Exponential Survivals with Censoring and Explanatory Variables
Author(s): R. L. Prentice
Source: *Biometrika*, Vol. 60, No. 2 (Aug., 1973), pp. 279-288
Published by: Oxford University Press on behalf of Biometrika Trust
Stable URL: https://www.jstor.org/stable/2334539
Accessed: 27-04-2020 09:51 UTC

Biometrika (1973), **60**, 2, *p.* 279
With 3 text-figures
Printed in Great Britain

Exponential survivals with censoring and explanatory variables

BY R. L. PRENTICE
University of Waterloo and State University of New York at Buffalo

SUMMARY

Consideration is given to the analysis of survival times, censored and uncensored, arising from a hazard function that varies exponentially with covariates and is otherwise constant. Significance tests and statements of probabilistic inference are derived for covariate parameters and for the hazard function at average values of the covariate. The results are extended to include a Weibull component in the hazard. Clinical trials on advanced lung cancer patients serve to illustrate the suggested procedures.

Some key words: Exponential and Weibull survivals; Concomitant variables; Exponential hazard; Censoring; Clinical trials; Hazard functions; Life tables; Structural probability.

1. INTRODUCTION

Preliminary analysis by sample survivor functions (Kaplan & Meier, 1958) of a number of clinical trials on advanced cancer patients under study at the Statistical Laboratory at State University of New York at Buffalo suggests an exponential survivor function. A number of explanatory variables describing patient etiology, general medical status and clinical stage of disease are recorded when a patient is taken on study. When such concomitant variables seem important, an exponential relation between failure rate and explanatory variable is frequently suggested. We are then led to consider a hazard function of the form

$$h(t) = \alpha \exp(-\mathbf{x}\boldsymbol{\beta}),$$

where t is survival time, $\mathbf{x} = (x_1, \ldots, x_p)$ is a vector of covariate values, while $\alpha > 0$ and $\boldsymbol{\beta}' = (\beta_1, \ldots, \beta_p)$ are unknown parameters.

Patients, when accepted for trial, are randomized to one of several therapy programmes. The objectives are to determine which covariates have an important relation with survival and to compare efficacy of treatments with respect to longevity, i.e. to compare α values adjusting for necessary covariates.

Techniques for the estimation and comparison of exponential parameters have long been available (Epstein, 1960). When arbitrary censoring is included suggested analyses have been based on the approximate normality of the logarithm of the failure rate (Zelen, 1959; Sampford & Taylor, 1959), or have been nonparametric (Gehan, 1965).

The above hazard function with a single covariate was suggested by Glasser (1967) for use in an application very similar indeed to those that motivated the present work. The same model was discussed by Fiegl & Zelen (1965). Likelihood methods for this model were presented by Sprott & Kalbfleisch (1969). Cox & Snell (1968) also considered this model ($p = 1$) to illustrate the analysis of residuals. Other authors have accommodated concomitant information by representing the exponential mean survival (Fiegl & Zelen, 1965) or failure rate (Cox, 1964; Zippin & Armitage, 1966) as a linear function of the covariate.

These latter models have the disadvantage that the range of values for the exponential parameter is restricted by the value of the covariate parameter. Most recently, Cox (1972) has considered a more general class of hazard functions with a similar relation between hazard and covariate.

Censoring is included by supposing that the only individuals subject to censor are those actually censored. Techniques for model building with respect to covariates and for comparing treatment parameters are included. The independence of the hazard from survival time is relaxed to include the two-parameter Weibull with covariates. Numerical results are presented. The techniques and notation of Fraser (1968, Chapter 2) are employed.

2. Uncensored exponential with covariates

2·1. *Inference on* $(\alpha, \boldsymbol{\beta})$

Suppose the ith individual under study has explanatory variables $\mathbf{x}_i = (x_{1i}, \ldots, x_{pi})$ $(i = 1, \ldots, n)$. Suppose further that corresponding survival times $t_1, \ldots, t_n$ arise from a hazard $h(t_i) = \alpha \exp(-\mathbf{z}_i\boldsymbol{\beta})$, where $\mathbf{z}_i = \{z_{ji} = x_{ji} - \bar{x}_j\}$ $(i = 1, \ldots, n; j = 1, \ldots, p)$. Covariate values have been standardized about the mean in order that α may be interpreted as the failure rate under an average set of covariate values.

The model can be written in the form of a linear regression model of $\log t$ on covariates with nonnormal error. Let $y_i = \log t_i$ and $\gamma = -\log\alpha$. Then $y_i = \gamma + \mathbf{z}_i\boldsymbol{\beta} + w_i$, where $f(w_i) = \exp(w_i - e^{w_i})$ for w_i taking real values.

