

Paper SAS4286-2020

Recent Developments in Survival Analysis with SAS® Software

G. Gordon Brown, SAS Institute Inc.

ABSTRACT

Are you interested in analyzing lifetime and survival data in SAS® software? SAS/STAT® and SAS® Visual Statistics offer a suite of procedures and survival analysis methods that enable you to overcome a variety of challenges that are frequently encountered in time-to-event data. This paper brings you up to date on new approaches and procedures in SAS software and provides an overview on when to use these procedures to overcome the challenges **inherent in conducting survival analysis in today's world. Procedures and methods that you** will learn about include performing model selection and model comparison with the PHSELECT procedure, fitting the Fine and Gray model and fitting Bayesian frailty models with the PHREG procedure, analyzing accelerated failure time models with the LIFEREG procedure, and handling interval-censored data with the ICLIFETEST and ICPHREG procedures. You will see how to deal with nonproportional hazards by using the LIFETEST procedure or the RMSTREG procedure.

INTRODUCTION

Special techniques are required to analyze survival data for two primary reasons: the data are almost always incomplete, and familiar parametric assumptions are frequently unjustifiable. Investigators usually follow subjects until they reach a prespecified endpoint (for example, death). However, subjects sometimes withdraw from a study, or the study is completed before the endpoint is reached. In these cases, the survival times (also known as failure times) are censored because the survival status beyond the censoring time is unknown. The uncensored survival times are sometimes called event times. Methods of survival analysis must account for both censored and uncensored data.

This paper presents eight SAS/STAT and SAS Visual Statistics procedures that enable you to overcome a variety of challenges that are frequently encountered in time-to-event data and guides you to the appropriate procedures and analytic methodologies for your data and research questions. The paper briefly describes the primary functionality of these procedures and provides links to papers of interest and online documentation that discuss these methods in detail.

Table 1 lists seven typical topics in survival analysis and the procedures that you can use for each topic. Each topic is described further in a section of the paper. All procedures except the PHSELECT procedure are available in SAS/STAT; the PHSELECT procedure is available only in SAS Visual Statistics. A license for SAS Visual Statistics also gives you access to SAS/STAT procedures.

Topic	Procedure	Procedure Description
Descriptive analysis <ul style="list-style-type: none"> • Create a graphical representation of the survival curves • Request statistical tests 	LIFETEST	Nonparametric methods for right-censored data
	ICLIFETEST	Nonparametric methods for interval censored data
Regression analysis <ul style="list-style-type: none"> • Perform regression analysis with respect to the time-to-event outcome • Perform Bayesian analysis of the proportional hazards model¹ 	PHREG	Proportional hazards models
	LIFEREG	Accelerated failure time models
	RMSTREG	Regression analysis of restricted mean survival time
Nonproportional hazards Perform data analysis that is based on the restricted mean survival time (RMST) when the proportional hazards assumption is violated	RMSTREG	Regression analysis of the RMST
	LIFETEST	Nonparametric estimation and hypothesis testing
Competing risks Perform regression analysis when multiple causes of failure exist	PHREG	Proportional hazards models
	LIFETEST	Nonparametric methods for right-censored data
Interval-censored data Perform data analysis when an event is known only to lie within an interval	ICLIFETEST	Nonparametric estimation and hypothesis testing
	ICPHREG	Parametric and semi-parametric proportional hazards models
Variable selection Perform automated variable selection to choose a subset of independent effects for the proportional hazards model	PHREG	Proportional hazards models
	PHSELECT	Variable selection for the proportional hazards model including LASSO
Sample survey design Perform data analysis of the proportional hazards model when the sample is selected from a sample design	SURVEYPHREG	Proportional hazards models with a sample design

Table 1. Survival Analysis Topics and Procedures

DESCRIPTIVE ANALYSIS

Conducting descriptive analysis for survival data typically implies plotting survival functions and calculating summary statistics. The LIFETEST and ICLIFETEST procedures in SAS/STAT enable you to create these plots of the survival curves. The two procedures share the same **analytic objectives: estimating and summarizing subjects' survival experiences and comparing them systematically**. The ICLIFETEST procedure is intended primarily for handling interval-censored data, whereas the LIFETEST procedure deals exclusively with right-censored data. You can use the ICLIFETEST procedure 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 the LIFETEST procedure because it provides specialized methods for dealing with right-censored data.

