subfaculteit der econometrie

RESEARCH MEMORANDUM

TILBURG UNIVERSITY
DEPARTMENT OF ECONOMICS
Postbus 90135 - 5000 LE Tilburg
Netherlands

FEW
104

MULTIPLE FAILURE RATES AND OBSERVATIONS OF TIME DEPENDENT COVARIABLES
(Part 1: Theory)

A.L. Hempenius and P.G.H. Mulder
november 1981

MULTIPLE FAILURE RATES AND OBSERVATIONS OF TIME DEPENDENT COVARIABLES
(Part 1: Theory)

A.L. Hempenius and P.G.H. Mulder

1. Introduction

In this paper a method is presented for analyzing, for censored failure time data, the relation between: (1) failure rates of a number of mutually exclusive and simultaneously acting (so-called "competing") failure causes and (2) values of time dependent covariables, including possibly the covariable "time" itself. These covariables are supposedly measured at predetermined time points.

Some examples from epidemiology, medicine, and economics are:

(i) The development over time of cardiovascular disease in an originally disease-free cohort of persons in relation to annual values of blood pressure, cholesterol level, ECG-anomalies, smoking, sex and age, taking into account the two other (competing) failure causes: mortality from other causes and disease-free withdrawal from the cohort.

(ii) The survival prognosis after operation on a certain type of cancer, in the presence of various competing mortality causes (including the given kind of cancer) of patients discharged from hospital, in relation to various medical therapies, the time since the operation, age, and sociodemographic characteristics.

(iii) The termination of the duration of unemployment, firstly of course by finding a job and secondly by other causes, such as transfer to some social security scheme (say a disability scheme), other than the unemployment scheme, or mortality. Possibly time-dependent covariables are educational and/or professional status, age, the duration of unemployment and the acquisition of new skills. See, for example, Nickel (1979) and Lancaster and Nickel (1980).

In Section 2 a review of the relevant theory of competing risks is presented, as this theory may be used to describe the distributions generating type and time of failure for a given "subject". The parameters of these distributions depend on the values of a number of explanatory variables characteristic for this subject. This dependency is also modelled in Section 2. In Section 3 the likelihood function suitable for estimating the relevant parameters of the model is presented.

2. Elements of the theory of competing risks[1)]

2.1. Theoretical and Observed Failure Times

A given subject, free from any "failures" at time point zero, is in the course of time subject to K (mutually exclusive) simultaneously acting or competing risks or failure causes. At some point in time, and by one of the K failure causes, the subject considered actually experiences "failure". This may be formalized as follows.

Let the observed random variable (r.v.) X denote the time of failure (from one of the K risk causes) and let the unobserved theoretical random variable $X_j \geq 0$ $(j = 1,\ldots,K)$ be the time of failure if risk j were the only risk present. Then X is equal to the smallest of these K theoretical failure times $X_1,\ldots,X_K$:

$$(2.1) \qquad X = \min(X_1,\ldots,X_K).$$

These theoretical failure times have a joint distribution defined by the so-called (supposedly continuous) joint survival function S:

$$(2.2) \qquad S(x_1,\ldots,x_K) = \Pr(X_1 > x_1,\ldots,X_K > x_K).$$

The (observable) probability of still being failure-free at time x (i.e. of surviving after time x) is a value of the survival function $\bar{F}(x)$ defined for the r.v. X of (2.1):

$$(2.3) \qquad \bar{F}(x) = \Pr(X > x).$$

1) See also David and Moeschberger (1978).

As X exceeds each X_j, according to (2.1), $\bar{F}(x)$ may be expressed as follows:

$$(2.4) \qquad \bar{F}(x) = Pr(X_1 > x,\dots,X_K > x)$$

$$= S(x,\dots,x).$$

The probability density function (p.d.f.) of X, f(x), thus is:

$$(2.5) \qquad f(x) = \frac{-d\bar{F}(x)}{dx}$$

$$= \frac{-dS(x,\dots,x)}{dx}.$$

The (observable) probability of failure, at some time, from risk cause j is denoted by π_j ($j = 1,\dots,K$):

$$(2.6) \qquad \pi_j = Pr[\min(X_1,\dots,X_K) = X_j]$$

with $\Sigma\pi_j = 1$. From (2.6) it follows that π_j may be expressed in terms of the joint distribution of the r.v.'s $X_1,\dots,X_K$, as specified by S in (2.2).