The model is in the form of Fraser's (1968, Chapter 3) structural model:

$$\prod_{i=1}^{n} f(w_i),$$

$$\mathbf{y} = \gamma\mathbf{1} + \mathbf{Z}\boldsymbol{\beta} + \mathbf{w},$$

where $\mathbf{y}' = (y_1, \ldots, y_n)$, $\mathbf{1}' = (1, \ldots, 1)$, $\mathbf{w}' = (w_1, \ldots, w_n)$ and $\mathbf{Z}'$ is a $p \times n$ matrix defined by $\mathbf{Z}' = (\mathbf{z}_1', \ldots, \mathbf{z}_n')$. The analysis is, in principle, direct. A suitable transformation variable, or conditionally pivotal quantity, is $\{\bar{y}, \tilde{\boldsymbol{\beta}}(\mathbf{y})\}$, where $\bar{y} = \Sigma y_i/n$ and $\tilde{\boldsymbol{\beta}}(\mathbf{y}) = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{y}$. The vector $\mathbf{d} = \mathbf{y} - \bar{y}\mathbf{1} - \mathbf{Z}\tilde{\boldsymbol{\beta}}(\mathbf{y})$ is the corresponding orbital reference point, or ancillary statistic.

If no assumption is made about the parameters $(\alpha, \boldsymbol{\beta})$, then observing $\{\bar{y}, \tilde{\boldsymbol{\beta}}(\mathbf{y})\}$ provides no 'information' about the values of the corresponding error, $\mathbf{w}$, quantities $\{\bar{w}, \tilde{\boldsymbol{\beta}}(\mathbf{w})\}$, in that whatever value of $\{\bar{w}, \tilde{\boldsymbol{\beta}}(\mathbf{w})\}$ is realized one can assert only $\{\bar{w}, \tilde{\boldsymbol{\beta}}(\mathbf{w})\} \in \mathscr{R}^{p+1}$ on the basis of the observed $\mathbf{y}$. This is the purpose of the group relation in Fraser's model.

Inference about $\{\bar{w}, \tilde{\boldsymbol{\beta}}(\mathbf{w})\}$ can then be made from

$$g\{\bar{w}, \tilde{\boldsymbol{\beta}}(\mathbf{w}) | \mathbf{d}\} = k(\mathbf{d}) \exp(n\bar{w}) \exp\left[e^{\bar{w}} \sum_{i=1}^{n} \exp\{\mathbf{z}_i\tilde{\boldsymbol{\beta}}(\mathbf{w}) + d_i\}\right], \tag{1}$$

where $k(\mathbf{d})$ is the integration constant.

Because $\bar{y} = \gamma + \bar{w}$ and $\tilde{\boldsymbol{\beta}}(\mathbf{y}) = \boldsymbol{\beta} + \tilde{\boldsymbol{\beta}}(\mathbf{w})$, an hypothesis $(\gamma_0, \boldsymbol{\beta}_0)$ is, in theory, assessed by comparing $\{\bar{y} - \gamma_0, \tilde{\boldsymbol{\beta}}(\mathbf{y}) - \boldsymbol{\beta}_0\}$ with the above distribution. Probabilistic statements of inference for $(\alpha, \boldsymbol{\beta})$ can correspondingly be made from

$$k'(\mathbf{t})\, \alpha^n \exp\left[-\alpha \sum_{i=1}^{n} \{\exp(-\mathbf{z}_i\boldsymbol{\beta})\, t_i\}\right] (d\alpha/\alpha)\, d\boldsymbol{\beta}, \tag{2}$$

where $k'(\mathbf{t})$ is again the integration constant. If $\boldsymbol{\beta} = \mathbf{0}$, (2) reduces to a $\Gamma(n)$ density for $\alpha\Sigma t_i$.

2·2. *Inferences on* $\boldsymbol{\beta}$

A usual first step in analysis of data as described in § 2·1 would be to decide which covariates merit inclusion in the model. Marginal significance tests for $\boldsymbol{\beta} = \boldsymbol{\beta}_0$ can be made by integrating $\overline{w}$ from expression (1), yielding

$$k(\mathbf{d})\,\Gamma(n)\Big[\sum_{i=1}^{n}\exp\{\mathbf{z}_i\tilde{\boldsymbol{\beta}}(\mathbf{w})+d_i\}\Big]^{-n}, \tag{3}$$

and by comparing $\tilde{\boldsymbol{\beta}}(\mathbf{y})-\boldsymbol{\beta}_0$ with (3). Corresponding to (3), marginal probability statements for $\boldsymbol{\beta}$ can be made from

$$k'(\mathbf{t})\,\Gamma(n)\Big\{\sum_{i=1}^{n}\exp(-\mathbf{z}_i\boldsymbol{\beta})t_i\Big\}^{-n}. \tag{4}$$

One can specialize (3) or (4) to the comparison of two exponentials by setting $x_i = 0$ $(i = 1, \ldots, n_1)$ and $x_i = 1\,(i = n_1+1, \ldots, n_1+n_2)$. Let F' be the ratio of sample means,

$$F' = n_2\sum_{i=1}^{n_1}t_i\Big/\Big(n_1\sum_{i=n_1+1}^{n_1+n_2}t_i\Big).$$

Expression (3) can be written as

$$k'F^{n_1-1}\,dF/(1+n_1/n_2F)^{n_1+n_2},$$

where

$$F = F'e^{\beta} = n_2\exp\{-\beta(\mathbf{w})\}\sum_{i=1}^{n_1}e^{d_i}\Big/\Big(n_1\sum_{i=n_1+1}^{n_1+n_2}e^{d_i}\Big).$$

That is, an hypothesis $\beta = \beta_0$ can be evaluated by comparing $F_0 = F'e^{\beta_0}$ with an $F(2n_1, 2n_2)$ distribution.

