
PARTITIONING THE EFFECTS OF CONCOMITANT VARIABLES
IN SURVIVAL ANALYSIS

by

H. Dennis Talley and Kerry L. Lee

Duke University Medical Center
Durham, North Carolina

and

Department of Biostatistics
University of North Carolina at Chapel Hill
Institute of Statistics Mimeo Series No. 1048

December 1975

PARTITIONING THE EFFECTS OF CONCOMITANT VARIABLES
IN SURVIVAL ANALYSIS

by

H. Dennis Tolley and Kerry L. Lee
Duke University Medical Center,
Durham, North Carolina

Summary

Using the Kalbfleisch-Prentice marginal likelihood based on Cox's regression and life model, a computationally simple method is presented for testing general linear hypotheses of the form $H_0: C\beta = 0$. The method, which does not require estimating β , is based on partitioning the asymptotic global chi-square statistic of Cox to produce the appropriate test statistic. Under H_0 , this statistic also has an asymptotic chi-square distribution. An examination of its performance in small and moderate samples reveals the method to be useful, particularly for testing the significance of subsets of the observed concomitant variables.

Some key words: survival analysis, concomitant variables, hazard functions, regression, significance tests, linear hypotheses, censoring.

1. INTRODUCTION

The problem of comparing failure times of individuals or objects which have been subjected to different experimental conditions is one of practical importance. For example in medical trials, researchers frequently seek to determine the best of several treatments by comparing differences in corresponding mortality experience. Methods for the analysis of such data, which may contain censored as well as tied observations, have prominently appeared in statistical literature over the past several years. Representative papers which discuss the analysis of life table data under varying conditions are Mantel (1966), Chiang (1968, chapt. 10) and Koch et al. (1972). These and numerous other papers now provide results for evaluating differences in survival of several population

Often in practice there are concomitant or explanatory variables observed on each individual in addition to the actual time of failure t . Methods suggested for adjusting the observed survival times to account for concomitant measurements have typically involved modelling the hazard function and the survivor distribution to reflect the form of the dependence on the covariates. A particularly attractive formulation incorporating explanatory variables in a regression-like manner was proposed by Cox (1972), and subsequently investigated and applied by several authors (cf. for example Kalbfleisch and Prentice (1973), Kalbfleisch (1974), and Holt and Prentice (1974).) The model is specified by the hazard function

$$\lambda(t; \underline{z}) = \lambda_0(t) \exp(\underline{z}\beta), \quad (1)$$

where \underline{z} is a set of r concomitant variables, and $\underline{\beta}$ is an $r \times 1$ vector of unknown regression coefficients. To examine the significance of the vector of parameters $\underline{\beta}$, Cox (1972) has provided a relatively simple procedure for testing the global null hypothesis, $\underline{\beta} = \underline{0}$. A significant result from this test will likely lead an investigator to inquire which covariables are actually responsible for affecting survival in a significant way. Cox suggested that significance tests about subsets of the parameters could be derived, for example, by performing the parameter estimation and comparing the maximum of log likelihoods achieved. Alternatively one can estimate $\underline{\beta}$ and the inverse of its covariance matrix, and use the asymptotic normality of MLE's to test various hypotheses of interest. In either case, iterative calculations are required to first estimate $\underline{\beta}$ before any tests can be performed.

Using the proportional hazard model (1), this paper presents a computationally simple method for assessing the significance of subsets of the observed concomitant variables. The procedure does not require formally estimating the vector $\underline{\beta}$, for the test statistic is derived by partitioning the asymptotic global chi-square statistic of Cox (1972) into statistics, one of which can be used for testing general hypotheses of the form $H_0: \underline{C}\underline{\beta} = \underline{0}$. Under H_0 , this simple statistic has an asymptotic chi-square distribution. Since the method assumes a particular form of the alternative hypothesis, its performance characteristics are investigated under a specific failure model with multiple populations. An examination of its properties in both censored and uncensored data reveals the method to be useful even in small and moderate sample experimental situations.

