197
6-23-81
JWA

LA-8834-SR

Status Report

B-5188

1

DR-2746

MASTER

**Applications of
Cox's Proportional Hazards Model to
Light Water Reactor Component Failure Data**

University of California

LOS ALAMOS SCIENTIFIC LABORATORY

Post Office Box 1663 Los Alamos, New Mexico 87545

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

APPLICATIONS OF COX'S PROPORTIONAL HAZARDS MODEL TO LIGHT WATER REACTOR COMPONENT FAILURE DATA

by

Jane Booker, Kathy Campbell, Aaron G. Goldman,
Mark E. Johnson, and Maurice C. Bryson

ABSTRACT

The use of Cox's proportional hazards model in analyzing light water reactor failure data is described. A small data base on reactor valve failures illustrates the method.

I. INTRODUCTION

Davis et al. (1979) proposed the use of regression techniques in the analysis of failure data on components in light water reactors (LWR). A key feature of their approach is that it pools all the available data in the analysis so that failure rates corresponding to components with very few failures can nevertheless be estimated. Traditional failure rate estimation methods, on the other hand, which are based upon analyzing subsets of the data separately, suffer in these applications. During the past year the statistics group at the Los Alamos National Laboratory (under contract IC0-13-3093) has continued to investigate the use of regression methods in analyzing LWR component failure data. The purpose of this report is to summarize the results of the past year's efforts and to explain some of their specific implications to LWR failure rate estimation problems.

By regression methods we are really referring to Cox's proportional hazards model (Cox 1972, Kalbfleisch and Prentice 1980). The Cox model has been widely used in medical applications, in particular, clinical trials and heart transplant studies (Crowley and Hu 1977). Davis et al. (1979) recognized its potential in the physical sciences and in application to operating data from reactors. Since the dominant application area of the Cox model has previously been biomedical, certain questions of particular interest to LWR component failure analyses have not been addressed. In particular, due to very low failure rates, there is extensive censoring in LWR data sets. That is, at the time at which the statistical analysis is to be performed, it is common that greater than 90% of the components have not failed. Moreover, certain valves under particular operating conditions may have experienced no failures whatsoever. An additional characteristic of these data includes correlated exogenous variables (covariates). This means, for example, that the covariate radiation level and operating temperature could be positively related. Theoretical developments (Kalbfleisch 1974) are generally restricted to orthogonal covariates.

In Sec. II, we describe a small data base consisting of failure data on valves in two reactors at the Westinghouse Zion I and Zion II power plants. This discussion will serve to illustrate problems encountered in applying the Cox model as contrasted to traditional engineering approaches in estimating failure rates. In Sec. III, we discuss the results presented in the two reports of Bryson et al. (1980) and Bryson and Johnson (1981). These efforts were designed to address some theoretical problems in the Cox model of particular interest to LWR applications. The results tend to confirm the efficacy of using the Cox model in these situations.

II. PRELIMINARY ANALYSIS AND SCREENING OF THE ZION I, II DATA BASE

In this section, some traditional analyses are discussed in reference to the Westinghouse PWR Zion I and II reactor valve failures. The data base contains information on 12 covariates that describe various valve characteristics such as composition, function, type, operation, and external and internal environments for 4 774 valves with 78 failures. This represents over 98% censoring. Such heavy censoring causes problems in many traditional

analyses of failure rates as mentioned in Sec. I. In addition, each covariate has from 2 to 26 levels (for example, there are 2 different plants, 26 systems, and 12 functions). In such a finely structured data base with so few failures, there are many covariate combinations (cells) that do not occur and many more that occur with no failures. This makes many comparisons of failures in these cells either imprecise or impossible.

One solution is to reduce the number of levels for covariates with a large number of levels by aggregating the data. For example, the 12 valve functions can be reduced to 3 more general functions: static, relief, and flow. Data aggregations were suggested by Bob Burns, Nuclear Engineer, Los Alamos National Laboratory Group Q-7. The aggregations can help prevent cells with no failures but can pose other difficulties, such as the loss of information and the undesirable result of combining covariate levels that have different failure rates. The latter problem occurred frequently in the Zion data base. Therefore, analysis was attempted on nonaggregated data wherever possible.

A. Data Base Screening

Prior to any statistical testing, the Zion data base was edited for obvious typographical and logical errors. Attempts were made to recover missing values and to identify nonsensical values. Preliminary analysis of the data set included checks for suspicious failure and censor times. This revealed, for example, the presence of one negative failure time!