For $p = 1$, expressions (3), and (4) are computationally feasible although the integration constant will generally have to be calculated numerically. For $p > 1$, some approximate procedures are desirable. One possibility is to apply a normal approximation to (4) with mean the modal value and covariance matrix minus the inverse of the matrix of second derivatives. This procedure is, for $\boldsymbol{\beta}$, precisely the same as a normal approximation to the original likelihood.

An alternative is to apply model building techniques similar to those used in normal linear regression. Enter covariates singly into the model and choose the one that is most highly significant. Suppose the lth covariate vector is selected. Let $t_i^* = t_i\exp(-z_{li}\hat{\beta}_l)$ be the new vector of observations; $\hat{\beta}_l$ is obtained as the mode of (4) with $\mathbf{Z}' = (\mathbf{z}_{l1}, \ldots, z_{ln})$. The quantities t_i^* are then assumed to arise from a hazard

$$h(t_i^*) = \alpha\exp(-\sum_{j\neq l}\beta_j^*z_{ji}^*),$$

where $\mathbf{Z}^*$ is an $n\times(p-1)$ matrix consisting of covariate vectors orthogonalized to the lth covariate. The above procedure would be repeated until no new covariates explain a significant portion of residual variation. The suggested procedure is similar to regression on residuals except that α is left unrestricted at each stage. The orthogonalizations imply that successive $\hat{\beta}$ values are asymptotically independent so that successive significance levels are also asymptotically independent.

In applications alluded to in the introduction, covariates are often very highly correlated so that only a small number need to be included in the model.

2·3. *Inference on* α

Integration of $\tilde{\boldsymbol{\beta}}(\mathbf{w})$ from (1), or $\boldsymbol{\beta}$ from (2), is generally not feasible and approximate procedures are again required. An asymptotic normal approximation to the original likelihood suggests that $\hat{\alpha}$, or $\hat{\gamma} = -\log\hat{\alpha}$, and $\hat{\boldsymbol{\beta}}$ are asymptotically independent and that

$$\hat{\gamma} = \log\left\{\sum_{i=1}^{n}\exp(-\mathbf{z}_i\hat{\boldsymbol{\beta}})\,t_i/n\right\} \text{ is approximately distributed as } N(\gamma, 1/n)$$

and $\hat{\boldsymbol{\beta}}$ can be obtained from the procedures of § 2·2. Alternatively, we may exploit the above asymptotic independence to insert the modal value of $\tilde{\boldsymbol{\beta}}(\mathbf{w})$ in (1), or $\hat{\boldsymbol{\beta}}$ in (2). Inference about α is then made by comparing

$$\alpha\sum_{i=1}^{n}\exp(-\mathbf{z}_i\hat{\boldsymbol{\beta}})\,t_i$$

with a $\Gamma(n)$ random variable. If the asymptotic properties are in doubt it would be wise to make this comparison for several values of $\boldsymbol{\beta}$ about $\hat{\boldsymbol{\beta}}$.

2·4. *Factorial arrangements of treatments*

The model of § 2·1 can be relaxed to allow α-values to vary with treatment classes. In the application mentioned in the introduction patients would be classified according to therapy received and may be further classified according to covariates that are of special interest or that are difficult to model.

For purposes of illustration suppose patients are classified according to two factors. Let t_{ijk} refer to the kth individual at the ith level of the first factor and jth level of the second factor. Let the corresponding covariate vector be $\mathbf{x}_{ijk} = (x_{1ijk}, \ldots, x_{pijk})$ with assumed hazard $h(t_{ijk}) = \alpha_{ij}\exp(-\mathbf{z}_{ijk}\boldsymbol{\beta})$ for $i = 1, \ldots, r$; $j = 1, \ldots, s$ and $k = 1, \ldots, n_{ij}$. As before set $z_{lijk} = x_{lijk} - \bar{x}_l$. A common covariate parameter has been assumed over classes. If this assumption is relaxed classes can be dealt with separately. The analysis is similar to that of § 2·1 and so will be dealt with briefly.

Let $y_{ijk} = \log t_{ijk}$ and $\gamma_{ij} = -\log\alpha_{ij}$. The model can be written, for $i = 1, \ldots, r$; $j = 1, \ldots, s$ and $k = 1, \ldots, n_{ij}$,

$$\prod_{i=1}^{r}\prod_{j=1}^{s}\prod_{k=1}^{n_{ij}}\exp\{w_{ijk} - \exp(w_{ijk})\},$$

$$y_{ijk} = \gamma_{ij} + \mathbf{z}_{ijk}\boldsymbol{\beta} + w_{ijk},$$

and is again in the form of Fraser's (1968, Chapter 3) regression model. A suitable transformation variable is supplied by the vector of class means $\bar{\mathbf{y}} = (\bar{y}_{11}, \ldots, \bar{y}_{1s}, \bar{y}_{21}, \ldots, \bar{y}_{rs})$ and $\tilde{\boldsymbol{\beta}}(\mathbf{y}) = (\mathbf{V}'\mathbf{V})^{-1}\mathbf{V}'\mathbf{y}$, where $\mathbf{V}$ is given by $v_{lijk} = x_{lijk} - \bar{x}_{lij}$ $(l = 1, \ldots, p)$. The quantities $d_{ijk} = y_{ijk} - \bar{y}_{ij} - \mathbf{v}_{ijk}\tilde{\boldsymbol{\beta}}(\mathbf{y})$ index orbits. Inference about the corresponding error quantities $\{\bar{\mathbf{w}}, \tilde{\boldsymbol{\beta}}(\mathbf{w})\}$ can be made from

