sample variance as

$$\hat{\sigma}_b^2(t) = \frac{1}{M-1} \sum_{k=1}^M [\hat{S}_m^k(t) - \bar{S}_m(t)]^2$$

where

$$\bar{S}_m(t) = \frac{\sum_{k=1}^M \hat{S}_m^k(t)}{M}$$

The method of multiple imputations exploits the fact that interval-censored data reduce to right-censored data when all interval observations of finite length shrink to single points ($L_i = R_i$). You obtain the k th imputed data set by replacing every interval-censored observation of finite length with an exact event time randomly drawn from the discrete survival function:

$$\hat{S}_i(p_j) = \frac{\hat{S}(q_j) - \hat{S}(R_i+)}{\hat{S}(L_i) - \hat{S}(R_i+)}, \quad q_j \in (L_i, R_i] \quad j = 1, \dots, m$$

Denote the Kaplan-Meier estimate of each imputed data set as $\hat{S}_{KM}^k(t)$. The variance of $\hat{S}(t)$ is estimated by

$$\hat{\sigma}_I^2(t) = \hat{S}^2(t) \sum_{q_j < t} \frac{d'_j}{n'_j(n'_j - d'_j)} + \frac{1}{M-1} \sum_{k=1}^M \left[\hat{S}_{KM}^k(t) - \bar{S}_{KM}(t) \right]^2$$

where

$$\bar{S}_{KM}(t) = \frac{1}{M} \sum_{k=1}^M \hat{S}_{KM}^k(t)$$

and

$$d'_j = \sum_{i=1}^n \frac{\alpha_{ij} [\hat{S}(p_{j-1}) - \hat{S}(p_j)]}{\sum_{j=1}^m \alpha_{ij} [\hat{S}(p_{j-1}) - \hat{S}(p_j)]}$$

and

$$n'_j = \sum_{k=j}^m d'_k$$

Comparison of Survival Distributions

For right-censored data, the log-rank test and its weighted family (Fleming and Harrington 1991) have been the most commonly used methods of comparing survival distributions between different groups. For interval-censored data, Finkelstein (1986) derives a score test based on the proportional hazards model, and it closely resembles the original test. Based on the discrete logistic model, Sun (1996) obtains a test statistic that has the same algebraic representation as the original log-rank test. Fay (1996) obtains another test under the proportional odds model, and he categorizes all three tests as a family of weighted log-rank tests in an attempt to construct a unified framework of testing equality of survival functions for interval-censored data.

Let $S_k(t)$ be the underlying survival function of the k th group, $k = 1, \dots, K$. The null and alternative hypotheses to be tested are, respectively,

$$H_0 : S_1(t) = S_2(t) = \dots = S_K(t) \text{ for all } t$$

$$H_1 : \text{at least one of the } S_k(t)\text{'s is different for some } t$$

Let N_k denote the number of subjects in group k , and let n denote the total number of subjects ($n = N_1 + \dots + N_K$). For the i th subject, let $\mathbf{z}_i = (z_{i1}, \dots, z_{iK})^T$ be a vector of K indicators that represent whether or not the subject belongs to the k th group.

Under the null hypothesis, the test statistic for the k th group takes the form

$$U_k = \sum_{j=1}^m U_{kj} = \sum_{j=1}^m v_j \left(d'_{kj} - \frac{n'_{kj}}{n'_j} d'_j \right)$$

where v_j are the weights, depending on the particular model specified (Fay 1999). The statistic for all the groups is $\mathbf{U} = (U_1, \dots, U_K)$.

If the number hypothesis is true, d'_{kj} estimates the expected number of events within I_j for the k th group, and it is computed as

$$d'_{kj} = \sum_{i=1}^n z_{ik} \frac{\alpha_{ij} \hat{\theta}_j}{\sum_{l=1}^m \alpha_{il} \hat{\theta}_l}$$

where d'_j is an estimate of the expected number of events within I_j for the whole sample and is computed as

$$d'_j = \sum_{k=1}^K d_{kj}$$

Similarly, n'_{kj} estimates the expected number of subjects at risk before entering I_j for the k th group; it can be estimated by $n'_{kj} = \sum_{l=j}^m d'_{kl}$. And n'_j is an estimate of the expected number of subjects at risk before entering I_j for all the groups: $n'_j = \sum_{k=1}^K n'_{kj}$.

The choices of weight function that PROC ICLIFETEST supports are given in Table 1.