¹ Bayesian analysis is supported by the BAYES statement in the PHREG and LIFEREG procedures. The RMSTREG procedure does not support Bayesian analysis.

LIFETEST PROCEDURE

If you have exact or right-censored data, you can use the LIFETEST procedure to compute nonparametric estimates of the survival function either by the product-limit method (also called the Kaplan-Meier method) or by the lifetable method (also called the actuarial method). A variety of hypothesis tests available in PROC LIFETEST enable you to examine the equality of the survival functions. You can visually examine the survival plot to check for nonproportional hazards. ODS Graphics automatically creates the graphical representation of the survival plots; methods you can use for modifying the survival plots are described in Kuhfeld and So (2013).

Online documentation for [the LIFETEST procedure](#)

ICLIFETEST PROCEDURE

The ICLIFETEST procedure is a counterpart of the LIFETEST procedure for analyzing interval-censored data. You can use the ICLIFETEST procedure to compute nonparametric estimates of the survival functions and to examine the equality of the survival functions through statistical tests. If you want to learn more about the ICLIFETEST procedure, see Guo, So, and Johnston (2014).

Online documentation for [the ICLIFETEST procedure](#)

REGRESSION ANALYSIS

Many types of models have been used for survival data. Two of the more popular types of models are the accelerated failure time model, available in the LIFEREG procedure, and the Cox proportional hazards model, available in the PHREG procedure. The accelerated failure time model assumes a parametric form for the effects of the explanatory variables and usually assumes a parametric form for the underlying survivor function. The Cox proportional hazards model also assumes a parametric form for the effects of the explanatory variables, but it allows an unspecified form for the underlying survivor function. Besides these classical methods, an alternative approach for analyzing time-to-event data is to conduct regression with respect to the restricted mean survival time (RMST), which is available in the RMSTREG procedure. The RMST is defined as the expected value of the time-to-event variable up to a prespecified time point τ . The focus of each procedure, the type of model employed, and the method used for parameter estimation are shown in Table 2.

Procedure	Focus	Model Type	Estimation Method
PHREG	Hazard function	Proportional hazards models	Partial likelihood
LIFEREG	Time-to-event	Accelerated failure time models	Likelihood
RMSTREG	Restricted mean survival time	Generalized linear models	Estimating equations

Table 2. Survival Modeling Procedures

PHREG PROCEDURE

The PHREG procedure fits the Cox proportional hazards model and its extensions, which include recurrent events models, shared frailty models, and models for competing-risks data. PROC PHREG also enables you to assess the predictive accuracy of survival models by using concordance statistics and time-dependent ROC curves. The BAYES statement enables you to perform Bayesian inference for the models available in this procedure, including Bayesian frailty models. For more information about performing Bayesian analysis, see Stokes, Chen, and Gunes (2015). For more information about concordance statistics and time-dependent ROC curves, see Guo, Ying, and Jang (2017).

Online documentation for [the PHREG procedure](#)

LIFEREG PROCEDURE

The LIFEREG procedure fits accelerated failure time models to failure time data, including data that are left-censored, right-censored, or interval-censored. The models for the response variable consist of a linear effect that is composed of the covariates and a random disturbance term. The distribution options for the random disturbance term include the extreme value, normal, logistic, exponential, Weibull, lognormal, log-logistic, and three-parameter gamma distributions.

Online documentation for [the LIFEREG procedure](#)

RMSTREG PROCEDURE

The restricted mean survival time (RMST) is defined as the expected value of the time-to-event variable up to a prespecified time point; it is an alternative measure that is more often reliably estimable than the mean and median of the event time. Survival analysis that uses the RMST has attracted practitioners for its straightforward interpretation (notably parameter estimates that are based on means rather than hazards ratios) and its capability to deal with nonproportional hazards. You use the RMSTREG procedure to fit linear and log-linear models to the RMST. For more information about using the RMSTREG procedure, see Guo and So (2019).