The (conditional) probability $f_j(x)dx$ of failure in the interval $(x,x + dx)$ from risk cause j, given failure from cause j, is:

$$(2.7) \qquad f_j(x)dx = \frac{S(x,\dots,x) - S(x,\dots,x + dx,\dots,x)}{\pi_j},$$

so that the (conditional) <u>p.d.f. of failure time given failure from cause j</u>, is:

$$(2.8) \qquad f_j(x) = \frac{-[\frac{\partial}{\partial x_j} S(x_1,\dots,x_j,\dots,x_K)]_{x_1 = \dots = x_K = x}}{\pi_j}.$$

<u>The (unconditional) probability of failure in the interval $(x, x + dx)$ from cause j thus is $\pi_j\ f_j(x)dx$.</u>

The following relation between f(x), $f_j(x)$ and π_j ($j = 1,\dots,K$) follows from the above definitions:

$$(2.9) \qquad f(x) = \sum_{j=1}^{K} \pi_j \, f_j(x).$$

2.2. Failure rates

By means of $\bar{F}(x)$ and $f(x)$ one defines the well-known failure rate function $\lambda(x)$ of X as follows:

$$(2.10) \qquad \lambda(x) = \frac{f(x)}{\bar{F}(x)} = -\frac{d}{dx} \ln \bar{F}(x).$$

The failure rate (function) defines the following probability: $\lambda(x)dx$ is the (conditional) probability that a subject, failure-free at time x, will experience failure in the interval (x, x + dx), from any of the K risk causes.

The marginal failure rate function for cause j ($j = 1,\dots,K$), in the presence of all K risks, is denoted by $\lambda_j(x)$,so that $\lambda_j(x)dx$ is the (conditional) probability of failure from risk cause j in the interval (x, x + dx), for a subject failure-free at time x:

$$(2.11) \qquad \lambda_j(x)dx = \frac{S(x,\dots,x) - S(x,\dots,x+dx,\dots,x)}{S(x,\dots,x)}.$$

By letting dx approach 0, one has:

$$(2.12) \qquad \lambda_j(x) = \frac{-\left[\frac{\partial}{\partial x_j} S(x_1,\dots,x_j,\dots,x_K)\right]_{x_1 = \dots = x_K = x}}{S(x,\dots,x)}.$$

From (2.8) and (2.12) the following relation follows:

$$(2.13) \qquad \pi_j \, f_j(x) = \lambda_j(x) \, \bar{F}(x).$$

The probability of failure in the interval (x, x + dx) from cause j may thus also be written as $\lambda_j(x) \, \bar{F}(x)dx$.

The failure rate function $\lambda(x)$ of the observed failure time X can be expressed in the failure rates $\lambda_j(x)$:

$$(2.14) \qquad \lambda(x) = \sum_{j=1}^{K} \lambda_j(x),$$

which follows from

$$(2.15)\quad \lambda(x) = f(x)/\bar{F}(x)$$

$$= -\left[\frac{d}{dx} S(x,\ldots,x)\right]/S(x,\ldots,x) \ ,$$

and from (2.12).

A special case of the marginal failure rate functions $\lambda_j(x)$:

$$(2.16)\quad \lambda_j(x) = c_j\,\lambda(x) \qquad (j = 1,\ldots,K)$$

defines the so-called proportional hazards model. In this model the ratios of the λ_j are independent of the time x. For the c_j one can easily prove:

$$(2.17)\quad c_j = \pi_j \qquad (j = 1,\ldots,K).$$

From (2.13), (2.15), (2.16) and (2.17) it follows that in this case

$$(2.18)\quad f_j(x) = f(x) \qquad (j = 1,\ldots,K),$$

so that in the proportional hazard model (2.16) all cause-specific failure time distributions $f_j(x)$ are equal to the overall failure time distribution f(x). Hence the proportional hazards model implies independency of time and cause of failure.