With the edited data, preliminary tests for covariate dependencies, differing failure rates, and hazard functions within covariates were performed.

B. Covariate Dependence

With 12 covariates in the data base, it is useful to examine how they relate to one another. This was accomplished using some nonparametric procedures (Conover 1971) to test failure rates and ranks of failure rates for two or three covariates. Even with the aggregated data base, many covariate comparisons were still not possible with the prevalence of empty cells and nonexistence of certain cells.

The results of tests indicate the following joint effects of covariates on the failure rate: valve function dependent on valve system; operation dependent on function; inlet size independent of function, operator, and

system; system independent of operator; and plant dependent on inlet size (see Table I). These results are based on a level of significance of 1% and utilized the aggregated data base. (The expression "level of significance of 1%" means that there is a 1% chance that these results are in error.)

C. Likelihood Ratio Tests

To test for the equality of failure rates ($\lambda_1 = \lambda_2 = \dots = \lambda_k$) for levels of a given covariate, the maximum likelihood ratio test statistic was used on several covariates for both aggregated and nonaggregated data sets. Here again, empty cells pose a problem in the analyses because zero failure rates are compared to positive failure rates.

For those tests that are possible, the results include: active systems failure rate differs from standby systems; Zion I failure rate differs from Zion II (see Table II); 4 aggregated types of valve operators have differing failure rates; all 6 inlet sizes have differing failure rates; 3 aggregated types of valve functions have differing failure rates; all 12 valve functions

TABLE I
DEPENDENCE OF PLANT ON INLET SIZE

(a) Observed Values

	Failures				Successes		
	<u>0-2"</u>	<u>2-12"</u>	<u>> 12"</u>		<u>0-2"</u>	<u>2-12"</u>	<u>> 12"</u>
Zion I	9	32	5	Zion I	1166	552	64
Zion II	7	21	3	Zion II	1376	755	83

(b) Expected Values (if plant and size are independent)

	Failures				Successes		
	<u>0-2"</u>	<u>2-12"</u>	<u>> 12"</u>		<u>0-2"</u>	<u>2-12"</u>	<u>> 12"</u>
Zion I	21.7	11.5	1.3	Zion I	1126.4	598.8	68.2
Zion II	26.6	14.2	1.6	Zion II	1383.3	735.4	83.8

(c) CONCLUSION: The observed differs from the expected at < 1% significance level with $\chi^2_2 = 78.9$.

TABLE II
SUMMARY OF ZION I, ZION II COMPARISONS

Failure Rates (failures/h)

Zion I:	0.54009×10^{-6}
Zion II:	1.21701×10^{-6}

Likelihood Ratio: 12.69 (df = 1)

Significant at 1% level

Log Rank Statistic: 7.73 (df = 1)

Significant at 1% level

have differing failure rates; 3 aggregated valve inlet sizes have differing failure rates.

By sequentially deleting a level at a time in each of the above, the test results can and do change. This sensitivity illustrates the importance of having complete data on all levels to assure accurate test results.

D. Log Rank Tests

Kalbfleisch and Prentice (1980) suggest testing for equality of hazard functions (see Sec. III.A.) for various levels of a covariate. This procedure actually parallels those in Sec. II.B. with the modification that the comparison is made between failures and survivors in the same covariate.

Briefly, the results indicate hazard functions differed with respect to the 2 plants Zion I and II (see bottom of Table II and Fig. 1), the 3 function categories, all 12 functions, all 3 body material types, all 8 operators, all 6 inlet sizes, all 18 systems, all 11 valve types, 4 aggregates of operators, and active and standby systems.

It is of interest to note that log rank test results on the data base in Davis, Campbell, and Schrader (1979) correspond to their Cox regression model results. Also, the above results support those in Sec. II.C.

Fig. 1.
Comparison of Zion I and Zion II survivor functions.

E. Conclusion

Two problems encountered in the Zion data base make many traditional analyses imprecise and difficult. These problems are (1) lack of failures in many covariate combinations or cells (for example, no failures in passive system valves with flow functions) and (2) lack of occurrence of some covariate combinations (for example, no passive system valves have inlet sizes larger than 24 in.). Even with combining or aggregating levels of a covariate, the above problems do not totally disappear, and, indeed, other difficulties may arise from the aggregation process. The analysis of the Zion data base, therefore, requires an approach that can overcome or at least account for the above limitations. Regression methods are clearly appropriate for these problems. Moreover, among the regression methods, the Cox model can be used without making the stringent assumption that the failure rates are constant.