Online documentation for [the RMSTREG procedure](#)

NONPROPORTIONAL HAZARDS

Classical hazard-based methods can be inappropriate or inefficient when the proportional hazards (PH) assumption is violated. A typical violation of the PH assumption occurs when two or more of the survival curves cross.

LIFETEST PROCEDURE

To check the validity of the proportional hazards assumption, you can use the LIFETEST procedure to produce a survival plot, as shown in Figure 1. If the proportional hazards assumption is violated, you can use the LIFETEST procedure (or the RMSTREG procedure, which is described in the previous section) to perform survival analysis. Both the LIFETEST procedure and the RMSTREG procedure analyze the restricted mean survival time (RMST), which is the expected value of the time-to-event variable up to a prespecified time horizon.

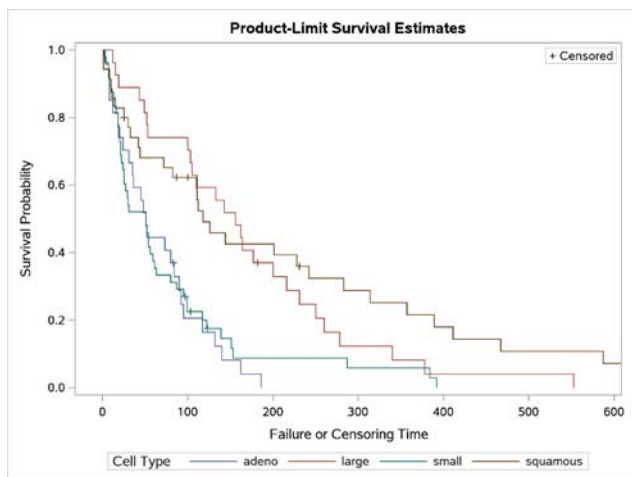


Figure 1. Survival Plot Showing Nonproportional Hazards

You can use the RMST option in the LIFETEST procedure to perform nonparametric analysis of the RMST. PROC LIFETEST enables you to estimate the RMST at times of interest and make appropriate group comparisons by performing multiple comparisons.

Online documentation for [the LIFETEST procedure](#)

RMSTREG PROCEDURE

You can also use the RMSTREG procedure to analyze nonproportional hazards. For more information about PROC RMSTREG, see its earlier description in the section "RMSTREG Procedure."

Online documentation for [the RMSTREG procedure](#)

COMPETING RISKS

Competing risks occur in studies where individuals are subject to several potential failure events and where the occurrence of one event might impede the occurrence of other events. The LIFETEST and PHREG procedures provide a comprehensive set of techniques for the analysis of competing-risks data.

LIFETEST PROCEDURE

The LIFETEST procedure focuses on nonparametric analysis and includes specialized features for analyzing the cumulative incidence function. You can use PROC LIFETEST to compute the nonparametric estimate of the cumulative incidence function and investigate group differences by using **Gray's (1988) test**.

Online documentation for [the LIFETEST procedure](#)

PHREG PROCEDURE

The PHREG procedure provides the following two regression approaches for analyzing competing-risks data:

- **Fine and Gray's method to directly model the cumulative incidence function** (Fine and Gray 1999)
- cause-specific proportional hazards models

PROC PHREG also enables you to make model-based predictions of the cumulative incidence functions. If you want to learn more about using SAS software to analyze competing risks, see So, Lin, and Johnston (2014) and Guo and So (2018).

Online documentation for [the PHREG procedure](#)

INTERVAL CENSORING

The ICLIFETEST and ICPHREG procedures are specifically designed to analyze interval-censored data; they are counterparts of the LIFETEST and PHREG procedures, which primarily deal with right-censored data. **Interval censoring happens when you don't know the exact time when an event occurred, but you do know the period during which it occurred.** The survival plot in Figure 2 is created from the ICLIFETEST procedure and demonstrates interval censored data.