2. PRELIMINARY BACKGROUND AND NOTATION

Consider an experiment in which the failure time, or censoring time is observed for each of n individuals. Suppose in addition that a set of r concomitant variables $\underline{z} = (z_1, z_2, \dots, z_r)$ is observed on each individual. These may be regression-type variables such as age, dosage level, etc., or simply indicator variables. Following Cox (1972), assume expression (1) to be the hazard function relating failure time to the observed covariables. Let $R(t_i)$ be the set of labels attached to the individuals at risk just prior to time t_i , i.e., individuals who have neither failed nor been censored prior to time t_i . Cox (1972), modeled the likelihood of a failure at time t_i for an individual as conditional on the set of observed failure times and on the set of individuals whose failure time is at least t_i .

Some justification of Cox's result is given by Kalbfleisch and Prentice (1973) who obtained a marginal likelihood for $\underline{\beta}$ from the marginal distribution of the ranks of the survival times. Restricting themselves to the class of hazard functions (1) for which $\lambda_0(t)$ is not identically zero over an open interval, they showed that because of the invariance of this class under the group of strictly monotone increasing transformations on the time scale, the rank vector

$$\underline{r} = [(1), (2), \dots, (n)]$$

is marginally sufficient for the estimation of $\underline{\beta}$. In the case where no ties occur, the marginal likelihood is identical to Cox's result. In the presence of tied data, the marginal likelihood differs from Cox's corresponding result. We have used the Kalbfleisch-Prentice likelihood; however, the procedure developed is equally applicable to Cox's model for tied data.

The logarithm of the likelihood given in Kalbfleisch and Prentice (1973) for the general case where tied observations may be present is

$$L(\underline{\beta}) = \sum_{i=1}^k \underline{s}_{(i)} \underline{\beta} + \sum_{i=1}^k \log \left\{ \sum_{p \in P_i} \prod_{j=1}^{m_i} \left(\sum_{\ell \in R(t_{(i)}, p_j)} \exp(\underline{z}_{\ell} \underline{\beta}) \right)^{-1} \right\}, \quad (2)$$

where $t_{(i)}$, $i=1,2,\dots,k$ are the distinct failure points, m_i individuals fail at $t_{(i)}$, $\underline{s}_{(i)}$ is the sum of the covariate vectors associated with the m_i failures at $t_{(i)}$, $P_i = (p_1, \dots, p_{m_i})$ is the set of permutations of the labels of these failures, and $R(t_{(i)}, p_j)$ is the set difference $R(t_{(i)}, p_j) = R(t_{(i)}) - (p_1, \dots, p_{j-1})$. By direct

differentiation, one may form

$$U_{\xi}(\underline{\beta}) = \left\{ \frac{\partial L(\underline{\beta})}{\partial \beta_{\xi}} \right\} \quad (3)$$

and

$$I_{\xi\eta}(\underline{\beta}) = - \left\{ \frac{\partial^2 L(\underline{\beta})}{\partial \beta_{\xi} \partial \beta_{\eta}} \right\}, \quad (4)$$

from which maximum likelihood estimates of $\underline{\beta}$ can be iteratively computed. Under certain regularity conditions, $\underline{U}(\underline{Q})$ has asymptotically a multivariate normal distribution with mean vector zero and covariance matrix $\underline{I}(\underline{Q})$ (cf. for example Kendall and Stuart (1961), chapter 24). Thus if the hypothesis $H_0: \underline{\beta} = \underline{Q}$ is true, the statistic

$$Q = \underline{U}'(\underline{Q}) \underline{I}^{-1}(\underline{Q}) \underline{U}(\underline{Q}) \quad (5)$$

has an asymptotic chi-square distribution with r degrees of freedom. Cox (1972) gives the statistic (5) for large sample testing of the global hypothesis $\beta=0$. The expressions for $U(\beta)$ and $I(\beta)$ based on the Kalbfleisch-Prentice likelihood (2) are given in Appendix 1. Note that for many tied failure times the iterative solution of these equations for β can become computationally prohibitive.

3. SIGNIFICANCE TESTS IN THE PROPORTIONAL HAZARD MODEL

Consider now a hypothesis of the form

$$H_0: C\beta = 0, \quad (6)$$

where C is a specified uxr matrix (usr) of rank u . We shall show that a statistic which is relatively simple computationally can be used for testing H_0 .