$$g\{\bar{\mathbf{w}}, \tilde{\boldsymbol{\beta}}(\mathbf{w})|\mathbf{d}\} = k(\mathbf{d})\exp(\Sigma\Sigma n_{ij}\bar{w}_{ij})\exp\{-\Sigma\Sigma\exp(\bar{w}_{ij})\,\Sigma\exp(\mathbf{v}_{ijk}\tilde{\boldsymbol{\beta}}(\mathbf{w}) + d_{ijk})\}. \quad (5)$$

The observed $\{\bar{\mathbf{y}}, \tilde{\boldsymbol{\beta}}(\mathbf{y})\}$ are related to $\{\bar{\mathbf{w}}, \tilde{\boldsymbol{\beta}}(\mathbf{w})\}$ by

$$\bar{y}_{ij} = \gamma_{ij} + \sum_{l=1}^{p}\beta_l(\bar{x}_{lij.} - \bar{x}_l) + \bar{w}_{ij} \quad (i = 1, \ldots, r;\ y = 1, \ldots, s),$$

$$\tilde{\boldsymbol{\beta}}(\mathbf{y}) = \boldsymbol{\beta} + \tilde{\boldsymbol{\beta}}(\mathbf{w}).$$

Expression (5) provides the basis for tests about the γ's and $\boldsymbol{\beta}$ while the corresponding probability statement about the α's and $\boldsymbol{\beta}$ is

$$k'(\mathbf{t}) \prod_{i=1}^{r} \prod_{j=1}^{s} \alpha_{ij}^{(n_{ij}-1)} \exp\{\sum_{i,j,k} \alpha_{ij} \exp(-\mathbf{z}_{ijk}\boldsymbol{\beta})\, t_{ijk}\}\, d\boldsymbol{\alpha}\, d\boldsymbol{\beta}.$$

Tests of significance for $\boldsymbol{\beta} = \boldsymbol{\beta}_0$ are made by comparing $\tilde{\boldsymbol{\beta}}(\mathbf{w}) = \tilde{\boldsymbol{\beta}}(\mathbf{y}) - \boldsymbol{\beta}_0$ with

$$g'\{\tilde{\boldsymbol{\beta}}(\mathbf{w}) | \mathbf{d}\} = K(\mathbf{d}) \prod_{i=1}^{r} \prod_{j=1}^{s} \left\{ \sum_{k=1}^{n_{ij}} \exp(\mathbf{v}_{ijk}\tilde{\boldsymbol{\beta}}(\mathbf{w}) + d_{ijk}) \right\}^{-n_{ij}} \tag{6}$$

and marginal probability statements for $\boldsymbol{\beta}$ are made from

$$K'(t) \prod_{i=1}^{r} \prod_{j=1}^{s} \left\{ \sum_{k=1}^{n_{ij}} \exp(-\mathbf{z}_{ijk}\boldsymbol{\beta})\, t_{ijk} \right\}^{-n_{ij}}. \tag{7}$$

The remarks of § 2·2 apply equally well to expressions (6) and (7) while the remarks of § 2·3 need to be altered only in that the indicator vector for the (i,j)th class is not orthogonal to the columns of $\mathbf{Z}$; the vector dot product with the lth column of $\mathbf{Z}$ is $\bar{x}_{lij.} - \bar{x}_{l}$. The justification for inserting maximum likelihood estimates for $\boldsymbol{\beta}$ in making inference about the α's is then weakened in that $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$ are not asymptotically independent. It is then necessary to rely on $\hat{\boldsymbol{\beta}}$ being precisely estimated on approximate orthogonality of the abovementioned vectors in making such an insertion.

3. Exponential survivals with covariates and censored data

In extending the model of § 2 to include censoring, the point of view is taken that the censoring point may be ignored for individuals that die during the study period. In the context of experiments mentioned in the introduction this assumption is not too unrealistic. Typically, entry into the study occurs as soon as the qualified patient is present with disease. The time from study activation to analysis is often large with respect to expected survival. Thus for patients entering the study early the censoring point is of no importance, while those entering late very likely give rise to censored survivals. Further, if the study has been terminated for some time, the only censored observations occur for patients lost to follow up. Such patients are, in a real sense, the only ones subject to censoring.

Suppose, then, that $t_1, \ldots, t_r$ are observed survival times while $t_{r+1}, \ldots, t_n$ are censored at $t_{r+1}^0, \ldots, t_n^0$, respectively. Assume a hazard $h(t_i) = \alpha \exp(-\mathbf{z}_i\boldsymbol{\beta})$, as before, where $z_{ji} = x_{ji} - \bar{x}_{(j)}$ and $\bar{x}_{(j)}$ is the jth covariate mean for individuals that have died.