Table 1 Weight Functions for Various Tests

Test	Weights
Sun (1996)	1.0
Fay (1996)	$\hat{S}(p_{j-1})$
Finkelstein (1986)	$\frac{\hat{S}(p_{j-1})[\log \hat{S}(p_{j-1}) - \log \hat{S}(p_j)]}{\hat{S}(p_{j-1}) - \hat{S}(p_j)}$
Harrington-Fleming (p, q)	$[\hat{S}(p_{j-1})]^p [1 - \hat{S}(p_{j-1})]^q, p \geq 0, q \geq 0$

The conventional score test uses the observed or expected information as an estimate of the variance of the score statistic and computes a p -value based on the asymptotic chi-square distribution under the null hypothesis. This becomes problematic in the interval-censored setting. Unless stringent assumptions are made, it is not clear whether the variance estimate is accurate and whether the asymptotic performance can be characterized by the chi-square distribution. Ad hoc methods have been developed. PROC ICLIFETEST implements the method of multiple imputations developed by Huang, Lee, and Yu (2008).

In order to generate an imputed data set, you need to randomly generate a survival time for every subject in the sample. For the i th subject, a time T_i^* is randomly generated based on the discrete survival function

$$\hat{S}_i(T_i^* = p_j) = \frac{\hat{S}(q_j) - \hat{S}(R_i +)}{\hat{S}(L_i) - \hat{S}(R_i +)}, \quad q_j \in (L_i, R_i], \quad j = 1, \dots, m$$

where $(L_i, R_i]$ denotes the interval observation for the subject.

For the h th imputed data set ($h = 1, \dots, H$), let d_{kj}^h and n_{kj}^h denote the numbers of failures and subjects at risk, respectively, by counting the imputed T_i^* 's for group k . Let d_j^h and n_j^h denote the corresponding pooled numbers.

You can perform the standard log-rank test for right-censored data on each of the imputed data sets (Huang, Lee, and Yu 2008). The test statistic is

$$\mathbf{U}^h = (U_1^h, \dots, U_K^h)^T$$

where

$$U_k^h = \sum_{j=1}^m v_j \left(d_{kj} - \frac{n_{kj}}{n_j} d_j \right)$$

Its variance-covariance matrix is estimated by the Greenwood formula as

$$\mathbf{V}^h = \mathbf{V}_1^h + \dots + \mathbf{V}_m^h$$

where

$$(\mathbf{V}_j^h)_{l_1 l_2} = \begin{cases} v_j^2 n_{l_1 j}^h (n_j^h - n_{l_1 j}^h d_j^h (n_j^h - d_j^h) (n_j^h)^{-2} (n_j^h - 1)^{-1}) & \text{when } l_1 = l_2 \\ -v_j^2 n_{l_1 j}^h n_{l_2 j}^h d_j^h (n_j^h - d_j^h) (n_j^h)^{-2} (n_j^h - 1)^{-1} & \text{when } l_1 \neq l_2 \end{cases}$$

After analyzing each imputed data set, you can estimate the variance-covariance matrix of \mathbf{U} by pooling the results as

$$\hat{\mathbf{V}} = \frac{1}{H} \sum_{h=1}^H \mathbf{V}^h - \frac{1}{H-1} \sum_{h=1}^H [\mathbf{U}^h - \bar{\mathbf{U}}] [\mathbf{U}^h - \bar{\mathbf{U}}]^T$$

where

$$\bar{\mathbf{U}} = \frac{1}{H} \sum_{h=1}^H \mathbf{U}^h$$

The overall test statistic is formed as $\mathbf{U}^T \mathbf{V}^- \mathbf{U}$, where \mathbf{V}^- is the generalized inverse of \mathbf{V} . Under the null hypothesis, the statistic has a chi-square distribution with degrees of freedom equal to the rank of \mathbf{V} .

Besides the multiple imputation method, an alternative approach to the testing is to use permutations (Fay and Shih 2012). Note that an equivalent way to write the weighted log-rank test statistic is

$$U_k = \sum_{i=1}^n z_{ik} a_i$$

where a_i is the subject-specific score for the i th subject.

This form suggests that a conventional permutation test can be performed on the individual scores. Because censoring is present, the validity of the permutation test depends on the assumption that the groups under comparison are subject to the same mechanism of censoring. Although simulation studies have suggested that the impact of unequal censoring tends to be negligible in many realistic settings (Fay and Shih 2012), PROC ICLIFETEST has not yet incorporated the permutation test.

COMPARISON WITH PROC LIFETEST

The ICLIFETEST procedure compares most closely to the LIFETEST procedure. The two procedures share the same analytical objectives: estimating and summarizing subjects' survival experiences and comparing them systematically. The main distinction between these procedures lies in the types of data that they are designed to analyze. The ICLIFETEST procedure is primarily intended to handle interval-censored data, whereas the LIFETEST procedure deals exclusively with right-censored data. You can use PROC ICLIFETEST to analyze data that are left-censored, interval-censored, or right-censored. However, if the data to be analyzed contain only exact or right-censored observations, it is recommended that you use PROC LIFETEST, because it has specialized methods of dealing with right-censored data that are more efficient.