Let

$$I^{-1}(0) = II', \quad (7)$$

where I is a non-singular rxr (triangular) matrix. We next define the rxr matrix

$$D = \begin{pmatrix} D_1 \\ D_2 \end{pmatrix}, \quad (8)$$

where $D_1 = CI$ and D_2 is an orthogonal complement of D_1 , of full rank, i.e., $D_1 D_2' = 0$, $D_2 D_1' = 0$. We next define $C_2 = D_2 I^{-1}$ and write

$$C^* = \begin{pmatrix} C \\ C_2 \end{pmatrix}. \quad (9)$$

We shall also use

$$C^{*-1} = (C^+, C_2^+), \quad (10)$$

where C^+ and C_2^+ are respectively Moore-Penrose generalized inverses of the matrices C and C_2 , (cf. for example Rao and Mitra (1971, chapt. 3)).

If we now set $\underline{\theta} = \underline{C}^* \underline{\beta}$, it follows that

$$\underline{U}^*(\underline{\theta}) = \left\{ \frac{\partial L(\underline{\theta})}{\partial \theta_{\xi}} \right\} = (\underline{C}^{*-1})' \underline{U}(\underline{\beta}), \quad (11)$$

and

$$\underline{I}^*(\underline{\theta}) = - \left\{ \frac{\partial^2 L(\underline{\theta})}{\partial \theta_{\xi} \partial \theta_{\eta}} \right\} = (\underline{C}^{*-1})' \underline{I}(\underline{\beta}) (\underline{C}^{*-1}) \quad (12)$$

It is easily shown that the quadratic form

$$Q = \underline{U}^{*'}(\underline{0}) \underline{I}^{*-1}(\underline{0}) \underline{U}^*(\underline{0}) \quad (13)$$

is precisely the same as that given in (5). Moreover, by substituting expression (12) for $\underline{I}^*(\underline{0})$, we may write (13) as

$$Q = \underline{U}^{*'}(\underline{0}) \begin{pmatrix} \underline{D}_1 \underline{D}_1' & \underline{0} \\ \underline{0} & \underline{D}_2 \underline{D}_2' \end{pmatrix} \underline{U}^*(\underline{0}) \quad (14)$$

$$= Q_1 + Q_2, \quad (15)$$

where
$$Q_1 = \underline{U}'(\underline{0}) \underline{C}^+ \underline{D}_1 \underline{D}_1' \underline{C}^{+'} \underline{U}(\underline{0}) \quad (16)$$

and
$$Q_2 = \underline{U}'(\underline{0}) \underline{C}_2^+ \underline{D}_2 \underline{D}_2' \underline{C}_2^{+'} \underline{U}(\underline{0}) \quad (17)$$

We have thus partitioned the global chi-square statistic into two orthogonal components, one of which (Q_1) is a direct measure of the source of variation due to $\underline{C}\underline{\beta}$. Under H_0 , Q_1 will have an asymptotic chi-square distribution with u degrees of freedom, and therefore can be used to test the hypothesis $\underline{C}\underline{\beta}=\underline{0}$. By appropriate substitution, expression (16) may be rewritten

$$Q_1 = \underline{U}'(\underline{\theta}) \underline{I}^{-1}(\underline{\theta}) \underline{C}' (\underline{C} \underline{I}^{-1}(\underline{\theta}) \underline{C}')^{-1} \underline{C} \underline{I}^{-1}(\underline{\theta}) \underline{U}(\underline{\theta}). \quad (18)$$

With this formulation, the significance of subsets of the observed concomitant variables can be assessed with only simple matrix calculations.

When H_0 does not hold, but local alternatives are considered, certain other properties follow. Recall that if a sequence \underline{t}_N of consistent asymptotically normal maximum likelihood estimators of $\underline{\theta}$ exists, where N is an index of sample size, then

$$-\sqrt{N} \underline{U}^*(\underline{\theta}) = \sqrt{N} \underline{I}^*(\underline{\theta}) (\underline{t}_N - \underline{\theta}) + o_p\left(\frac{1}{\sqrt{N}}\right) \quad (19)$$

is asymptotically $N(\underline{0}, \underline{I}_0(\underline{\theta}))$, where $\underline{I}_0(\underline{\theta}) = \lim_{N \rightarrow \infty} \frac{1}{\sqrt{N}} \cdot \underline{I}(\underline{\theta})$.