The data can be written, in the notation of § 2·1, as $y_i = \gamma + \mathbf{z}_i\boldsymbol{\beta} + w_i$ $(i = 1, \ldots, n)$ with $y_1, \ldots, y_r$ observed and $y_i > y_i^0 = \log t_i^0$ for $i = r+1, \ldots, n$. If we define $\bar{y}$, $\bar{w}$, $\tilde{\boldsymbol{\beta}}(\mathbf{y})$ and $\tilde{\boldsymbol{\beta}}(\mathbf{w})$ as in § 2·1 but based only on uncensored individuals and set $d_i = y_i - \bar{y} - \mathbf{z}_i\tilde{\boldsymbol{\beta}}(\mathbf{y})$ for $i = 1, \ldots, n$ and $d_i^0 = y_i^0 - \bar{y} - \mathbf{z}_i\tilde{\boldsymbol{\beta}}(\mathbf{y})$ for $i = r+1, \ldots, n$, then the model and data may be written

$$\begin{gathered} g\{\bar{w}, \tilde{\boldsymbol{\beta}}(\mathbf{w}) \mid d_1, \ldots, d_r, d_{r+1} > d_{r+1}^0, \ldots, d_n > d_n^0\}, \\ \bar{y} = \gamma + \bar{w}, \quad \tilde{\boldsymbol{\beta}}(\mathbf{y}) = \boldsymbol{\beta} + \tilde{\boldsymbol{\beta}}(\mathbf{w}), \quad \mathbf{d} \in \mathscr{D}, \end{gathered} \tag{8}$$

where $\mathscr{D}$ is the event defined by $d_1, \ldots, d_r, d_{r+1} > d_{r+1}^0, \ldots, d_n > d_n^0$. The model (8) is again in the form of Fraser's structural model provided the set of censored individuals is held fixed. The quantity $\{\bar{y}, \tilde{\boldsymbol{\beta}}(\mathbf{y})\}$ acts as transformation variable and $\mathscr{D}$ indexes the orbit. Inference about the realized $\{\bar{w}, \tilde{\boldsymbol{\beta}}(\mathbf{w})\}$ can be made from

$$g\{\bar{w}, \tilde{\boldsymbol{\beta}}(\mathbf{w}) | \mathscr{D}\} = k(\mathscr{D}) \exp(r\bar{w}) \exp\left[-e^{\bar{w}} \sum_{i=1}^{n} \exp\{\mathbf{z}_i\tilde{\boldsymbol{\beta}}(\mathbf{w}) + d_i\} \right],$$

where d_i has been written for d_i^0 for $i > r$; t_i will also be written for t_i^0. Marginal tests of significance are made by comparing $\tilde{\boldsymbol{\beta}}_0(\mathbf{w}) = \tilde{\boldsymbol{\beta}}(\mathbf{y}) - \boldsymbol{\beta}_0$ with

$$K(\mathscr{D})\left[\sum_{i=1}^{n} \exp\{\mathbf{z}_i \tilde{\boldsymbol{\beta}}(\mathbf{w}) + d_i\}\right]^{-r}.$$

The corresponding probability statement for $(\alpha, \boldsymbol{\beta})$ is

$$k'(\mathbf{t})\,\alpha^{r-1} \exp\left\{-\alpha \sum_{i=1}^{n} \exp(-\mathbf{z}_i\boldsymbol{\beta})\, t_i\right\} d\alpha\, d\boldsymbol{\beta}$$

and for $\boldsymbol{\beta}$ alone is

$$K'(\mathbf{t})\left\{\sum_{i=1}^{n} \exp(-\mathbf{z}_i\boldsymbol{\beta})\, t_i\right\}^{-r} d\boldsymbol{\beta}.$$

The remarks of §§ 2·2, 2·3 and 2·4 apply equally well here with r replacing n as required; in expressions (6) and (7) the number of deaths in the (i,j)th class replaces $n_{ij}+h$ when n_{ij} appears as an exponent.

4. The Weibull distribution with covariates

It may be desired to relax the independence of the hazard from survival time to accommodate a broader range of applications or as a test for the suitability of the above models. A Weibull component may be inserted in the model of § 2 and the hazard written

$$h(t) = \alpha t^{\lambda-1} \exp(-\mathbf{z}\boldsymbol{\beta}).$$

The additional parameter λ can be modelled as a scale parameter. On the basis of survivals $t_1, \ldots, t_n$ and with $\gamma = -\log(\alpha/\lambda)/\lambda$, $\boldsymbol{\delta} = \boldsymbol{\beta}/\lambda$ and $y_i = \log t_i$ $(i = 1, \ldots, n)$, the model can be written

$$\prod_{i=1}^{n} \exp\{w_i - \exp(w_i)\},$$

$$\mathbf{y} = \gamma\mathbf{1} + \mathbf{Z}\boldsymbol{\delta} + \mathbf{w}/\lambda.$$

The quantities $\{\bar{y}, \tilde{\boldsymbol{\delta}}(\mathbf{y}), s(\mathbf{y})\}$ provide a transformation variable, where $\bar{y}$ and $\tilde{\boldsymbol{\delta}}(\bar{y})$ are the sample mean and least squares regression coefficients for $\mathbf{y}$ as before and $s(y)$ is the length of the residual vector, $s^2(\mathbf{y}) = \Sigma(y_i - \bar{y} - \mathbf{z}_i\tilde{\boldsymbol{\delta}})^2$, while the quantities $d_i = s^{-1}(\mathbf{y})(y_i - \bar{y} - \mathbf{z}_i\tilde{\boldsymbol{\delta}})$ index orbits.