Similarities and Differences in Syntax

For the most part, PROC ICLIFETEST shares a similar syntax with PROC LIFETEST. That means that you can build on your knowledge of the LIFETEST procedure to get started with PROC ICLIFETEST. A major difference is that in PROC ICLIFETEST, the TEST statement has been redesigned to perform statistical testing and the STRATA statement no longer performs the testing by default. Unlike the LIFETEST procedure, PROC ICLIFETEST does not support a similar test to detect association with multiple covariates. The STRATA statement has two functions. If you specify it without the TEST statement, it requests separate analysis for each group that the STRATA variables generate. If you specify both the STRATA and TEST statements, the testing is performed by stratifying on the populations as defined by the STRATA variables. The variables in the STRATA statement must differ from the variables in the TEST statement.

The following code examples demonstrate the difference in syntax when you perform the stratified test by using PROC ICLIFETEST and PROC LIFETEST:

```

proc iclifetest;
  strata s;
  time (l, r);
  test g;
run;

proc lifetest;
  time t*c(0);
  strata s / group=g;
run;

```

PROC ICLIFETEST provides a trend test for testing an ordered alternative hypothesis when you have more than two groups. The formation of the test statistic is the same as in PROC LIFETEST. The following code examples demonstrate the difference in syntax when you perform the trend test by using PROC ICLIFETEST and PROC LIFETEST:

```

proc iclifetest;
  time (l, r);
  test g / trend;
run;

proc lifetest;
  time t*c(0);
  strata g / trend;
run;

```

Differences in Statistical Methods

Both PROC ICLIFETEST and PROC LIFETEST estimate the survival functions nonparametrically. The difference is that in PROC LIFETEST, the Kaplan-Meier estimate produces a right-continuous step function, whereas the estimate that PROC ICLIFETEST produces is undefined for the Turnbull intervals. For statistical inferences about the whole survival curve, the most common way to fill the gaps is to move the probability mass associated with a Turnbull interval to the right end. This practice can be justified by the theory that the new estimate is still asymptotically unbiased (Groeneboom and Wellner 1992).

Most statistical methods that PROC ICLIFETEST implements are generalizations of the methods of analyzing right-censored data. So you might expect that the results of applying these methods to right-censored data will be the same as you would get by using the right-censored methods. This is true in some cases, but exceptions are not uncommon. For example, the nonparametric algorithms in PROC ICLIFETEST should yield the same results as the Kaplan-Meier method when applied to right-censored data because they maximize the same nonparametric likelihood function. On the other hand, you would not expect the generalized log-rank test to produce the same result as the log-rank test, because the two tests differ in how they deal with right-censored observations.

COMPARISON WITH THE %EMICM AND %ICSTEST MACROS

Two SAS macros, %EMICM and %ICSTEST, were specially developed to analyze interval-censored data. The %EMICM macro estimates survival functions by using the same estimation algorithms as PROC ICLIFETEST. The difference lies in how exact observations ($L = R$) are handled. The %EMICM macro applies an ad hoc transformation by subtracting 0.0001 from L . In contrast, PROC ICLIFETEST adopts the proposal by Ng (2002) so that exact observations remain a different class.

The %ICSTEST macro implements a generalized log-rank test (Zhao and Sun 2004) for comparing survival curves between different groups. This test computes the same test statistic that PROC ICLIFETEST computes under the default option (WEIGHT=SUN). Like PROC ICLIFETEST, the %ICSTEST macro uses multiple imputations to compute the covariance matrix, and the estimate consists of two parts. However, the second part is subtracted from the first part, whereas PROC ICLIFETEST does the opposite. Huang, Lee, and Yu (2008) provide a statistical perspective to compare the two approaches.

The %ICSTEST macro also offers the test developed by Sun, Zhao, and Zhao (2005). Its test statistic resembles that of the permutation form of the generalized log-rank test (Fay 1999). The %ICSTEST macro does not support the stratified test, trend test, or multiple comparisons.

SUMMARY

The new ICLIFETEST procedure in SAS/STAT is a counterpart to the LIFETEST procedure. You use PROC LIFETEST for right-censored data, and you use PROC ICLIFETEST for interval-censored data. PROC ICLIFETEST provides a set of nonparametric statistical methods of estimating survival functions and comparing survival distributions.

Table 2 compares the major capabilities of PROC LIFETEST and PROC ICLIFETEST.

