Phase-Type Models and Their Extension to Competing Risks

Bo Henry Lindqvist and Susanne Hodneland Kjølen

Abstract We present an extension of the phase-type methodology for modeling of lifetime distributions to include the case of competing risks. This is done by considering finite state Markov chains in continuous time with more than one absorbing state, letting each absorbing state correspond to a particular risk. The special structure of Coxian phase-type models is considered in particular. The chapter emphasizes the use of phase-type models in statistical modeling and inference for survival and competing risks data.

Keywords Phase-type distribution · Coxian distribution · Competing risks · Identifiability

1 Introduction

Phase-type distributions represent the time to absorption for a finite state Markov chain in continuous time. The simplest examples are mixtures and convolutions of exponential distributions and phase-type distributions have therefore received much attention in applied probability, in particular in queuing theory. Here they generalize the celebrated Erlang distribution. Nowadays, phase-type distributions are applied in various areas such as reliability analysis and medical statistics.

In its generality, the class of phase-type distributions is both flexible and conceptually simple to work with. Interestingly, the class of phase-type distributions is dense in the sense that any lifetime distribution can be approximated arbitrarily close by a phase-type distribution. For a comprehensive introduction to the topic we refer to Neuts [16], while a shorter and very useful introduction is given by Aalen [1].

B.H. Lindqvist (✉)
Norwegian University of Science and Technology, Alfred Getz Vei 1,
7491 Trondheim, Norway
e-mail: bo.lindqvist@ntnu.no

S.H. Kjølen
Visma Software Labs, Karenlyst Allé 56, 0277 Oslo, Norway
e-mail: susanne.kjolen@gmail.com

A. Lisnianski et al. (eds.), *Recent Advances in Multi-state Systems Reliability*,
Springer Series in Reliability Engineering, DOI 10.1007/978-3-319-63423-4_6

The potential usefulness of phase-type distributions in statistical modeling and inference has more recently been revealed in the literature. In statistical applications there seems to be a particular interest in the use of so-called Coxian phase-type models, first suggested by Cox [8]. These are models for phenomena where the units go through stages (phases) in a specified order, and may transit to the absorbing state (corresponding to the event of interest) at any stage. Coxian phase-type models have recently been successfully applied in health care studies ([10, 11, 15, 18]). The problem of fitting more general phase-type distributions to lifetime data has also been considered in the literature, both in a frequentist setting using the EM-algorithm (Asmussen et al. [3]), and in a Bayesian setting using MCMC (Bladt et al. [5]).

The main purpose of the present chapter is to give the necessary tools and results in order to extend the phase-type methodology to include competing risks. The latter concept has been introduced for cases where one in addition to a lifetime have information about the specific cause of failure or death. The classical examples of competing risks consider individuals subjected to multiple causes of death. A famous example is due to David Bernoulli who around 1760 studied the problem of how to disentangle the risk of dying from smallpox from other causes. In cancer research one may consider both the age at onset of cancer and the cancer type. In reliability engineering, one may observe both the time to breakdown of a mechanical component and the root cause, for example vibration or corrosion. An introduction to the theory can be found in, e.g., Lawless [13, Chap. 9].

The basic ingredient in a competing risks phase-type model is a finite state Markov chain in continuous time with more than one absorbing state, where each absorbing state corresponds to a particular risk. Expressions for cause specific hazard functions, cumulative incidence functions etc. can now be given in terms of the transition matrix of the underlying Markov chain. Special structures like Coxian models may still be studied in the competing risks framework. Statistical inference for competing risks using phase-type models is of particular interest in the chapter. This extends approaches in the literature for ordinary phase-type models, and some basic aspects of this extension will be emphasized by studying simple examples involving Coxian models.

2 Phase-Type Distributions

A phase-type distribution can be described in terms of a Markov process $\{X(t); t \geq 0\}$, say, where the system moves through some or all of K transient states, or phases, before moving to a single absorbing state $K + 1$. The time of absorption, T, is then said to have a phase-type distribution. A simple illustration is given in Fig. 1, where $K = 7$ and state $K + 1 = 8$ is absorbing (state 9 will be considered later).