For the sake of brevity only the statements of probabilistic inference from this model are recorded. Tests of significance, equivalent to the type given above, can be computed as tail area probabilities from the structural probability statements. Inference about $(\alpha, \boldsymbol{\delta}, \lambda)$ can be made from

$$k(\mathbf{t})\,\alpha^{n-1}\lambda^{-1} \prod_{i=1}^{n} t_i^{\lambda-1} \exp\left[-(\alpha/\lambda)\left\{\sum_{i=1}^{n} t_i^{\lambda} \exp(-\lambda\mathbf{z}_i\boldsymbol{\delta})\right\}\right] d\alpha\, d\boldsymbol{\delta}\, d\lambda,$$

and α can be integrated out to yield

$$k'(\mathbf{t})\,\lambda^{n-1} \prod_{i=1}^{n} t_i^{\lambda-1} \left\{\sum_{i=1}^{n} t_i^{\lambda} \exp(-\lambda\mathbf{z}_i\boldsymbol{\delta})\right\}^{-n} d\boldsymbol{\delta}\, d\lambda. \tag{9}$$

Marginal inference about λ can be made by integrating (9) over $\boldsymbol{\delta}$. For $p = 1$ this is possible, but generally to test $\lambda = 1$ it would be more convenient to classify data on covariate values, or intervals, and to compare $\lambda = 1$ to

$$K'(\mathbf{t})\,\lambda^{m-1} \prod_{i=1}^{m} t_i^{\lambda-1} \left(\sum_{i=1}^{m} t_i^{\lambda}\right)^{-m}$$

within each such class, m denoting the number of individuals in such a class. In the mood of § 3 expression (9) becomes

$$k'(\mathbf{t})\,\lambda^{r-1}\prod_{i=1}^{r}t_i^{\lambda-1}\left\{\sum_{i=1}^{n}t_i^{\lambda}\exp\left(-\lambda\mathbf{z}_i\boldsymbol{\delta}\right)\right\}^{-r}d\boldsymbol{\delta}\,d\lambda \tag{10}$$

if $t_{r+1},\ldots,t_n$ are censored.

If data are initially classified as in § 2·4, joint statements for $(\boldsymbol{\delta},\lambda)$ are a product of terms (10) over classes.

5. An illustration

Table 1 presents survival data, double daggers indicate censoring, on 137 advanced lung cancer patients as collected by the Veterans Administration Lung Cancer Study Group. Patients were randomized according to one of two chemotherapeutic agents (1, standard; 2, test). Of particular interest is the possible differential effects of therapy on tumour cell

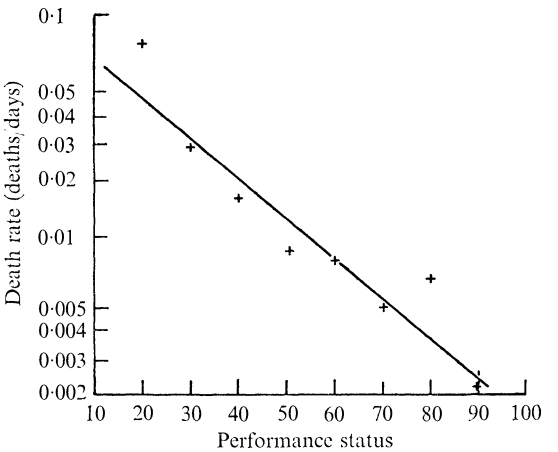

Fig. 1. Death rate versus performance status.

type. Tumours are classified into one of four broad groups (1, squamous; 2, small; 3, adeno; 4, large). The covariates recorded when a patient is taken on study include performance status, a measure of general medical status on a scale 10, 20, ..., 90; 10, 20, 30 – completely hospitalized, 40, 50, 60 – partial confinement to hospital, 70, 80, 90 – able to care for self, time from diagnosis to starting on study (months), age, and previous therapy (0, no; 10, yes). Denote covariate vectors by $\mathbf{x}_1,\ldots,\mathbf{x}_4$, respectively.

The assumed hazard for the survival t_{ijk} days of the kth patient in treatment group i and tumour cell type j is

$$\alpha_{ij}\exp\left\{-\sum_{l=1}^{4}\beta_l(x_{lijk}-\bar{x}_l)\right\}. \tag{11}$$

The exponential relation between covariate and hazard was suggested by plots of exponential death rate *versus* covariate value. Figure 1 is a plot of death rate on a log scale versus performance status. A linear relation is suggested.