III. CONSIDERATIONS IN APPLYING THE COX MODEL IN LWR APPLICATIONS

Two reports were written this past year to address questions about the Cox model that are particularly relevant to failure time analysis of LWR components. Bryson, Johnson, and Goldman (1980) performed a Monte Carlo simulation study to investigate the performance of estimators under situations similar to those encountered in our applications. A second report by Bryson and Johnson (1981) provides guidelines to the (unwary) analyst who attempts to execute Cox model computer packages. This work has implications as well to the modeling process. Incorporating too many covariates can lead to absurd results due to an overfit of the data. However, stratification methods can be used to obviate some of these problems.

A. Simulation Study

Much of the work since Cox's (1972) landmark paper has been concerned with "large-sample" (asymptotic) behavior in the proportional hazards model. These asymptotic results are useful for constructing confidence intervals, testing hypotheses, and assessing goodness of fit. The immediate question pronounced by practitioners is "how large is large enough?" A simulation study was performed to determine sample sizes for which asymptotic results appear to be valid.

The Cox model assumes a hazard function $\lambda(t) = \lambda_0(t) \cdot \exp(\underline{\beta} \underline{Z})$, where $\underline{Z}$ is a vector of covariates, $\underline{\beta}$ is a vector of regression parameters, and $\lambda_0(t)$ is the "nominal" failure rate. The function $\lambda_0(t)$ can be viewed as the failure rate with no covariate effects ($\underline{Z} = \underline{0}$). Two LWR components with different covariate values will be assumed to have proportional hazard rates. The first step in applying the Cox model is to estimate the vector $\underline{\beta}$ of unknown regression parameters. Our attention has thus far been primarily focused on the estimation of $\underline{\beta}$, because subsequent estimation of $\lambda_0(t)$ is fairly straightforward (Kalbfleisch and Prentice 1980). A summary of the key results and their implications to LWR problems will now be given.

1. Sample-Size Effects with Complete Data. For this problem we considered a two-covariate problem with frequencies $P[Z_1 = 0] = P[Z_1 = 1] = 0.5$

and $P[Z_2 = 0] = P[Z_2 = 0.25] = P[Z_2 = 0.50] = P[Z_2 = 0.75] = P[Z_2 = 1.] = 0.2$. In other words, we consider a 2×5 factorial structure in the covariates. Random samples of size 20, 40, 60, 80, and 100 were simulated from an exponential distribution with failure rate $\exp(\beta_1 Z_1 + \beta_2 Z_2)$ with the "true" values of the regression parameters $\beta_1 = \beta_2 = 1$. For samples of size 40 or more, we found that the sample variances of $\hat{\beta}_i$ were close to the asymptotic variance, and for all cases the estimates were within Monte Carlo sampling variability of β_i . Inferences regarding $\underline{\beta}$ (for example testing the hypothesis that $\underline{\beta} = \underline{0}$ versus $\underline{\beta} \neq \underline{0}$ --that is, testing for a covariate effect) could be accurately performed for all the sample sizes tested.

2. Effects of Censoring. We considered simulated samples of size 60 with censoring mechanisms yielding average censoring rates of 17, 56, 73, and 11% (these rates are determined by the censoring mechanism.) Again, we found that the estimates of β_1 and β_2 appeared to be unbiased and the sample variances of $\hat{\beta}_i$ agreed closely with the large sample values. The statistic based on the chi-square distribution for testing for covariate effects performed correctly. Hence, accurate confidence intervals for β_1 , β_2 , and jointly β_1 and β_2 can be constructed.

3. Effects of Covariate Structure. In practice, it is unlikely that the covariate structure will be that given in item 1. In fact, design constraints will frequently introduce imbalance and correlation among exogenous variables. We considered a variety of covariate distributions, including correlated covariates and unequal variances for the two covariates. No substantive effects were observed concerning estimator bias or the suitability of the chi-square test distribution.

In summary, each of the three topics examined yielded a similar picture, namely that the small sample performance of the regression estimators performed reasonably well over a variety of covariate and censoring situations. These limited results do not provide any evidence to dispel the use of the Cox model in LWR applications. We are currently in the process of refining this study in view of the results obtained in the Bryson and Johnson (1981) report, which will now be described.