In contrast to $\lambda_j(x)$, which is defined in the presence of all risks, a theoretical marginal failure rate function in the absence of all other risks, $r_j(x)$, is defined analogous to (2.10) as:

$$(2.19)\quad r_j(x) = \frac{p_j(x)}{\bar{P}_j(x)} \ ,$$

with $p_j(x)$ the p.d.f. of the unobserved r.v. X_j and $\bar{P}_j$ its decumulative distribution function.

If the r.v.'s $X_1,\ldots,X_K$ are independent r.v.'s, then

(2.20) $\lambda_j(x) = r_j(x)$.

This may be proved directly from (2.12) and also as follows. The (conditional) probability of failure from risk cause j (and not from other causes) in the interval (x, x + dx), given survival (no failure) until time x, is:

$$(2.21) \quad \lambda_j(x)dx = \frac{[p_j(x)dx] \times \prod_{i \neq j} \bar{P}_j(x)}{\bar{F}(x)}$$

$$= \frac{p_j(x)dx}{\bar{P}_j(x)} ,$$

where use has been made of $\bar{F}(x) = Pr(X_1 > x,\ldots,X_K > x) = Pr(X_1 > x) \times \ldots \times Pr(X_K > x) = \bar{P}_1(x) \times \ldots \times \bar{P}_K(x)$.

Result (2.20) implies that

$$(2.22) \quad \lambda(x) = \sum_{j=1}^{K} r_j(x) ,$$

if the unobservable r.v.'s $X_1,\ldots,X_K$ are independent.

Some relevant probabilities can be expressed in terms of the $\lambda_j(x)$ and $\lambda(x)$. For example, $\bar{F}(x)$, the probability of failure later than x, is according to (2.10):

$$(2.23) \quad \bar{F}(x) = \exp[-\int_0^x \lambda(t)dt].$$

The so-called <u>crude probability</u> of failure (i.e. in the presence of all risks) from cause j in the interval (a,b), given survival until time a, $Q_j(a,b)$, is equal to

$$(2.24) \quad Q_j(a,b) = \frac{1}{\bar{F}(a)} \int_a^b \pi_j \ f_j(x) \ dx = \int_a^b \lambda_j(x) \ \exp[-\int_a^x \lambda(t)dt] \ dx$$

as follows from (2.13) and (2.23).

2.3. Modelling the dependency of the failure process on the covariables[2)]

In this subsection the covariables enter the failure process. As will be seen in Section 3, the likelihood function may be expressed in terms of $\bar{F}(x)$ and the $\lambda_j(x)$ $(j = 1,\ldots,K)$, with $\bar{F}(x)$ the (overall) survival function (2.3) and $\lambda_j(x)$ the marginal failure rate function for cause j, in the presence of all risks. As $\bar{F}(x)$ can be expressed in the $\lambda_j(x)$, see (2.14) and (2.23), it thus is sufficient to describe the dependency of the $\lambda_j(x)$ $(j = 1,\ldots,K)$ on the covariables, which will be assembled in the vector z. In the context of the competing risk theory of Section 2.2, which uses the theoretical r.v.'s $X_1,\ldots,X_K$, it is natural to specify first the dependency of $S(x_1,\ldots,x_K\ z)$ on z and then to derive the $\lambda_j(x\ z)$ $(j = 1,\ldots,K)$.

As directly modelling the dependency of the $\lambda_j(x\ z)$ also seems very natural, it will first be checked whether both ways, modelling the dependency of the joint survival function and modelling the marginal failure rates, are equivalent ways. For the most general survival function this is not the case:

$$(2.25)\qquad \lambda_j(x) = \frac{-[\frac{\partial}{\partial x_j} S(x_1,\ldots,x_j,\ldots,x_K)]_{x_1 = \ldots = x_K = x}}{\bar{F}(x)},$$

so that knowledge of the $\lambda_j(x)$ $\bar{F}(x)$ does not allow for a solution of $S(x_1,\ldots,x_K)$. This is just an expression of the non-identifiability of S from data on failure time and failure cause; see David and Moeschberger (1978, Chapter 4) on this phenomenon. If the $X_1,\ldots,X_K$ are assumed to be independent r.v.'s, then both ways are equivalent:

$$(2.26)\qquad \lambda_j(x) = r_j(x) = \frac{-\frac{d}{dx}\bar{P}_j(x)}{\bar{P}_j(x)},$$

