

ANALYSIS OF SURVIVAL DATA: A COMPARISON OF THREE MAJOR STATISTICAL PACKAGES (SAS, SPSS, BMDP)

Corey J. Pelz and John P. Klein, Medical College of Wisconsin
 Corey J. Pelz, Medical College of Wisconsin, 8701 Watertown Plank Rd., Milwaukee, WI 53226

KEY WORDS: Survival analysis, SAS, SPSS, BMDP

Survival analysis techniques have become standard tools for the statistician in medical research. The application of survival models to data is valid when the endpoint of interest is the "time to the occurrence of a particular event." Survival models may be applied to a variety of fields such as biology, medicine, engineering, and economics. With modern computing technology, the analysis of "time-to-event" data has become inexpensive in terms of time. There are several statistical packages on the market today that can be used to do survival analyses. The most commonly used packages are SAS, SPSS, and BMDP. These three packages are compared based upon their capabilities, accuracy, and user-friendliness as applied to survival analysis. Example data sets are used to demonstrate standard and nonstandard conditions that occur when modelling survival data in each of the packages. Several survival analysis applications are presented to determine the agreement among the three packages. Both the univariate and multivariate survival analysis procedures are presented for each package.

1. INTRODUCTION

The application of survival models to data is valid when the endpoint of interest is the "time to the occurrence of a particular event." Survival models may be applied to a variety of fields such as biology, medicine, engineering, and economics. An example of an application in engineering is to model the time it takes for a ball-bearing to wear. The focus will be on applications in biology and medicine where the event of interest may be time to death or time to a particular event such as relapse of a disease. The standard statistical techniques for data analysis are usually not applicable to survival data. First of all survival data are typically not symmetric. A histogram of survival times will indicate that they tend to be positively skewed. As a result it is not reasonable to assume data of this type to be normally distributed. Another feature of survival data that makes it difficult to use standard techniques is that

survival times are frequently "censored." The survival time of an individual is said to be censored when the endpoint of interest has not been observed.

Right censoring, which is the most common form, occurs when the exact survival time is not known. All that is known is that the exact survival time exceeds the recorded value. This type of situation can occur if the subjects do not experience the event of interest when the study terminates or they are lost to follow-up. Such data cannot be analyzed by ignoring the censored observations because in general those who tend to live longer are more likely to be censored.

Another feature of survival data is the potential for truncation. For left truncation only subjects that experience a certain intermediate event are made known to the investigator. For example, if the focus of the study is to look at relapse of leukemia prior to death, left truncation occurs because only those who experience the intermediate event (relapse) are observed.

There are both parametric and nonparametric techniques available to model survival data. The parametric methods of estimation assume that the probability density function of the time to a particular event follows a specific distribution, such as the exponential distribution, while the nonparametric methods do not. The three major statistical packages (SAS, SPSS, and BMDP) are compared for both parametric and nonparametric survival analysis methods. Recommendations are given as to when each package is superior under both standard and nonstandard conditions. Several datasets are analyzed by each of the packages so that direct comparisons can be made. These include: (1) ovarian cancer data, Edmunson *et al.* (1979); (2) the Stanford heart transplant data, Crowley and Hu (1977); (3) larynx cancer data, Kardaun (1983); (4) breast feeding data, National Labor Survey of Youth (NLSY); and (5) melanoma data, Lee (1992).

2. COMPARISON OF THE PACKAGES

There are a number of similarities between SAS(version 6.09), SPSS (version 5.0), and BMDP (version 1990) in terms of computational methods for survival analysis. For the most part, the three packages agree with one another with respect to parameter estimation and calculation of available statistical tests. Table 1 lists the procedures that are found in each of the three statistical packages that perform the major survival analysis techniques: Kaplan-Meier method (Kaplan and Meier, 1958), life table methods (Gehan, 1969), Cox proportional hazards models (Cox, 1972), and the accelerated failure time model (Andersen, Borgan, Gill, Keiding, 1993). The life table method is not considered in this discussion since it is no longer commonly used in medical applications. Each of the packages can handle right censored data easily. The major differences among the packages are summarized in Table 2.