Following Stroud (1971), if the hypotheses are of the form

$$\begin{aligned} H_0: \underline{C} \underline{\beta} &= \underline{0} \\ H_A: \underline{C} \underline{\beta} &= \underline{S} / \sqrt{N} \end{aligned}$$

for some unspecified constant vector \underline{S} , then it is relatively easy to show from an expansion similar to (19) that Q_1 has asymptotically a non-central chi-square distribution with noncentrality parameter

$$\begin{aligned} \lambda &= 0, & \text{if } H_0 \text{ is true} \\ &= \underline{S}' \underline{D}_1 \underline{D}_1' \underline{S}, & \text{if } H_A \text{ is true} \end{aligned}$$

If the value of the statistic Q_1 is small and non-significant, one may choose to remove certain parameters from the hazard model (1). Such deletion would cause the likelihood equations to change; however, certain relationships are of interest. If we denote

$$U_{\xi}^{(2)}(\underline{\theta}_2) = \frac{\partial L^{(2)}(\underline{\theta}_2)}{\partial \theta_{2\xi}},$$

where the superscript (2) indicates the likelihood model after deleting $\underline{\theta}_1 = \underline{C}\beta$ from (1), then $U_{\xi}^{(2)}(0) = U_{\xi}^*(0)$, for all θ_{ξ} which remain in the model. Similarly it can be shown that $\left(\underline{I}_{\xi\eta}^{(2)}(0) \right)^{-1} = D_2 D_2'$. For the reduced model, the Cox statistic for testing the global null hypothesis $\underline{\theta}_2=0$ is simply

$$\begin{aligned} Q^{(2)} &= \underline{U}^{(2)'}(0) \underline{I}^{(2)-1}(0) \underline{U}^{(2)}(0) \\ &= Q_2. \end{aligned}$$

Therefore the statistic for testing $H_0: \underline{C}\beta = 0$ could be calculated by forming the difference between the global statistic for the initial model and the global statistic for the reduced model. In practice this is unnecessary, however, as expression (18) provides a straightforward formula for computing the desired statistic. However, the chi-square accounted for by deleting $\underline{C}\beta$ from the hazard model is equal to Q_1 , and a partitioning of the total global chi-square for various hypotheses of interest can be achieved. Since the procedure does not actually estimate β , but uses procedures similar to regression analysis by ranks, implementation in involved failure models where many concomitant variables are recorded is feasible.

4. PERFORMANCE OF THE TEST PROCEDURE

In order to assess the usefulness of this procedure in small and moderate samples, we have performed extensive numerical power calculations. Since the value of the test statistic depends only on the values of the covariate vector \underline{z} and the order in which the failures occur, some exact power calculations are possible. The probability of any particular permutation of the n distinct failure times is given by

$$\frac{\prod_{i=1}^n \exp(z_{p_i} \beta)}{\prod_{k=1}^n \left\{ \sum_{i=0}^{k-1} \exp(z_{p_{n-i}} \beta) \right\}}, \quad (20)$$

where the notation z_{p_i} indicates the vector of concomitant variables for the i th failure component of the particular permutation of the failure times. For each permutation, the value of the statistic can be calculated and weighted by the probability associated with that permutation. In this way the power can be evaluated for a given set of regression parameters β . If the z_{p_i} are distinct for all n individuals, the $n!$ possible permutations and their associated probabilities are impractical to compute except for rather small n . In order to examine power for moderate sample sizes, we have therefore conducted a number of sampling experiments. The basic comparisons involve model (1) with two components z_1 and z_2 in the covariate vector, and hence two parameters $\beta' = (\beta_1, \beta_2)$. The quantities z_1 and z_2 were chosen to be 0-1 indicator variables, giving four possible different values of \underline{z} , i.e., four populations. We focus on the hypothesis $C\beta = 0$, where C is $(0,1)$ or $(1,0)$, so

that the test is to decide whether a particular subset of the parameter vector has a significant effect in modelling the survival data. Since the power is invariant to $\lambda_0(t)$, we consider an exponential model where $\lambda_0(t) = \lambda$ is constant, and provide in Table 1 a representative comparison of exact power versus empirically simulated power for the test of $H_0: \beta_1 = \beta_2 = 0$.