2) See also Prentice, Kalbfleisch et.al.(1978).

which for given $\lambda_j(x)$ may be solved for $\bar{P}_j(x)$ $(j = 1,\ldots,K)$, specifying $S = \bar{P}_1 \times \ldots \times \bar{P}_K$.

Evidently there are two convenient ways of introducing the dependency on the covariables z. The first way is to make assumptions about the $\lambda_j(x|z)$ $(j = 1,\ldots,K)$, without assuming anything about the S-function or even mentioning it. The second way is to assume $S = \bar{P}_1 \times \ldots \times \bar{P}_K$, i.e. independent r.v.'s $X_1,\ldots,X_K$, and subsequently also assume something about the marginal failure rates $-[d\ \bar{P}_j(x|z)\ /\ dx]/\bar{P}_j(x|z)$. For the resulting likelihood function these two ways of introducing z are indifferent. As in the previous section the theoretical r.v.'s $X_1,\ldots,X_K$ have been used to model the competing risk problem, the second way is the consistent one, and will be used here.

Assuming independent r.v.'s $X_1,\ldots,X_k$, the dependency will be modelled as follows:

$$(2.27)\qquad \lambda_j(x|z) = \exp(\beta_j' z)\ \lambda_{0j}(x),$$

where the β_j $(j = 1,\ldots,K)$ are column vectors of cause specific regression coefficients and $\lambda_{0j}(x)$ is the failure rate for $\beta_j = 0$, i.e. in the absence of any influence from the covariables z.

A special case of (2.27) is

$$(2.28)\qquad \lambda_j(x|z) = \exp(\beta_j' z)\ e^{\gamma_j}\ \lambda_0(x),$$

in which case the marginal failure rates λ_{0j} are proportional to each other. For a subject, with given z, the ratios of the λ_j are independent of x, which defines the proportional hazards model.[3] The probability π_j that a subject with covariables z ever fails of cause j, then becomes:

$$(2.29) \quad \pi_j = \frac{\lambda_j}{\sum_{i=1}^{K} \lambda_i} = \frac{\exp(\gamma_j + \beta_j' z)}{\sum_{i=1}^{K} \exp(\gamma_i + \beta_i' z)}$$

which defines a logistic function in z.

3. The Likelihood Function

3.1. Introduction

In this section, the likelihood function, initially for a cohort of subjects followed for some period of time, will be developed.
In Section 3.2, this is done for the case that failure rates and covariables influencing the failure rates may change continuously with time. In Section 3.3, this same case is treated by means of Cox's partial likelihood approach. Finally, in Section 3.4, the practical case of failure rates and covariables being constant within time intervals is considered. Section 3.4 forms the core of this paper.

3.2. The likelihood for continuously changing rates and covariables

Subject i of the failure-free cohort H is followed during the time interval $[0, t_i]$. By "time" is meant follow-up or study time rather than calendar time, so that although different subjects enter the follow-up study at the same study time 0, the calendar times of entering the study may be different. The only type of censoring present is of type I, i.e., the total follow-up time t_i of subject i is either predetermined or stochastic and it is stochastic only if one of the K failure causes produces time t_i. Correspondingly, two types of events are defined for each subject $i \in H$:

3) If several of the covariables are time dependent, so that $z = z(x)$, then a proportional hazards model requires all K cause specific regression coefficients associated with such a time dependent covariable to be the same.

(i) event M_0, occurring if the follow-up ends without failure at predetermined time t_i;

(ii) event M_j ($j = 1,\ldots,K$), occurring if the follow-up ends at stochastic time t_i, produced by failure cause j ($j = 1,\ldots,K$).