Covariates 1 to 4 were entered singly into the hazard (11) and significance levels, twice the tail area, for $\beta_j=0$ $(j=1,\ldots,4)$ were calculated. These values are 0·00, 0·61, 0·72 and 0·57, respectively, while the corresponding $\hat{\beta}$'s are 0·029, −0·005, −0·003 and −0·012. The marginal probability densities for the β_j's are very nearly normal with slight irregularities

Table 1. *Data for lung cancer patients: days of survival (t), performance status (x_1), months from diagnosis (x_2), age in years (x_3), and prior* therapy (x_4)*

t	x_1	x_2	x_3	x_4
Standard, squamous†				
72	60	7	69	0
411	70	5	64	10
228	60	3	38	0
126	60	9	63	10
118	70	11	65	10
10	20	5	49	0
82	40	10	69	10
110	80	29	68	0
314	50	18	43	0
100‡	70	6	70	0
42	60	4	81	0
8	40	58	63	10
144	30	4	63	0
25‡	80	9	52	10
11	70	11	48	10
Standard, small				
30	60	3	61	0
384	60	9	42	0
4	40	2	35	0
54	80	4	63	10
13	60	4	56	0
123‡	40	3	55	0
97‡	60	5	67	0
153	60	14	63	10
59	30	2	65	0
117	80	3	46	0
16	30	4	53	10
151	50	12	69	0
22	60	4	68	0
56	80	12	43	10
21	40	2	55	10
18	20	15	42	0
139	80	2	64	0
20	30	5	65	0
31	75	3	65	0
52	70	2	55	0
287	60	25	66	10
18	30	4	60	0
51	60	1	67	0
122	80	28	53	0
27	60	8	62	0
54	70	1	67	0
7	50	7	72	0
63	50	11	48	0
392	40	4	68	0
10	40	23	67	10
Standard, adeno				
8	20	19	61	10
92	70	10	60	0
35	40	6	62	0
117	80	2	38	0
132	80	5	50	0
12	50	4	63	10
162	80	5	64	0
3	30	3	43	0
95	80	4	34	0
Standard, large				
177	50	16	66	10
162	80	5	62	0
216	50	15	52	0
553	70	2	47	0
278	60	12	63	0
12	40	12	68	10
260	80	5	45	0
200	80	12	41	10
156	70	2	66	0
182‡	90	2	62	0
143	90	8	60	0
105	80	11	66	0
103	80	5	38	0
250	70	8	53	10
100	60	13	37	10
Test, squamous				
999	90	12	54	10
112	80	6	60	0
87‡	80	3	48	0
231‡	50	8	52	10
242	50	1	70	0
991	70	7	50	10
111	70	3	62	0
1	20	21	65	10
587	60	3	58	0
389	90	2	62	0
33	30	6	64	0
25	20	36	63	0
357	70	13	58	0
467	90	2	64	0
201	80	28	52	10
1	50	7	35	0
30	70	11	63	0
44	60	13	70	10
283	90	2	51	0
15	50	13	40	10
Test, small				
25	30	2	69	0
103‡	70	22	36	10
21	20	4	71	0
13	30	2	62	0
87	60	2	60	0
2	40	36	44	10
20	30	9	54	10
7	20	11	66	0
24	60	8	49	0
99	70	3	72	0
8	80	2	68	0
99	85	4	62	0
61	70	2	71	0
25	70	2	70	0
95	70	1	61	0
80	50	17	71	0
51	30	87	59	10
29	40	8	67	0
Test, adeno				
24	40	2	60	0
18	40	5	69	10
83‡	99	3	57	0
31	80	3	39	0
51	60	5	62	0
90	60	22	50	10
52	60	3	43	0
73	60	3	70	0
8	50	5	66	0
36	70	8	61	0
48	10	4	81	0
7	40	4	58	0
140	70	3	63	0
186	90	3	60	0
84	80	4	62	10
19	50	10	42	0
45	40	3	69	0
80	40	4	63	0
Test, large				
52	60	4	45	0
164	70	15	68	10
19	30	4	39	10
53	60	12	66	0
15	30	5	63	0
43	60	11	49	10
340	80	10	64	10
133	75	1	65	0
111	60	5	64	0
231	70	18	67	10
378	80	4	65	0
49	30	3	37	0

* 0, no prior therapy, 10, prior therapy.
† Standard therapy, squamous tumour cell type.
‡ Censored survival.

in the tails. Figure 2 provides a plot of this density for β_1. Since $\beta_1 = 0$ was the most highly significant, $\mathbf{x}_2$, $\mathbf{x}_3$ and $\mathbf{x}_4$ were orthogonalized to $\mathbf{x}_1$ and $\beta_1 = \hat{\beta}_1$ was inserted. Significance levels for the new coefficients $\beta_j^* = 0\,(j = 2, 3, 4)$ were calculated as the orthogonalized vectors were entered singly into the hazard. These are 0·80, 0·56 and 0·46, respectively, with corresponding β_j^* values 0·001, 0·005 and $-0{\cdot}015$.

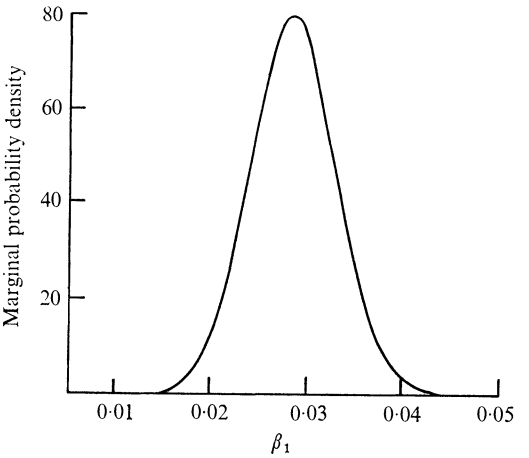

Fig. 2. Marginal probability density for β_1.