2.1 Kaplan-Meier Estimates and Tests

The Kaplan-Meier estimates of the survival function are available in all three packages along with standard errors of the survival function calculated by Greenwood's formula (Greenwood, 1926). The three packages provide the results of the log-rank (Collett, 1994) and the Wilcoxon tests (Gehan, 1969) for comparing the survival of two or more groups. The Tarone-Ware test (Tarone and Ware, 1977) is available in SPSS and BMDP but not in SAS. The Peto-Prentice (Peto and Peto, 1972) test is available only in BMDP. SPSS has the ability to calculate all pairwise comparisons among the groups by issuing a single command (/compare=pairwise). More work is needed in SAS and BMDP to obtain the pairwise results.

SAS conflicts with SPSS and BMDP in the way the mean survival times are calculated. All three packages underestimate the mean when the last observation is censored, but SAS does so with a larger magnitude. SAS estimates the mean only up to the last event time, where both SPSS and BMDP estimate the mean up to the last observation. The packages agree in estimating the median survival time and its variance. The number at risk provided as standard output is incorrect for each of the packages. The risk set should be determined an instant before each of the event times. Therefore the numbers at risk should all be increased by one in

each of the package's output. Both SPSS and BMDP have built-in modules for the test for trend and to do stratified analyses (Klein and Moeschberger, 1996). SAS/IML can be used to obtain these tests in SAS by using the proper test statistics. The results agree among the three packages for stratified analyses and the test for trend.

2.2 Cox Proportional Hazards Model

Each of the packages fits the Cox proportional hazards model using the Newton-Raphson procedure to maximize the partial likelihood. The estimated regression parameters and their standard errors are provided as standard output. SAS's PROC PHREG has four likelihoods (Breslow, exact, discrete, and Efron) that adjust for ties in the survival times. The Breslow likelihood is due to Breslow (1974). The exact likelihood is due to Kalbfleisch and Prentice (1980). The discrete likelihood replaces the proportional hazards model by the discrete logistic model. The Efron likelihood is due to Efron (1977). Each of these can be used to adjust for tied survival times. They are equivalent when there are no ties. When there are a large number of tied survival times the exact and the discrete likelihoods are preferred. Both perform well but the discrete likelihood is computationally easier and faster in terms of computer time. When there are only a few ties both the Efron and Breslow likelihoods perform well. The Breslow likelihood is usually preferred to Efron's since it is quite straightforward to compute. The SAS default is the Breslow likelihood. This is the only likelihood that is available in SPSS and BMDP.

A nice feature of each of the packages is the ability to construct 2 df tests. SAS uses a "test" statement to obtain the results of the Wald test of linear hypotheses about the regression parameters. SPSS uses its "/categorical" and "/contrast" subcommands to obtain 2 df Wald tests. In BMDP, 2 df Wald, likelihood ratio, and score tests are available in the "/TEST" paragraph using the "ELIM" statement.

Each package also allows the output of other useful information such as the estimated survival and hazard functions, the martingale residuals, and the number of subjects at risk for inspection or graphics. Each package can account for time-dependent covariates and they can be incorporated to test the proportional hazards assumption. A graphical method is also available in each package to test the

proportional hazards assumption. SAS, SPSS, and BMDP each have procedures to do a stratified analysis. The results agree among the packages.

SPSS has a feature (/missing=include) that includes cases in the analysis that have *missing values* for covariates that are specified in the model. SAS and BMDP automatically exclude such cases from the analysis. The results of using the “/missing=include” subcommand are peculiar. What SPSS does is actually use the value of the covariate even though it is defined as missing. For example, a covariate that has -9 as its defined missing value is used in the analysis as a -9. We recommend not using this option. Use the default option “/missing=exclude”. These results agree with those of SAS and BMDP using Breslow’s likelihood.

2.3 Left Truncated Data

Left truncated data frequently arises in the analysis of medical data. If the truncation is ignored your results may be severely biased (Klein and Zhang, 1996).

Only BMDP has a built-in routine to handle left truncation. BMDP uses a time-dependent strata to modify the risk sets to adjust for left truncation. It uses the “PASS” and “LEVEL” commands together to modify the risk sets. We have discovered a method that allows analysis of left truncated data in SAS. An augmented data set is constructed where at each death time a record is added for each individual who has entered the study at an earlier time. The censoring variable is set to zero and a variable named “strata” which is equal to the death time is included. A stratified Cox model is used stratifying on the constructed variable “strata”. Both the SAS and BMDP code for left truncated data follows:

```
SAS code using augmented data:
proc phreg;
  strata strata;
  model time*censor(0)=group;
run;
```

Original Data:

Time	Censor	Group	Enter
10	1	1	8
12	1	0	6
15	1	1	11
20	1	0	14

Augmented Data:

Time	Censor	Group	Strata
10	1	1	10
12	0	0	10
12	1	0	12
15	0	1	12
15	1	1	15
20	0	0	15
20	1	0	20

BMDP code using original data:

```
/REGRESS
  ADD = treat.
  AUX = group, enter.
  PASS = 2.
/TEST
  ELIM = treat.
  STAT = wald, lratio, score.
/FUNCTION
  LEVEL = 1.
  if (enter gt time) then LEVEL=0.
  treat=group.
/END
```

The results of running a stratified Cox model on the “augmented data” are identical to the results obtained using the “original data” in BMDP’s 2L. The method is not very practical at this time. If the augmented data set can be constructed in SAS/IML the method will be very useful to the SAS user.

2.4 Model Building

There are automated variable selection routines (forward, backward, and stepwise) available in each of the statistical packages for model building. The routines are very similar among the packages and tend to give the same results in terms of a final model. A nice feature of SPSS is the ability to use a 2 df Wald test to determine if a block of levels of a covariate should enter the model. This feature is superior to those in SAS and BMDP where only 1 df tests are available during the model-building process.

2.5 Nonstandard Conditions

There are several nonstandard conditions that may arise when modelling survival data. Two of the most commonly encountered ones are: 1) all the events times of one level of a covariate occur before the first event time of another level and 2) one covariate is a

linear combination of another covariate. Under the first condition, the estimate of the regression parameter is \pm infinity. The second condition leads to a singular matrix that is not invertible and therefore the regression parameters cannot be estimated. The results of how each package handles the nonstandard conditions differs among the packages. Both SPSS and BMDP give warning messages that one of the nonstandard conditions is present. SAS provides results without providing any information that a nonstandard condition is present. For example, when condition 1 is present, SAS provides the results of the estimated regression parameters after 15 iterations. These are not the correct estimates because the parameter estimate for the covariate with the condition present is diverging. This can be seen by using the “itprint” option, but no warning message is given. BMDP provides references that provide information on how to remedy the nonstandard conditions.

2.6 The Accelerated Failure Time Model

Only SAS and BMDP allow the use of the accelerated failure time model. There is an agreement in the results. SAS allows the use of the generalized gamma distribution. It can be useful in choosing which underlying distribution to use to model the data. One must be careful in interpreting the results of both packages, because the estimates that are provided are for the transformed logarithm of survival time. You will want to transform the estimates back to their original units. This can be done using the delta method.

3. CONCLUSION

SAS, SPSS, and BMDP are all very good packages for analyzing survival analysis applications. Each package has its advantages and pitfalls. There is not one identifiable superior package. Under an ideal situation, a researcher would want to have all three packages available. The choice of which package to buy is determined by the nature of analyses that will be done. For the ordinary researcher any of the packages would be sufficient.

SAS and SPSS both have matrix languages that allow the implementation of applications and tests that are not a standard part of the package. Even though BMDP does not have a built-in matrix language, it has incorporated techniques geared

toward biomedical applications such as left truncation.

ACKNOWLEDGMENTS

Research supported by contract 5 R01 CA54706-05 from the National Cancer Institute and the U.S. Army Medical Research and Development Command.

BIBLIOGRAPHY

Andersen, P.K., Borgan Ø., Gill, R.D., and Keiding N. (1993). *Statistical Models Based on Counting Processes*. Springer-Verlag, New York.

BMDP Statistical Software Manual, Volume 2 (1992). University of California Press, California.

Breslow, N.E. (1972). Contribution to the discussion of a paper by D.R. Cox. *Journal of the Royal Statistical Society, B*, **34**, 216-7.

Collett, D. (1994). *Modelling Survival Data in Medical Research*. Chapman and Hall, London.

Cox, D.R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society*, **74**, 187-220.

Cox, D.R. (1975). Partial likelihood. *Biometrika*, **62**, 269-76.

Crowley, J. and Hu, M. (1977). Covariance analysis of heart transplant survival data. *Journal of the American Statistical Association*, **72**, 27-36.

Edmunson, J.H., Fleming, T.R., Decker, G.D., Malkasan, G.D., Jorgensen, J.A., Jefferies, J.A., Webb, M.J., Kvols, L.K. (1979). Different chemotherapeutic sensitivities and host factors affecting prognosis in advanced ovarian carcinoma versus minimal residual disease. *Cancer Treatment Reports*, **63**, 241-7.