A Coxian phase-type distribution is obtained when all the transitions from the transient states are either from i to $i + 1$ or to the absorbing state $K + 1$, see Fig. 2. The resulting restriction on the permitted transitions is in fact not as strong as it may

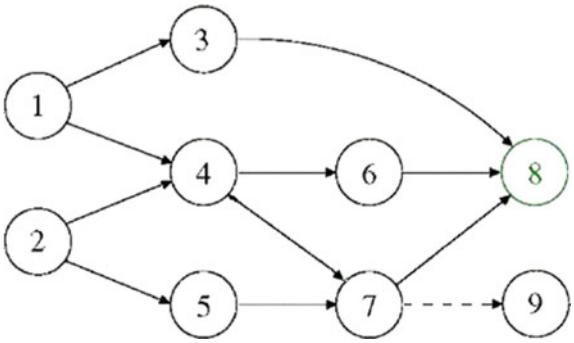

Fig. 1 The state-space and permitted transitions of an absorbing Markov chain, with absorbing state(s) 8 (8 and 9)

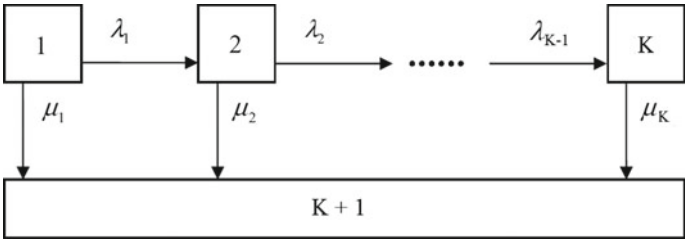

Fig. 2 A coxian phase-type model

look, since any phase-type distribution based on a Markov chain where each move is to a higher numbered state, can be brought on Coxian form (see, e.g., O'Cinneide [17]).

2.1 Model Specification

The infinitesimal transition matrix $\mathbf{A}$ of the Markov chain producing the phase-type distribution is a $(K+1)\times(K+1)$ matrix given on block form as

$$\mathbf{A} = \begin{bmatrix} \mathbf{Q} & \boldsymbol{\ell} \\ \mathbf{0} & 0 \end{bmatrix}. \tag{1}$$

Here $\mathbf{Q}$ is the $K \times K$ matrix corresponding to the transitions between the transient states; $\boldsymbol{\ell}$ is the $K \times 1$ vector defining direct transition intensities from the transient states to the absorbing state; while $\mathbf{0}$ is a $1 \times K$ vector of zeros. Letting $\mathbf{P}(t)$ be the matrix of transition probabilities $P_{ij}(t) = P(X(t) = j|X(0) = i)$ it is well known (e.g., Ross [19, Chap. 5]) that

$$\mathbf{P}(t) = e^{\mathbf{A}t} = \sum_{i=0}^{\infty} \mathbf{A}^i,$$

and it is then straightforward to show that (1) implies

$$\mathbf{P}(t) = \begin{bmatrix} e^{\mathbf{Q}t} & \mathbf{Q}^{-1}(e^{\mathbf{Q}t} - \mathbf{I})\boldsymbol{\ell} \\ \mathbf{0} & 1 \end{bmatrix}.$$

From this we obtain an expression for the cumulative distribution function of T,

$$F(t) = P(T \le t) = P(X(t) = K + 1) = \mathbf{p}\mathbf{Q}^{-1}(e^{\mathbf{Q}t} - \mathbf{I})\boldsymbol{\ell}.$$

Here $\mathbf{p}$ is the $1 \times K$-vector with entries $p_i = P(X(0) = i)$ for $i = 1, \ldots, K$, which defines the initial distribution of the Markov chain.

3 Classical Competing Risks

In survival analysis one basically considers the time to failure, T, of a unit. Suppose now that the unit can experience any one of k competing failure causes. Then for each unit one observes both the time to failure, T, and the cause of failure, $C \in \{1, 2, \ldots, k\}$. The pair (T, C) is the observation in the case of competing risks.

In the so called latent failure time approach to competing risks one assumes that the k causes are represented by potential failure times $T_1, \ldots, T_k$, where one only observes the smallest time, $T = \min_j T_j$ and its index $C = \arg\min_j T_j$.

3.1 Distributional Properties of Competing Risks

The joint distribution of the observed pair (T, C) is completely specified by the sub-distribution functions and their derivatives, the subdensities,

$$F_j(t) = P(T \le t, C = j), \quad f_j(t) = F_j'(t).$$

The interpretation of $F_j(t)$ is as the probability of failing from cause j before time t. In biostatistics literature, the $F_j(t)$ are also called cumulative incidence functions.

As an extension of the concept of hazard function of a lifetime distribution, one considers the cause-specific hazard functions,

$$\lambda_j(t) = \lim_{\Delta t \to 0} \frac{P(t < T \le t + \Delta t, C = j | T > t)}{\Delta t} = \frac{f_j(t)}{\bar{F}(t)}.$$

The interpretation is that $\lambda_j(t)$ is the failure rate from cause j conditional on survival up to time t.

3.2 The Identifiability Problem of Competing Risks

Consider here the latent failure time approach. The main interest is often in the joint and marginal distributions of the latent failure times $T_1, \ldots, T_k$. The classical problem of competing risks is, however, that the distribution of the observable pair (T, C) in general does not determine the distribution of the latent failure times. In fact, several different joint distributions of $T_1, \ldots, T_k$ will give rise to same distribution of (T, C). This *non-identifiability* property was noted by Cox [9] and formalized by Tsiatis [21]. The main result of Tsiatis is that for a given set of sub-distribution functions $F_j(t)$, there is always a unique (proxy) model with independent T_j yielding these $F_j(t)$.