The set H is partitioned into the sets $M_0, M_1,\ldots,M_K$. Each subject $i \in H$ belongs to exactly one of these sets. This is also denoted by the zero-one indicators $\delta_{0i}, \delta_{1i},\ldots,\delta_{Ki}$, which are defined as $\delta_{ji} = 1$ if $i \in M_j$ ($j = 0,1,\ldots,K$) and as $\delta_{ji} = 0$ otherwise.

The likelihood function L_i for subject $i \in H$ can now be written as:

$$(3.1) \qquad L_i = [\overline{F}_i(t_i)]^{\delta_{0i}} \prod_{j=1}^{K} [\pi_{ji} \, f_{ji}(t_i)]^{\delta_{ji}} ,$$

as follows from (2.3) and (2.8). In (3.1) $\overline{F}_i(t_i)$ denotes, for subject i, the survival probability for predetermined time t_i and $\pi_{ji} \, f_{ji}(t_i)dt_i = \lambda_{ji}(t_i)\overline{F}_i(t_i)dt_i$ denotes the probability of failure from cause j at stochastic time t_i. The likelihood function for all subjects in H is, because of independence:

$$(3.2) \qquad L = \prod_{i \in H} L_i = \prod_{i \in M_0} \overline{F}_i(t_i) \prod_{j=1}^{K} \prod_{i \in M_j} \lambda_{ji}(t_i) \, \overline{F}_i(t_i) .$$

This can be expressed into the rates λ_{ji}, as follows:

$$(3.3) \qquad L = \prod_{j=1}^{K} \left\{ \prod_{i \in H} \exp[-\int_0^{t_i} \lambda_{ji}(x)dx] \prod_{i \in M_j} \lambda_{ji}(t_i) \right\} ,$$

which follows from (2.14) and (2.23). This implies that $\ln L$ is additively separable with respect to the causes $j = 1,\ldots,K$:

$$(3.4) \qquad \ln L = \sum_{j=1}^{K} \left\{ \sum_{i \in M_j} \ln [\lambda_{ji}(t_i)] - \sum_{i \in H} \int_0^{t_i} \lambda_{ji}(x)dx \right\} .$$

A separate maximization for each failure cause j maximizes the log likelihood (3.4).

By introducing the dependency model (2.27) for the λ_{ji}, with time dependent covariables $z_i(x)$, one introduces cause-specific coefficients β_j. Assuming knowledge of the $z_i(x)$, these β_j can then be estimated by maximizing (3.4) with respect to the β_j. Of course, the λ_{0j} have to be specified. The λ_{0j} need not be made dependent on the subjects i, as this dependency is accomplished by means of the (subject dependent) vector of covariables z_i.

In Section 3.4 the assumption of knowing the $z_i(x)$ functions is replaced by the assumption of knowing values of $z_i(x)$ at (predetermined) points in time, say $x_1, x_2, \ldots, x_p$. Another, somewhat more "heroic", assumption is to suppose that the λ_{ji} depend on one measurement of the $z_i(x)$, say at time x_1. Of course, for some covariables this is justified. For example, a subject's sex does not change that easily.

3.3. Cox's partial likelihood approach

But first, for more insight, it is instructive to use a partial likelihood for estimating the β_j. The partial likelihood approach considers the complete likelihood as factored into parts which are relevant to the estimation problem at hand and other parts; see Cox (1972 and 1975) and Prentice and Kalbfleisch (1979).

A partial likelihood is found by conditioning on certain relevant events. Suppose, there are m_j failures from cause j, assembled in the set M_j ($j = 1,\ldots,K$). The times of failure corresponding to these failures are such that $0 < t_1 < \ldots < t_{m_j}$. (For ease of notation, a second index, indicating cause of failure, is suppressed.) Denoting by H_i the set of persons free from failure just before time t_i ($i = 1,\ldots,m_j$), the probability that subject i fails from cause j at time t_i, conditional on the set H_i, is

$$(3.5) \qquad L_{ji} = \frac{\lambda_{ji}(t_i)}{\sum_{r \in H_i} \lambda_{jr}(t_i)} .$$