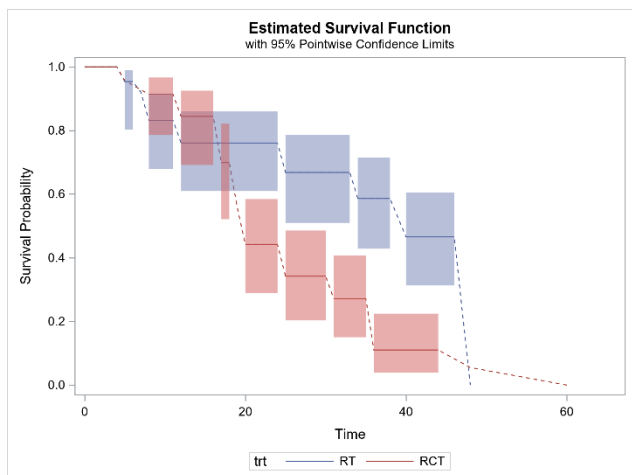


Figure 2. Survival Plot for Interval Censoring

ICLIFETEST PROCEDURE

The ICLIFETEST procedure implements nonparametric statistical methods of analyzing interval-censored data. You can compute nonparametric estimates of survivor functions and compare the underlying survival distributions of two or more samples of interval-censored data.

Online documentation for [the ICLIFETEST procedure](#)

THE ICPHREG PROCEDURE

The ICPHREG procedure fits proportional hazards regression models to interval-censored data. It enables you to fit parametric and semiparametric models. For parametric models, you select a piecewise constant function, a cubic spline, or a discrete hazard as the baseline hazard function. The semiparametric models in PROC ICPHREG are analyzed on the basis of the nonparametric likelihood function, and you can make statistically efficient inferences about the model effects without prespecifying the baseline hazard function.

Online documentation for [the ICPHREG procedure](#)

VARIABLE SELECTION

You can perform variable selection for the Cox proportional hazards model by using automated variable selection methods in the PHSELECT procedure in SAS Visual Statistics or

the PHREG procedure in SAS/STAT. The PHSELECT procedure has more options for variable selection than the PHREG procedure has.

THE PHSELECT PROCEDURE

The PHSELECT procedure in SAS Visual Statistics fits the Cox proportional hazards regression model for survival data and performs variable selection.

The models that PROC PHSELECT supports can contain main effects that consist of both continuous and classification variables and interaction effects of these variables. The models can also include constructed effects such as splines. The procedure offers several effect selection methods, including stepwise methods and modern LASSO methods. It also offers extensive capabilities for customizing the model selection by using a wide variety of selection and stopping criteria, from those that are computationally efficient and based on a significance level to those that are modern, computationally intensive, and based on validation. PROC PHSELECT also provides a variety of Cox regression diagnostics that are conditional on the selected model.

Online documentation for [the PHSELECT procedure](#)

PHREG PROCEDURE

The PHREG procedure in SAS/STAT provides stepwise model selection methods including backward, forward, score, and stepwise. PROC PHREG enables you to specify significance levels for entry and removal of effects, add effects in a sequential order, specify the number of variables in the model for forward or backward selection, and select the best subsets. The PHREG procedure does not offer the LASSO method, which is available in the PHSELECT procedure.

Online documentation for [the PHREG procedure](#).

SURVEY SAMPLE DESIGN

Data sets frequently have a sample design that includes elements such as strata, clusters, and analysis weights. The purpose of the sample design is to ensure that an analysis returns results that have appropriate point and variance estimates. If you choose to ignore the sample design, you risk obtaining biased parameter and variances estimates, which can lead to low-quality inference and poor decisions. In order to correctly account for a sample design when you have time-to-event data, it is highly recommended that you use the SURVEYPHREG procedure.

SURVEYPHREG PROCEDURE

The SURVEYPHREG procedure in SAS/STAT offers the Cox proportional hazards model for survival data (the same model that is available in the PHREG procedure).

Many of the options available in the other survival procedures are not available for data that have elements of a sample design. To correctly analyze survey data, it is highly recommended that you use a procedure that has built in capabilities to account for a sample design. Typically, inference obtained from a **procedure that can't account for a sample design** is suspect and requires additional adjustments.

If you want to learn more about PROC SURVEYPHREG, see Mukhopadhyay (2010) and Gardiner (2015).