Table 1 will be inserted here.

Results are given for the sample size $n=8$ and several different parameter combinations. In this particular case, the samples were comprised of two observations from each of the four populations. Since the test statistic has an asymptotic chi-square distribution with 1 df under H_0 , the values reported are based on the use of 3.84 as the critical value. The actual level of significance is 0.039. The uniformly distributed variates used for generating simulated exponential variates were obtained from a Tausworthe shift register generator for which an algorithm appears in Whittlesey (1968). All calculations were performed on a Zerox Sigma V computer at the Duke University Medical Center. Because of the close agreement between exact and simulated power, we report only simulation results for larger values of n .

Table 2 indicates the improvement in power as the sample size increases.

Table 2 will be inserted here.

With only a single observation from each of the four populations ($n = 4$), the power is zero for the parameter values considered. With so little data, this is not unexpected, however. The results improve substantially with the addition of only one more observation per population ($n = 8$). The values for $n = 48$ and $n = 96$, which are based on 12 and 24 observations respectively in

each population, reveal that for moderate and large samples the test is rather sensitive in detecting even relatively small deviations from the null hypothesis. To intuitively appreciate the magnitude of the difference in the underlying survival distribution when $\beta_2 = 0.5$ as compared with $\beta_2 = 0$, the density functions for the four populations are illustrated in Figure 1. In this case the test is

Figure 1 will be inserted here

actually whether the two upper curves are different and/or whether the two lower curves differ. Even though the differences are not very pronounced, the test distinguishes them reasonably well with less than 100 observations (power = 0.57 for $n = 96$). If the magnitude of the parameter β_1 is different from that shown, the power values reported also change slightly. If β_1 is small in absolute magnitude, the level of significance is slightly higher than when β_1 is large. This produces a corresponding change in the power. The values reported are representative, however, of the general relationships which prevail.

We have also examined the performance of the test procedure when some of the data are censored. The censoring model and sampling procedure used are identical to that employed by Gehan and Thomas (1969). Individuals are assumed to enter the study at a constant rate in the interval $(0, T)$, and then fail according to the exponential distribution. The value of T can be determined so that the expected proportion of censored observations for each sample is fixed. In order to examine the effect on performance of different censoring rates, we have considered fixed proportions of 25%, 50%, and 75%. These results are presented in Table 3. As expected, the power generally decreases as the censoring rate

Table 3 will be inserted here

increases. The size of the test remains rather constant, however.

5. CONCLUSION

Judging from the empirical results and our experience thus far with actual survival data, this is a useful method for linear hypothesis tests of the covariate parameter vector in the proportional hazards model. Because of its simplicity, it can easily be incorporated into survival analysis computer programs.

REFERENCES

- Chiang, C. L. (1968). Introduction to Stochastic Processes in Biostatistics. New York: Wiley.
- Cox, D. R. (1972). Regression models and life tables. J. R. Statist. Soc. B 34, 187-220.
- Gehan, E. A. and Thomas, D. G. (1969). The performance of some two-sample tests in small samples with and without censoring. Biometrika 56, 127-32.
- Holt, J. D. and Prentice, R. L. (1974). Survival studies in twin studies and matched pair experiments. Biometrika 61, 17-30.
- Kalbfleisch, J. D. (1974). Some efficiency calculations for survival distributions. Biometrika 61, 31-8.
- Kalbfleisch, J. D. and Prentice, R. L. (1973). Marginal likelihoods based on Cox's regression and life model. Biometrika 60, 267-78.
- Kendall, M. G. and Stuart, A. (1961). The Advanced Theory of Statistics. New York: Hafner.
- Koch, G. G., Johnson, W. D., and Tolley, H. D. (1972). A linear models approach to the analysis of survival and extent of disease in multidimensional contingency tables. J. Am. Statist. Assoc. 67, 783-96.
- Mantel, N. (1966). Evaluation of survival data and two new rank order statistics in its consideration. Cancer Chemotherapy Reports 50, 163-70.
- Rao, C. R. and Mitra, S. K. (1971). Generalized Inverse of Matrices and its Applications. New York: Wiley.
- Stroud, T. W. F. (1971). On obtaining large sample tests from asymptotically normal estimators. Ann. Math Statist. 42, 1412-24.
- Whittlesey, J. R. (1968). A comparison of the correlational behavior of random number generators for the IBM 360. Comm. of the ACM. 11, 641-44.