Introducing the dependency model (2.27), the log partial likelihood for all subjects of M_j becomes:

$$(3.6)\qquad \ln L_j(\beta_j) = \sum_{i=1}^{m_j} \ln L_{ji}(\beta_j)$$

$$= \beta_j' \sum_{i=1}^{m_j} z_i(t_i) - \sum_{i=1}^{m_j} \ln \left[\sum_{r\in H_i} \exp(\beta_j' z_r(t_i))\right] .$$

The notation $\ln L_j(\beta_j)$ stresses that (3.6) [and also (3.5)] only depends on β_j and not on λ_{0j}. This in contrast to the log likelihood (3.4).

The relation of (3.6) to (3.4) - actually its j-th sum - is, that (3.6) may be considered as a special "maximized" case of (3.4); see Breslow (1974). In order to show this, the total follow-up time is divided into m_j intervals $(t_{i-1}, t_i]$ $(i = 1, \ldots, m_j;\ t_0 = 0)$. Considering only failure cause j within interval $(t_{i-1}, t_i]$ and considering (not quite correctly) the "risk" set H_i as defined above, the log likelihood for each interval follows from (3.4) with only its j-th sum considered. Considering all the intervals, gives:

$$(3.7)\qquad \ln L_j = \sum_{i=1}^{m_j} \ln [\lambda_{ji}(t_i)] - \sum_{i=1}^{m_j} \sum_{r\in H_i} \int_{t_{i-1}}^{t_i} \lambda_{jr}(x)dx .$$

Now, introduce the dependency (2.27) and consider the λ_{0j} as constant within the intervals $(t_{i-1}, t_i]$. Writing these constants as $\exp(\alpha_{ji})$, - (3.7) must then be maximized with respect to the α_{ji} and β_j. The α_{ji} may be solved from the first order conditions as:

$$(3.8)\qquad \exp(\alpha_{ji}) = \left[(t_i - t_{i-1}) \sum_{r\in H_i} \exp(\beta_j' z_r(t_i))\right]^{-1} .$$

Substituting (3.8) into (3.7) gives, apart from a constant, the log partial likelihood (3.6).

When the number of subjects m_j is large, giving rise to many constants α_{ji}, this approach of using the log likelihood (3.6) for estimating the β_j and (3.8) for also estimating the λ_{0j} as a step function of time, may not be so efficient. By sufficiently parameterizing the functions $\lambda_{0j}(x)$ and treating them as continuous functions of time, one may increase the efficiency of the ML estimates by using the log likelihood (3.4). Two other arguments are in favour of the approach leading to (3.4). Firstly, this approach copes more easily with dependency models other than (2.27). And, secondly, it copes more easily with a subdivision of total follow-up time as determined, for instance, by the measurement times of the covariables, which is the subject of the next section.

3.4. Rates and covariables constant within time intervals

Although rates and covariables, when they change with time, in most cases change continuously, there will be little or no harm in supposing constancy in predetermined time intervals. This complies with the actual practice of follow-up studies in which the total follow-up time is divided into predetermined intervals, at the start of which a number of covariables is measured in subjects of the remaining failure-free cohort. This division of total follow-up time may be, and ideally is, done in such a way as to practically ensure the above mentioned constancy within intervals.

A reformulation of the log likelihood (3.4), incorporating the information from several observation points, requires double sums over subjects and time intervals. A simpler notation is achieved in the following way. The time intervals during which a given subject of the cohort is followed, are called subject-intervals. The set of subject-intervals for all subjects of H is denoted by H'. It is assumed that the indexes i, denoting subject-intervals for the same subject, form a set of consecutive integers. A subject-interval $i \in H'$ is denoted by $(t'_i, t'_i + w_i]$, with t'_i the starting point and w_i the width of the interval. The widths of subject-intervals are predetermined, possibly different, numbers determined by the observation points t'_i of the covariables. Only if the subject-interval considered is the last time interval for this subject and if failure occurs by one of the K failure causes, then the interval width is stochastic. Accordingly, the set H' is partitioned as follows:

M'_0: the set of subject-intervals with w_i predetermined;

M'_j $(j = 1,\dots,K)$: the set of subject-intervals with w_i the stochastic time of failure, as measured from the start of the interval.