It appears that if (11) is appropriate only $\mathbf{x}_1$ needs to be included in the model. The quantities

$$\hat{\theta}_{ij} = \hat{\alpha}_{ij}^{-1} = \left[\sum_{k=1}^{n_{ij}} \exp\{-\hat{\beta}_1(x_{1ijk} - \bar{x}_1)\}\, t_{ijk}\right] \Big/ r_{ij}$$

are presented in Table 2. The number of deaths r_{ij} is bracketed.

Table 2. *Expected survival estimates in days adjusted for performance status*

	Cell type			
Therapy	squamous	small	adeno	large
Standard	142·53 (13)	101·79 (28)	48·01 (9)	140·47 (14)
Test	190·63 (18)	53·04 (17)	61·67 (17)	96·50 (12)

Since β_1 is rather precisely estimated (Fig. 2), the quantities $\hat{\theta}_{2j}/\hat{\theta}_{1j}$ $(j = 1, \ldots, 4)$ may reasonably be regarded as arising from an $F(2r_{2j}, 2r_{1j})$ distribution for purposes of testing therapy equality within cell types.

Finally the appropriateness of the exponential assumption is examined by writing the hazard as $\alpha_{ij} t_{ijk}^{\lambda-1} \exp\{-\beta_1(x_{1ijk} - \bar{x}_1)\}$. The marginal probability distribution for λ given by

$$k(\mathbf{t}) \int \prod_{i=1}^{2} \prod_{j=1}^{4} \left(\prod_{k=1}^{r_{ij}} t_{ijk}^{\lambda-1} \lambda^{r_{ij}} \Big/ \left[\sum_{k=1}^{n_{ij}} \exp\{-\beta_1(x_{1ijk} - \bar{x}_1)\}\, t_{ijk}^{\lambda}\right]^{r_{ij}} \right) \lambda^{-1} d\beta_1$$

is plotted in Fig. 3. The hypothesis $\lambda = 1$ seems to be reasonably in keeping with the data.

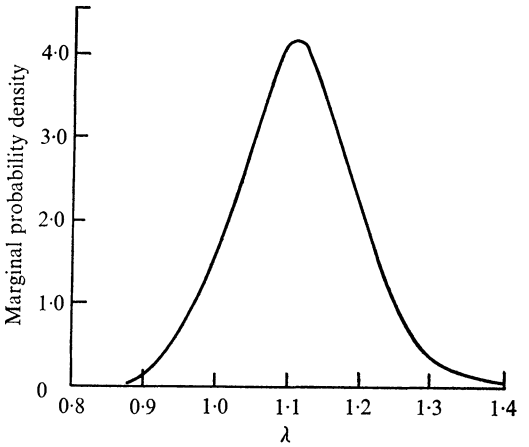

Fig. 3. Marginal probability density for λ.

The author would like to thank the Veterans Administration Lung Cancer Study Group for permission to use their data. He would also like to thank Professors J. D. Kalbfleisch and M. Zelen for helpful suggestions and Mr R. Greenberg for computational assistance. The work was supported by a Veterans Administration Grant.

REFERENCES

Cox, D. R. (1964). Some applications of exponential ordered scores. *J. R. Statist. Soc.* B **26**, 103–10.

Cox, D. R. (1972). Regression models and life tables (with discussion). *J. R. Statist. Soc.* B **34**, 187–220.

Cox, D. R. & SNELL, E. J. (1968). A general definition of residuals (with discussion). *J. R. Statist. Soc* B **30**, 248–75.

EPSTEIN, B. (1960). Estimation from life test data. *Technometrics* **2**, 447–54.

FIEGL, P. & ZELEN, M. (1965). Estimation of exponential survival probabilities with concomitant information. *Biometrics* **21**, 826–38.

FRASER, D. A. S. (1968). *The Structure of Inference.* New York: Wiley.

GEHAN, E. A. (1965). A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. *Biometrika* **52**, 203–23.

GLASSER, M. (1967). Exponential survival with covariance. *J. Am. Statist. Assoc.* **62**, 561–8.

KAPLAN, E. L. & MEIER, P. (1958). Nonparametric estimation from incomplete observations. *J. Am. Statist. Assoc.* **53**, 457–81.

SAMPFORD, M. R. & TAYLOR, J. (1959). Censored observations in randomized block experiments. *J. R. Statist. Soc.* B **21**, 214–37.

SPROTT, D. A. & KALBFLEISCH, J. D. (1969). Examples of likelihood and comparison with point estimates and large sample approximations. *J. Am. Statist. Assoc.* **64**, 468–84.

ZELEN, M. (1959). Factorial experiments in life testing. *Technometrics* **1**, 269–88.

ZIPPIN, C. & ARMITAGE, P. (1966). Use of concomitant variables and exponential survival parameter. *Biometrics* **22**, 665–72.

[*Received March* 1972. *Revised September* 1972]