TABLE 1. Comparison of exact and simulated power for $H_0: \zeta\beta = \beta_2 = 0$. The sample size is $n=8$, and $\beta_1 = -2$. Simulation results are each based on 1000 samples of exponential failures.

β_2	Exact power	Simulated power
-4.00	0.939	0.94
-3.00	0.796	0.78
-2.00	0.481	0.48
-1.00	0.152	0.16
-0.50	0.066	0.06
-0.25	0.045	0.04
0.0	0.039	0.04
0.25	0.045	0.05
0.50	0.066	0.07
1.00	0.152	0.15
2.00	0.481	0.49
3.00	0.796	0.81
4.00	0.939	0.93

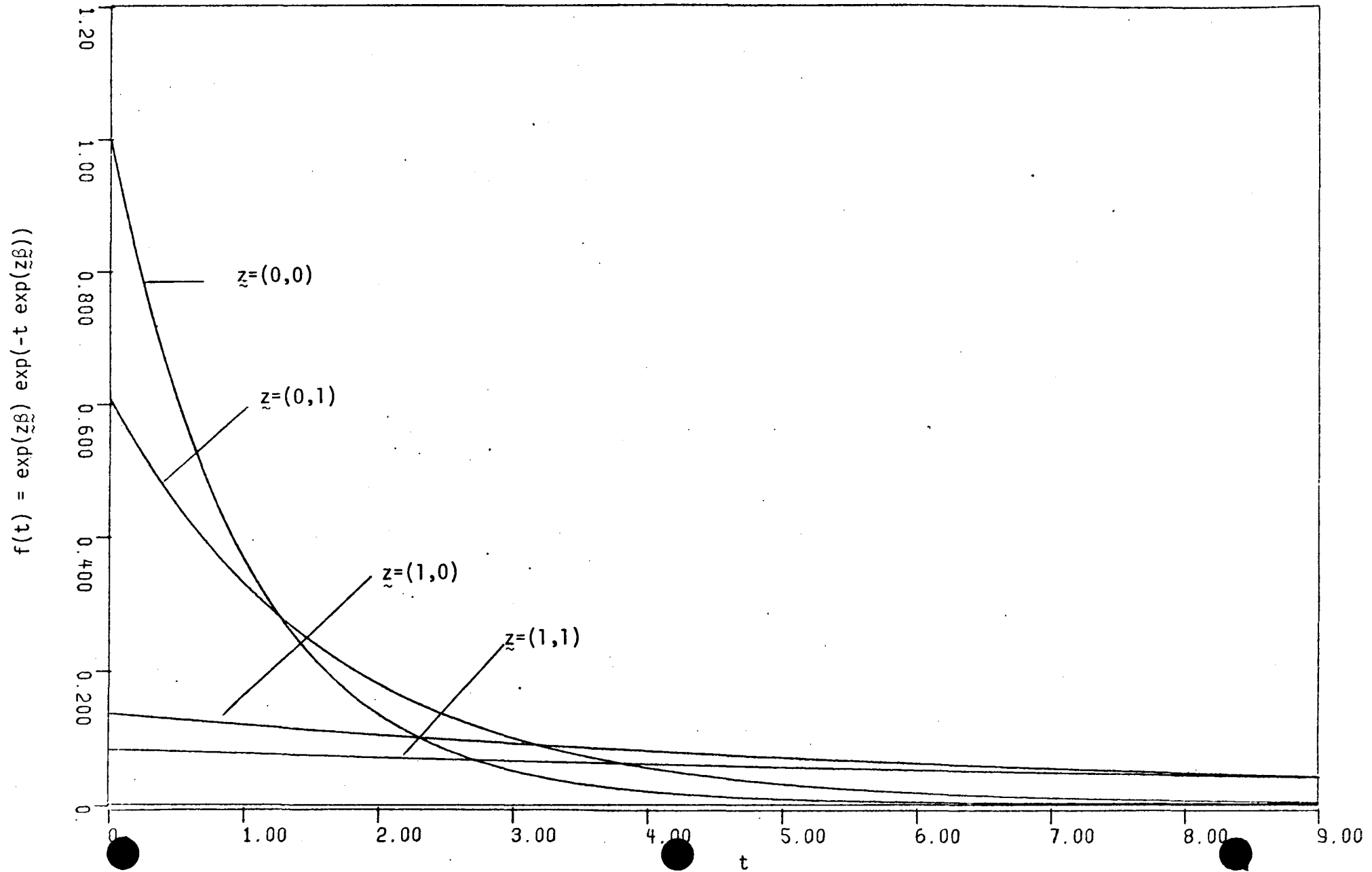
TABLE 2. Comparison of power versus sample size. Entries for $n > 8$ are each based on 1000 samples of exponential failures.