Efron, B. (1977). The efficiency of Cox's likelihood function for censored data. *Journal of the American Statistical Association*, **72**, 557-65.

Gehan, E.A., (1969). Estimating survival functions for the life table. *Journal of Chronic Diseases*, **21**, 629-44.

Greenwood, M. (1926). The errors of sampling of the survivorship tables. *Reports on Public Health and Statistical Subjects*, number 33, Appendix 1, HMSO, London.

Lee, E.T. (1992). *Statistical Models and Methods for Lifetime Data*, Wiley, New York.

Kalbfleisch, J.D., and Prentice, R.L. (1980) *The Statistical Analysis of Failure Time Data*, Wiley, New York.

Kaplan E.L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, **53**, 457-81.

Kardaun, O. (1983). Statistical survival analysis of male larynx-cancer patients -- A case study. *Statistica Neerlandica*, **37**, 103-25.

Klein, J.P. and Zhang, M.J. (1996). Statistical challenges in comparing chemotherapy and bone marrow transplantation as a treatment for leukemia, *Lifetime Data: Models in Reliability and Survival Analysis*, N.P. Jewel, 175-85.

Klein, J.P. and Moeschberger M.L. (1996) *Survival Analysis*. Springer-Verlag, New York (in press).

Peto, R. And Peto, J. (1972) Asymptotically efficient rank invariant procedures. *Journal of the Royal Statistical Society, A*, **135**, 185-207.

SAS/STAT User's Guide, Version 6, Volume 1 (1990), SAS Institute, Inc., Cary, NC.

SAS/STAT User's Guide, Version 6, Volume 2 (1990), SAS Institute, Inc., Cary, NC.

SPSS for Unix, Advanced Statistics, Release 5 (1993), SPSS, Inc.

SPSS Statistical Algorithms, 2nd Edition (1993), SPSS, Inc.

Tarone, R.E. and Ware, J, (1977). On distribution-free tests for equality of survival distribution. *Biometrika*, **64**, 156-60.

Table 1: Listing of the procedures by survival analysis topic and statistical package.

Survival Analysis Topic	SAS	BMDP	SPSS
Kaplan-Meier Method	PROC LIFETEST	1L	KM
Life Tables Method	PROC LIFETEST	1L	SURVIVAL
Cox Proportional Hazards Model	PROC PHREG	2L	COXREG
Accelerated Failure Time Model	PROC LIFEREG	2L	N/A*

* Not Available

Table 2. Survival analysis applications available in SAS, SPSS, and BMDP with recommendations and comments.

Application	Packages Available	Agree in results	Recommended Package(s)	Comments
Kaplan-Meier estimates	SAS, SPSS, BMDP	YES	N/D**	Mean survival time is underestimated with a greater degree in SAS
Testing equality of strata	SAS, SPSS, BMDP	YES	SPSS	SPSS : Tarone-Ware Test and pairwise comparisons
Test for trend	SAS*, SPSS, BMDP	YES	SPSS, BMDP	Not built into SAS
Stratified tests	SAS*, SPSS, BMDP	YES	SPSS, BMDP	Not built into SAS
Estimation in the Cox model	SAS, SPSS, BMDP	YES	N/D	Each use maximum partial likelihood estimation
Likelihoods for ties	SAS	N/A	SAS	SAS has four likelihoods
Test of regression parameters	SAS, SPSS, BMDP	YES	BMDP	BMDP has Wald, score, and likelihood ratio tests for 2 df tests
Time-dependent covariates	SAS, SPSS, BMDP	YES	SAS	Syntax easier in SAS
Left truncation	SAS, SPSS, BMDP	YES	BMDP	Built into BMDP
Model building techniques	SAS, SPSS, BMDP	YES	SPSS	SPSS: 2 df Wald tests during selection process
Testing proportional hazards	SAS, SPSS, BMDP	YES	N/D	2 methods available in each package
Stratified analysis	SAS, SPSS, BMDP	YES	N/D	Easy to implement in each package
Nonstandard conditions	N/A***	NO	BMDP	BMDP: warnings and references
Parametric regression	SAS, BMDP	YES	SAS	SAS includes generalized gamma distribution

* Available in the package with additional programming (SAS/IML) or calculation.

** N/D = Not Distinguishable

*** N/A = Not Applicable