B. Anomalies in the Cox Model and Their Resolution

The past few years have seen the development of computer programs to facilitate failure rate estimation with the Cox model (e.g., Kalbfleisch and Prentice, 1980, provide computer codes). The Kalbfleisch-Prentice programs do not check for anomalous situations which could lead to fatal errors. This problem can be illustrated with a simple example. Suppose a data set consists of the following failure times and corresponding covariates

<u>Failure Time t_i</u>	<u>Z_{1i}</u>	<u>Z_{2i}</u>
405	-1	1
427	-1	1
456	1	1
470	1	1
489	-1	-1
521	-1	-1
523	1	1
546	-1	-1

Assume further that all remaining components have not yet failed (that is, they are censored) and there are no other components with covariate values $Z_{1i} = -1$ and $Z_{2i} = 1$. The Kalbfleisch-Prentice code will result in the message "NEWTON-RAPHSON DID NOT CONVERGE IN 10 ITERATIONS." Bryson and Johnson (1981) showed that for this example, $\hat{\beta}_1 = -\infty$ and $\hat{\beta}_2 = +\infty$. Hence, nonconvergence is expected. Note that these covariate values are commonly used to model exogenous factors. For example, we might have $Z_{1i} = 1$ if component i is in a pressurized water reactor and $Z_{1i} = -1$ if component i is in a boiling water reactor; $Z_{2i} = 1$ if component i is a motor-operated valve, and $Z_{2i} = -1$ otherwise. Key questions here are how often such anomalies occur and what should be done about them. Bryson and Johnson (1981) give a table of expected fraction of infinite estimates assuming an exponential model. A general statement about corrective action is that if $\hat{\beta}_i = \pm\infty$, it is inappropriate to pool the failure data for this covariate. Bryson and Johnson (1981) suggest the use of stratification to handle the problem. They also point out that perhaps the wrong approach is to ignore troublesome covariates. Examples are given to highlight the problem with this naive method.

The clue that an anomaly exists is the nonconvergence of the Newton-Raphson solution to the likelihood equations. When such nonconvergence occurs, the analyst should take the following steps.

1. List the covariate values corresponding to the ordered failure times.
2. Identify those covariates whose values are monotone.
3. If there are no monotone covariates, identify any subsets of covariates whose sums or differences are monotone. (In the example Z_{1i} , Z_{2i} is monotone.)
4. Identify the appropriate regression coefficient estimates as $+\infty$ or $-\infty$, depending on the direction of monotonicity.
5. Estimate other coefficients according to the stratification procedure outlined by Bryson and Johnson (1981). If convergence problems persist, return to step 3.

IV. DISCUSSION

This report was intended to summarize the results obtained during the past year in applying the Cox model in LWR applications. The two recent reports of Bryson and Johnson (1981) and Bryson et al. (1980) were written more for mathematical statisticians than for an applied audience. The mathematical details were not repeated here, as these reports are available from the authors.

ACKNOWLEDGEMENTS

The authors are grateful for many helpful comments from Art du Charme and Bob Easterling who are affiliated with Sandia National Laboratories.

REFERENCES

Maurice C. Bryson, Mark E. Johnson, and Aaron S. Goldman, "A Simulation-Based Analysis of Cox's Proportional Hazard Rate Model," Los Alamos National Laboratory, unpublished data, 1980 (LA-UR-80-830).

Maurice C. Bryson and Mark E. Johnson, "The Incidence of Monotone Likelihood in the Cox Model," submitted to Technometrics, 1981, Los Alamos National Laboratory (LA-UR-80-2710).

W. J. Conover, Practical Nonparametric Statistics (Wiley, New York, 1971).

David R. Cox, "Regression Models and Life Tables (With Discussion)," *Journal of the Royal Statistical Society, Series B*, 34, 187-220 (1972).

John Crowley and Marie Hu, "Covariance Analysis of Heart Transplant Survival Data," *Journal of the American Statistical Association*, 72, 27-36 (1977).

Herbert T. Davis, Katherine Campbell, and Ronald M. Schrader, "Improving the Analysis of LWR Component Failure Data," Los Alamos National Laboratory, unpublished data, 1979 (LA-UR-80-92).

John D. Kalbfleisch, "Some Efficiency Calculations for Survival Distributions," *Biometrika* 61, 31-38 (1974).

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