Online documentation for [the SURVEYPHREG procedure](#)

CONCLUSION

This paper provides an overview of the SAS/STAT and SAS Visual Statistics procedures you can use to analyze survival data. If you want to learn more about the statistical procedures available in SAS, visit support.sas.com/statistics. For each procedure, you will find documentation, detailed examples that show how to use the syntax and how to interpret the results, papers that highlight important features of the procedures, and short videos about using the procedures.

REFERENCES

- Fine, J. P., and Gray, R. J. (1999). "A Proportional Hazards Model for the Subdistribution of a Competing Risk." *Journal of the American Statistical Association* 94: 496–509.
- Gardiner, J. C. (2016). "Competing Risks Analyses: Overview of Regression Models." In *Proceedings of the SAS Global Forum 2016 Conference*. Cary, NC: SAS Institute Inc. Available at <https://support.sas.com/resources/papers/proceedings15/2040-2015.pdf>
- Gray, R. J. (1988). "A Class of K-Sample Tests for Comparing the Cumulative Incidence of a Competing Risk." *Annals of Statistics* 16: 1141–1154.
- Guo, C., and So, Y. (2018). "Cause-Specific Analysis of Competing Risks Using the PHREG Procedure." In *Proceedings of the SAS Global Forum 2018 Conference*. Cary, NC: SAS Institute Inc. Available at <https://www.sas.com/content/dam/SAS/support/en/sas-global-forum-proceedings/2018/2159-2018.pdf>
- Guo, C., and So, Y. (2019). "Analyzing Restricted Mean Survival Time Using SAS/STAT®." In *Proceedings of the SAS Global Forum 2019 Conference*. Cary, NC: SAS Institute Inc. Available at <https://www.sas.com/content/dam/SAS/support/en/sas-global-forum-proceedings/2019/3013-2019.pdf>
- Guo, C., So, Y., and Jang, W. (2017). "Evaluating Predictive Accuracy of Survival Model with PROC PHREG." In *Proceedings of the SAS Global Forum 2017 Conference*. Cary, NC: SAS Institute Inc. Available at <http://support.sas.com/resources/papers/proceedings17/SAS0462-2017.pdf>
- Guo, C., So, Y., and Johnston, G. (2014). "Analyzing Interval-Censored Data with the ICLIFETEST Procedure." In *Proceedings of the SAS Global Forum 2014 Conference*. Cary, NC: SAS Institute Inc. Available at <http://support.sas.com/resources/papers/proceedings14/SAS279-2014.pdf>
- Kuhfeld, W., and So, Y. (2013). "Creating and Customizing the Kaplan-Meier Survival Plot in PROC LIFETEST." In *Proceedings of the SAS Global Forum 2013 Conference*. Cary, NC: SAS Institute Inc. Available at <http://support.sas.com/resources/papers/proceedings13/427-2013.pdf>
- Mukhopadhyay, P. (2010). "Not Hazardous to Your Health: Proportional Hazards Modeling for Survey Data with the SURVEYPHREG Procedure." In *Proceedings of the SAS Global Forum 2010 Conference*. Cary, NC: SAS Institute Inc. Available at <http://support.sas.com/resources/papers/proceedings10/254-2010.pdf>
- So, Y., Lin, G., and Johnston, G. (2014). "Using the PHREG Procedure to Analyze Competing-Risks Data." In *Proceedings of the SAS Global Forum 2014 Conference*. Cary, NC: SAS Institute Inc. Available at <http://support.sas.com/rnd/app/stat/papers/2014/competingrisk2014.pdf>
- Stokes, C., Chen, F., and Gunes, F. (2015). "An Introduction to Bayesian Analysis with the SAS/STAT Software." In *Proceedings of the SAS Global Forum 2014 Conference*. Cary, NC: SAS Institute Inc. Available at <http://support.sas.com/resources/papers/proceedings15/SAS1775-2015.pdf>

ACKNOWLEDGMENTS

The author is grateful to Changbin Guo and Pushpal Mukhopadhyay for their helpful suggestions and comments and to Anne Baxter for her editing.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

G. Gordon Brown
SAS Institute Inc.
SAS Campus Drive
Cary, NC 27513
Gordon.Brown@sas.com
support.sas.com/statistics

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.