

A Method to Generate Kaplan-Meier and Adjusted Survival Curves using SAS

Bharat Thakkar, Kwan Hur, William G. Henderson, and Charles Oprian
Department of Veteran affairs cooperative Studies Program Hines, IL

Abstract

Kaplan-Meier(K-M) survival curves can be generated using SAS PROC LIFETEST. In the presence of significant covariates, adjusting the survival curves is necessary. In this paper a step by step method is presented to produce survival curves that adjusted for the covariates. This is accomplished using SAS PROC PHREG and SAS PROC GLPOT.

Introduction

In clinical research often a comparison of survival curves is made between treatment groups. The SAS PROC LIFETEST produces K-M survival curves that depend only on time. However, the survival curves may depend on both time and covariates. Here, adjusting the survival curves for these covariates is necessary.

The problem faced is to devise a method of generating survival curves adjusted for the covariates. To obtain the adjusted survival curves, one must adjust for the covariates that were statistically significant in the regression model. This regression model then estimates the survival function.

This paper presents a method to generate survival curves adjusted for the covariates from a Cox proportional hazard model. An example taken from the VA Cooperative Study #90, "Prognosis and Outcome Following Heart Valve Replacement", consisting of a total of 938 aortic patients (299 medical patients and 639 surgical patients) was used to illustrate this method.

Method

Model Building

Significant predictors of survival can be identified using Cox regression model if the proportional hazard assumption is valid. A stepwise procedure can be used to select the significant covariates. For example, the following SAS code produces a stepwise model including the entry or removal of the

covariates at each step.

Other selection methods such as forward selection or backward elimination may be

```
proc phreg=clean;
  model times*censor(0)= treat var_1- var_p /
  selection = stepwise details;
run;
used to identify significant covariates. The variables that are statistically significant in the final step of the selection procedure are the significant covariates in the model.
```

Survival Function Estimation

After determining the covariates in the final model PROC PHREG is run with the BASELINE statement. The BASELINE statement in PROC PHREG creates a new data set which contains the survival function estimates at the event times of each stratum for each list of covariates in the COVARIATE = *data set*. By default the data set contains the survival function estimates corresponding to the means of the covariates for each stratum.

One may be interested in computing the survival function estimates for a set of covariates with specific values other than the means. Here, the COX regression model can be rerun with the COVARIATES= *data set* in the BASELINE statement. This data set contains specific values in the model. For example, the following SAS code will generate the survival function estimates for the set of given covariates.

```
proc phreg data = clean;
  model times*censor(0)=treat var_1 - var_p;
  baseline covariates = cov out = result
  survival =s nomean;
run;
```

Example

A subset of 938 aortic patients taken from the VA Study #90 was used for the analysis. There were 299 medical patients and 639 surgical patients followed for 20 years. The observed survival rate for the medical group was 28.76 % (86

patients), compared to 27.86 % (178 patients) for the surgical group. The medical group had a mean survival time of 8.08 years with a standard deviation of 6.36, while the surgical group had a mean survival time of 8.97 years with a standard deviation of 5.80.

The K-M survival curves were calculated using PROC LIFETEST and are displayed in Figure 1.

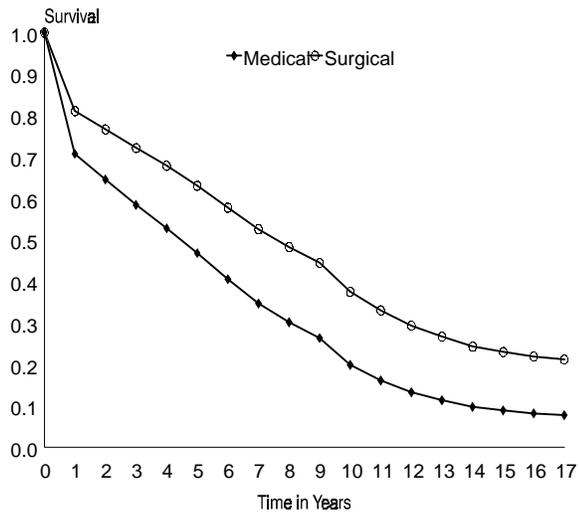
Figure 1. K-M Survival Curves

The surgical subgroup had longer survival times than the medical subgroup.

The Cox regression model for this data included the following covariates: treatment group, age, functional class, hypertension, smoking, hemoglobin, and number of occluded arteries. Using the means of these covariates in the Cox regression model the estimated survival functions were calculated for each subgroup.

The data set that contains the estimated survival functions can be read by PROC GPLOT. Figure 2 displays the adjusted survival curves for each subgroup.

Figure 2. Adjusted Survival Curves



After adjusting for the significant covariates the survival times for the surgical subgroup were longer than the medical subgroup. The difference in the survival probabilities between the two subgroups after 12 years is more pronounced in the adjusted survival curves than the K-M survival curves.

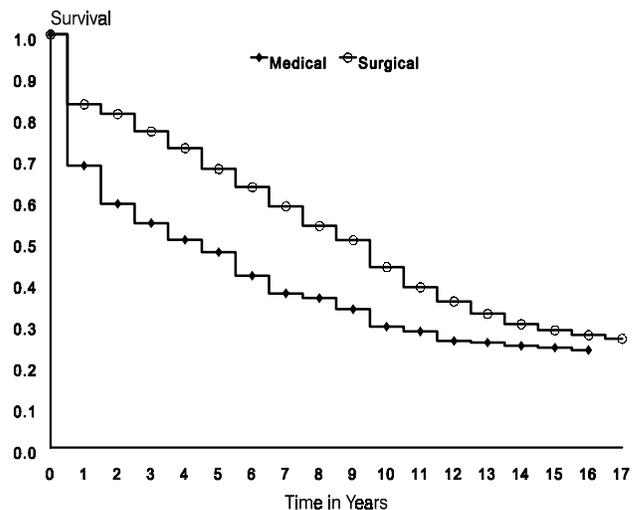
Conclusion

With the method described above it is possible to create K-M survival curves and survival curves adjusted for the covariates in a regression model. A comparison between the K-M and adjusted survival curves as illustrated in the above example shows the importance of adjusting for the significant covariates.

Adjusted survival curves can be constructed for different sets of covariates. In this manner the survival functions corresponding to different covariates may be compared.

References

1. SAS Institute Inc. (1990), "The LifeTest Procedure", Chapter 26 of SAS/STAT User's Guide Version 6, Fourth Edition, Volume 2, Cary, NC SAS Institute Inc.
2. Marubini, Ettore and Valsecchi, Maria Grazia Analyzing Survival Data from Clinical Trials and Observational Studies New York, NY John Wiley & Sons, 1995
3. SAS Institute Inc. (1992), "The PHREG Procedure", Chapter 19 of SAS Technical Report P-229, Cary, NC SAS Institute Inc.
4. Collett, D Modelling Survival Data in Medical Research London, England Chapman & Hall, 1994



Author Contact

Department of Veteran Affairs
Edward Hines Jr. VA Hospital
Bharat Thakkar 151A
5th and Roosevelt road
Hines IL. 60141
(708)343-7200 x3555
e-mail:
Thakkar@Research.Hines.Med.VA.GOV