β_2	Sample size					
	4	8	16	32	48	96
-4.00	0.00	0.94				
-3.00	0.00	0.80	1.00			
-2.00	0.00	0.48	0.88	1.00	1.00	
-1.00	0.00	0.15	0.34	0.66	0.87	1.00
-0.50	0.00	0.07	0.11	0.21	0.31	0.57
-0.25	0.00	0.05	0.06	0.07	0.11	0.17
0.0	0.00	0.04	0.04	0.04	0.03	0.03
0.25	0.00	0.05	0.06	0.08	0.12	0.17
0.50	0.00	0.07	0.11	0.19	0.32	0.57
1.00	0.00	0.15	0.33	0.67	0.86	1.00
2.00	0.00	0.48	0.85	1.00	1.00	
3.00	0.00	0.80	1.00			
4.00	0.00	0.94				

TABLE 3. Performance of the test procedure with censoring. Each entry is based on 1000 samples.

Sample size	β_2	no censoring	25%	50%	75%
16	-3.00	1.00	0.93	0.73	0.40
	-2.00	0.88	0.70	0.48	0.24
	-1.00	0.34	0.29	0.20	0.10
	-0.50	0.11	0.11	0.09	0.07
	-0.25	0.06	0.05	0.05	0.04
	0.0	0.04	0.04	0.04	0.04
	0.25	0.06	0.05	0.05	0.04
	0.50	0.11	0.09	0.09	0.05
	1.00	0.33	0.27	0.20	0.10
	2.00	0.85	0.71	0.50	0.24
	3.00	1.00	0.94	0.72	0.36
32	-2.00	1.00	0.96	0.82	0.46
	1.00	0.66	0.53	0.36	0.21
	0.50	0.21	0.18	0.14	0.10
	-0.25	0.07	0.07	0.05	0.05
	0.0	0.04	0.04	0.04	0.03
	0.25	0.08	0.08	0.06	0.06
	0.50	0.19	0.18	0.14	0.08
	1.00	0.67	0.53	0.36	0.18
	2.00	1.00	0.97	0.82	0.46

FIG. 1. Exponential density functions for different concomitant values in the hazard function $\lambda(t)=\exp(\underline{z}\underline{\beta})$, where $\underline{\beta}'=(-2.0 \ -0.5)$.