Table 2 Similarities In Capability

Capability	PROC LIFETEST	PROC ICLIFETEST
Censoring type	Right	Interval
Nonparametric survival estimation	Yes	Yes
Methods of standard errors	Greenwood formula	Multiple imputations or bootstrap
Survival equality testing	Yes	Yes
Stratified estimation	Yes	Yes
Stratified testing	Yes	Yes
Trend test	Yes	Yes
Customized plots	Yes	Yes
Hazard plot	Yes	Yes

Table 3 lists the major differences in syntax between PROC LIFETEST and PROC ICLIFETEST.

Table 3 Syntax Differences Between LIFETEST and ICLIFETEST

Feature	PROC LIFETEST	PROC ICLIFETEST
Statement for testing groups	STRATA	TEST
Statement for stratified testing	STRATA with GROUP= option	STRATA and TEST
Tests for survival equality	Log-rank test (Default) Wilcoxon test (Default)	Constant weight (Default or WEIGHT=SUN) Wilcoxon-type weight (WEIGHT=FAY)

REFERENCES

- Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977), "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society, Series B*, 39, 1–38.
- Fay, M. P. (1996), "Rank Invariant Tests for Interval Censored Data under the Grouped Continuous Model," *Biometrics*, 52, 811–822.
- Fay, M. P. (1999), "Comparing Several Score Tests for Interval Censored Data," *Statistics in Medicine*, 18, 273–285.
- Fay, M. P. and Shih, J. H. (2012), "Weighted Logrank Tests for Interval Censored Data When Assessment Times Depend on Treatment," *Statistics in Medicine*, 31, 3760–3772.
- Finkelstein, D. M. (1986), "A Proportional Hazards Model for Interval-Censored Failure Time Data," *Biometrics*, 42, 845–854.
- Finkelstein, D. M. and Wolfe, R. A. (1985), "A Semiparametric Model for Regression Analysis of Interval-Censored Failure Time Data," *Biometrics*, 41, 933–945.

- Fleming, T. R. and Harrington, D. P. (1991), *Counting Processes and Survival Analysis*, New York: John Wiley & Sons.
- Gentleman, R. and Geyer, C. J. (1994), "Maximum Likelihood for Interval Censored Data: Consistency and Computation," *Biometrika*, 81, 618–623.
- Goodall, R. L., Dunn, D. T., and Babiker, A. G. (2004), "Interval-Censored Survival Time Data: Confidence Intervals for the Non-parametric Survivor Function," *Statistics in Medicine*, 23, 1131–1145.
- Groeneboom, P. and Wellner, J. A. (1992), *Information Bounds and Nonparametric Maximum Likelihood Estimation*, New York: Birkhäuser.
- Huang, J., Lee, C., and Yu, Q. (2008), "A Generalized Log-Rank Test for Interval-Censored Failure Time Data via Multiple Imputation," *Statistics in Medicine*, 27, 3217–3226.
- Ng, M. P. (2002), "A Modification of Peto's Nonparametric Estimation of Survival Curves for Interval-Censored Data," *Biometrics*, 58, 439–442.
- Peto, R. (1973), "Experimental Survival Curves for Interval-Censored Data," *Applied Statistics*, 22, 86–91.
- Sun, J. (1996), "A Nonparametric Test for Interval-Censored Failure Time Data with Application to AIDS Studies," *Statistics in Medicine*, 15, 1387–1395.
- Sun, J. (2001), "Variance Estimation of a Survival Function for Interval-Censored Survival Data," *Statistics in Medicine*, 20, 1249–1257.
- Sun, J., Zhao, Q., and Zhao, X. (2005), "Generalized Log-Rank Test for Interval-Censored Failure Time Data," *Scandinavian Journal of Statistics*, 32, 49–57.
- Turnbull, B. W. (1976), "The Empirical Distribution Function with Arbitrarily Grouped, Censored, and Truncated Data," *Journal of the Royal Statistical Society, Series B*, 38, 290–295.
- Wellner, J. A. and Zhan, Y. (1997), "A Hybrid Algorithm for Computation of the Nonparametric Maximum Likelihood Estimator from Censored Data," *Journal of the American Statistical Association*, 92, 945–959.
- Zhao, Q. and Sun, J. (2004), "Generalized Log-Rank Test for Mixed Interval-Censored Failure Time Data," *Statistics in Medicine*, 23, 1621–1629.

ACKNOWLEDGMENTS

The authors thank Bob Rodriguez and Ed Huddleston for their valuable editorial assistance in preparing this paper.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors:

Changbin Guo	Ying So	Gordon Johnston
SAS Institute Inc.	SAS Institute Inc.	SAS Institute Inc.
SAS Campus Drive	SAS Campus Drive	SAS Campus Drive
Cary, NC 27513	Cary, NC 27513	Cary, NC 27513
changbin.guo@sas.com	ying.so@sas.com	gordon.johnston@sas.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.