As an example, let there be three subjects, two observation points at times 0 and 1, two failure causes and total follow-up interval [0,2]. The set H' may be as follows: intervals 1 and 2, being [0,1] and (1,2], for failure-free subject 1, hence $w_1 = w_2 = 1$ (predetermined); intervals 3 and 4, being [0,1] and $(1, 1 + w_4]$ for subject 2 failing from, say, cause 1, hence $w_3 = 1$ (predetermined) and w_4 is stochastic; interval 5, being $[0,w_5]$ for subject 3 failing from, say, cause 2, hence w_5 is stochastic. The set M'_0 then consists of the subject-intervals 1, 2 and 3, the set M'_1 of subject-interval 4 and the set M'_2 of subject-interval 5.

Assuming constancy of the failure rates λ_{ji}, with index j denoting cause j and index i subject-interval i, within the time-intervals of the follow-up study, the log likelihood becomes:

$$(3.9) \qquad \ell n\, L = \sum_{j=1}^{K} \{ \sum_{i \in M'_j} \ell n\, \lambda_{ji} - \sum_{i \in H'} w_i \lambda_{ji} \} .$$

For the proof of (3.9) the same steps as the ones leading from (3.1) to (3.4) may be followed. So, let the index i in (3.1) now represent a subject-interval from H'. Further, let all probabilities (and rates) in (3.1) now be conditional upon survival until the start of interval i. Time is measured from the start of the intervals.

The likelihood for all subject-intervals $i \in H'$ then is the product of the L_i, with the added motivation that for the same subject the likelihoods L_i of different intervals may be multiplied because L_i specifies a likelihood conditional upon survival until the start of the intervals. This leads to (3.4), with M_j and H replaced by M'_j and H', respectively, and with the integral having lower bound t'_i and upper bound $t'_i + w_i$. Taking natural logarithms gives (3.9).

The log likelihood (3.9) is not unique: there are other assumptions, than the one used in deriving it, that lead to this form. In the above a cohort is followed in time and each subject of the cohort contributes a number of intervals to H'. By, for example, following a cohort of subjects initially of age forty one observes the development of the relevant covariables and rates over time (= age). One obtains the same log likelihood if one samples the subjects by age and measures the relevant covariables just once. Each person contributes one interval, with its covariables, including possibly "age", to H'. What matters are the "information sets", the subject-intervals with their covariables and their type of endpoint, characterized by one of the failure causes or by no failure. (Of course, as is well known, this does not necessarily mean that the conclusions drawn within one context may be transplanted to another context.)

Assuming that to each subject-interval $i \in H'$ there belongs a value of z_i, the vector of covariables measured at the start of this interval for the relevant subject, the rates λ_{ji}, which are supposedly constant within the intervals, may be made dependent on z_i by means of the dependency model (2.27). In the previous subsection it was concluded that the time dependence of the λ_{0j} may be satisfactorily modelled by sufficiently parameterizing the λ_{0j} as functions of time. This may be accomplished by incorporating "time" as one of the covariables in the z_i vectors. Doing this, the dependency model (2.27) for subject-interval i is reformulated as:

$$(3.10) \qquad \lambda_{ji}(z_i) = \exp(\beta_{j0} + \beta'_j z_i) ,$$

where, moreover, one of the covariables has been explicitly stated as the constant 1.

Substituting (3.10) into (3.9) gives:

$$(3.11) \qquad \ell n\, L = \sum_{j=1}^{K} \{ m_j \beta_{j0} + \sum_{i \in M'_j} \beta'_j z_i - \exp(\beta_{j0}) \sum_{i \in H'} w_i \exp(\beta'_j z_i) \} ,$$

where m_j is again the number of subjects failing from cause j. This likelihood is to be maximized with respect to β_{j0} and β'_j , $j = 1,\ldots,K$. Details may be found in the appendix.

APPENDIX

Maximizing the log likelihood (3.11).