APPENDIX

Define the following: (The quantities defined are functions of β ; however the β argument is suppressed)

$$A_{ir\xi} = \sum_{\ell \in R(t_i, p_r)} z_{\ell\xi} \exp(z_{\ell}\beta)$$

$$A_{ir} = \sum_{\ell \in R(t_i, p_r)} \exp(z_{\ell}\beta)$$

$$A_{ir\xi\eta} = \sum_{\ell \in R(t_i, p_r)} z_{\ell\xi} z_{\ell\eta} \exp(z_{\ell}\beta)$$

$$B_{i \cdot r v} = \prod_{\substack{u=1 \\ u \neq r \\ u \neq v}}^{m_i} \left\{ \sum_{\ell \in R(t_i, p_u)} \exp(z_{\ell}\beta) \right\}^{-1}$$

$$D_{i\xi} = \sum_{P \in Q_i} \sum_{r=1}^{m_i} \frac{A_{ir\xi}}{(A_{ir})^2} \cdot B_{i \cdot r 0}$$

$$E_i = \sum_{P \in Q_i} \prod_{r=1}^{m_i} \frac{1}{A_{ir}}$$

$$D_{i\xi\eta} = \sum_{P \in Q_i} \sum_{r=1}^{m_i} \frac{A_{ir\xi\eta}}{(A_{ir})^2} \cdot B_{i \cdot r 0}$$

Then, from the Kalbfleisch-Prentice model (2) we determine

$$U_{\xi}(\beta) = \sum_i s_{i\xi} - D_{i\xi} / E_i$$

and

$$I_{\xi\eta}(\beta) = \left[\frac{D_{i\xi} D_{i\eta}}{E_i} + D_{i\xi\eta} - 2 \left\{ \sum_{p \in Q_i} \sum_{r=1}^{m_i} \frac{A_{ir\xi} A_{ir\eta}}{(A_{ir})^3} \cdot B_{i \cdot r0} \right. \right. \\ \left. \left. + \sum_{\substack{v=1 \\ v \neq r}}^{m_i} \frac{A_{ir\xi} A_{iv\eta}}{(A_{ir})^2 (A_{iv})^2} \cdot B_{i \cdot rv} \right\} \right] / E_i$$

To simplify the above formula for the point $\beta = 0$ define the following:

n_i = number of individuals at risk at $R(t_i)$

$n_{ir} = n_{i-r+1}$

$R^+(t_i) = R(t_{i+})$ = set of individuals at risk at time t_i who do not fail at time t_i

$$s_{i\xi} = \sum_{l \in R^+(t_i)} z_{l\xi}$$

$$s_{i\xi}^+ = \sum_{l \in R^+(t_i)} z_{l\xi}^+$$

$$s_{i\xi\eta} = \sum_{l \in R^+(t_i)} z_{l\xi} z_{l\eta}$$

$$s_{i\xi\eta}^+ = \sum_{l \in R^+(t_i)} z_{l\xi}^+ z_{l\eta}$$

$$H_{i\xi} = \sum_{r=1}^{m_i} \frac{1}{n_{ir}} \cdot \left\{ s_{i\xi}^+ + \frac{m_i - r + 1}{m_i} s_{i\xi} \right\}$$

Then

$$U_{\xi} (0) = \sum_{i=1}^k s_{i\xi} - H_{i\xi}$$

$$I_{\xi\eta}(0) = \sum_{i=1}^k H_{i\xi} H_{i\eta} + \sum_{v=1}^{m_i} \frac{1}{n_{iv}} \left(s_{i\xi\eta} \cdot \frac{m_i - v + 1}{m_i} + s_{i\xi\eta}^{\dagger} \right)$$

$$2 \sum_{r \leq v} \sum \frac{1}{n_{ir} n_{iv}} \left(\frac{(m-v+1)(m-r)}{m(m-1)} s_{i\xi} \cdot s_{i\eta} + \left(\frac{1}{m} - \frac{m-r}{m(m-1)} \right) (m-v+1) s_{i\xi\eta} + s_{i\xi}^{\dagger} s_{i\eta}^{\dagger} \right)$$

$$- \left(s_{i\xi} s_{i\eta}^{\dagger} + s_{i\eta} s_{i\xi}^{\dagger} \right) \left(\sum_{r=1}^{m_i} \sum_{v=1}^{m_i} \frac{1}{n_{ir} n_{iv}} \cdot \frac{(m_i - v + 1)}{m_i} \cdot \frac{(m_i - r + 1)}{m_i} \right)$$