The log likelihood function (3.11) has to be maximized with respect to the vectors of coefficients (β_{j0}, β_j'), for given m_j, w_i and z_i $(j = 1,\ldots,K;\ i \in H')$. As the log likelihood function is separable with respect to j, its maximization can be performed separately and analogously for each failure cause j $(j = 1,\ldots,K)$. Therefore, the index j in (β_{j0}, β_j'), m_j and M_j' is suppressed in the following.

The ML-estimators for (β_0,β') follow from the condition of zero first order partial derivatives of (3.11) with respect to (β_0,β') :

(A.1a) $$\frac{\partial \ln L}{\partial \beta_0} = m - B \exp(\beta_0) = 0 ,$$

(A.1b) $$\frac{\partial \ln L}{\partial \beta} = \sum_{i \in M'} z_i - \frac{\partial B}{\partial \beta} \exp(\beta_0) = 0 ,$$

where

(A.2a) $$B = \sum_{i \in H'} w_i \exp(\beta' z_i) ,$$

(A.2b) $$\frac{\partial B}{\partial \beta} = \sum_{i \in H'} z_i w_i \exp(\beta' z_i) .$$

Solving $\exp(\beta_0)$ from (A.1a) as

(A.3a) $$\exp(\beta_0) = \frac{m}{B}$$

and substituting this into (A.1b) yields

(A.3b) $$\bar{z} - \frac{1}{\beta} \frac{\partial B}{\partial \beta} = 0 ,$$

where

(A.4) $$\bar{z} = \frac{1}{m} \sum_{i \in M'} z_i \ .$$

The estimate for β follows from numerically solving (A.3b) to β, whereafter the estimate for β_0 directly follows from (A.3a).

The estimated (asymptotic) covariance matrix of the ML-estimators for (β_0, β') is

(A.5) $$\begin{bmatrix} -\frac{\partial^2 \ln L}{\partial \beta_0^2} & -\frac{\partial^2 \ln L}{\partial \beta_0 \, \partial \beta'} \\ -\frac{\partial^2 \ln L}{\partial \beta_0 \, \partial \beta} & -\frac{\partial^2 \ln L}{\partial \beta \, \partial \beta'} \end{bmatrix}^{-1} = \frac{1}{m} \begin{bmatrix} 1 & \bar{z}' \\ \bar{z} & \frac{1}{B} \frac{\partial^2 B}{\partial \beta \, \partial \beta'} \end{bmatrix}^{-1} ,$$

where

(A.6) $$\frac{\partial^2 B}{\partial \beta \, \partial \beta'} = \sum_{i \in H'} z_i z_i' w_i \exp(\beta' z_i) \ .$$

Hence, the estimated (co)variances are inversely proportional to m, and systematically insensitive to the total number of subject-intervals $i \in H'$.

From (A.3b) it follows that the estimated β-vectors are systematically insensitive to m and also to the total number of subject-intervals $i \in H'$, while the estimated β_0 directly varies with the ratio of these numbers, see (A.3a).

REFERENCES

Breslow, N., "Covariance Analysis of Censored Survival Data", Biometrics, Vol. 30 (1974).

Cox, D.R., "Regression Models and Life-Tables", Journal of the Royal Statistical Society, Series B, Vol. 34 (1972).

Cox, D.R., "Partial Likelihood", Biometrika, Vol. 62 (1975).

David, H.A., and M.L. Moeschberger, The Theory of Competing Risks, Charles Griffin & Company Ltd., London (1978).

Lancaster, T. and S. Nickel, "The Analysis of Re-Employment Probabilities for the Unemployed", Journal of the Royal Statistical Society, Series A, Vol. 143 (1980).

Nickel, S., "Estimating the Probability of Leaving Unemployment", Econometrica, Vol. 47 (1979).

Prentice, R.L., and J.D. Kalbfleisch, "Hazard Rate Models with Covariates", Biometrics, Vol. 35 (1979).

Prentice, R.L.; J.D. Kalbfleisch et. al., "The Analysis of Failure Times in the Prensence of Competing Risks", Biometrics, Vol. 34 (1978).