Biostatisticians have for several decades abandoned the latent failure time approach and claim that statistical conclusions from data only should be based on observable (i.e., identifiable) quantities like the cumulative incidence functions and the cause-specific hazard functions.

4 Phase-Type Models for Competing Risks

As already indicated in Fig. 1, a Markov chain may have more than one absorbing state. In the figure, both states 8 and 9 are absorbing. If we let T be the time of absorption, and C be the identity of the absorbing state, then it is seen that (T, C) is of the form of the observation of a competing risks case.

More generally, consider the general setup of Sect. 2 where the Markov process $\{X(t); t \geq 0\}$ moves among the K transient states before it is absorbed in state $K + 1$. Suppose now instead that there are $m > 1$ absorbing states, named $K + 1, K + 2, \ldots, K + m$, say. Letting T be the time of absorption (in any one of the absorbing states), and letting the cause C represent the state where absorption occurs, by defining $C = K + j$ if $X(T) = K + j$; $j = 1, 2, \ldots, m$, the pair (T, C) can be viewed as an observation from a classical competing risks process with possible causes $K + 1, \ldots, K + m$.

The Coxian phase-type model can now in a straightforward manner be extended to the competing risks case by allowing transitions to any of the m absorbing states $K + 1, \ldots, K + m$ from each of the transient states. The case $m = 2$ is illustrated in Fig. 3.

4.1 Model Specification for Phase-Type Based Competing Risks

By extending the matrix (1) to encompass m absorbing states, we obtain the infinitesimal matrix of the modified Markov process to be the $(K + m) \times (K + m)$ matrix given on block form as

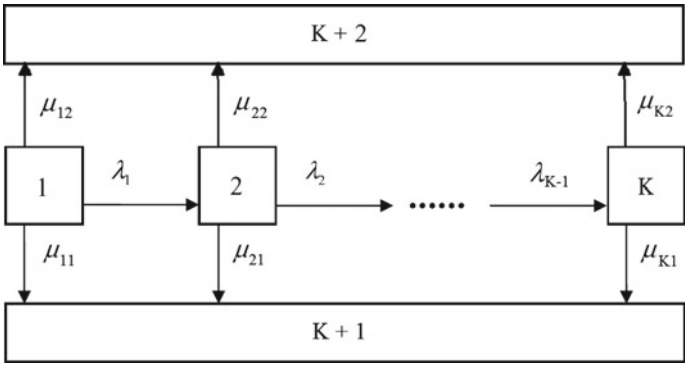

Fig. 3 A coxian phase-type model for $m = 2$ competing risks

$$\mathbf{A} = \begin{bmatrix} \mathbf{Q} & \mathbf{L} \\ \mathbf{0}_1 & \mathbf{0}_2 \end{bmatrix}. \tag{2}$$

As before, $\mathbf{Q}$ is the $K \times K$ matrix corresponding to the transitions between the transient states. The vector $\boldsymbol{\ell}$ is now replaced by the $K \times m$ matrix $\mathbf{L}$ which contains transition intensities from the transient states to the absorbing states. Further, $\mathbf{0}_1$ and $\mathbf{0}_2$ are, respectively, $m \times K$ and $m \times m$ matrices of zeros.

It is rather straightforward to show that (2) implies that the matrix of transition probabilities $P_{ij}(t)$ is given by

$$\mathbf{P}(t) = \begin{bmatrix} e^{\mathbf{Q}t} & \mathbf{Q}^{-1}(e^{\mathbf{Q}t} - \mathbf{I})\mathbf{L} \\ \mathbf{0}_1 & \mathbf{I} \end{bmatrix}, \tag{3}$$

where $\mathbf{I}$ is the $K \times K$ identity matrix. From (3) we obtain expressions for the subdistribution functions, given by

$$F_j(t) = P(T \le t, C = j) = P(X(t) = j) = \mathbf{p}\mathbf{Q}^{-1}(e^{\mathbf{Q}t} - \mathbf{I})\mathbf{L}\mathbf{v}_j \tag{4}$$

for $j = 1, \ldots, m$. By differentiation we get the subdensities

$$f_j(t) = F_j'(t) = \mathbf{p}\mathrm{e}^{\mathbf{Q}t}\mathbf{L}\mathbf{v}_\mathbf{j}. \tag{5}$$

In these formulas, $\mathbf{p}$ is the K-vector defining the initial distribution of the Markov chain. It is often natural to assume $p_1 = 1$. Further, $\mathbf{v}_j$ is the m-vector with jth element equal to 1 and the rest equal to 0.

Finally, the cause-specific hazard rate is given by

$$\lambda_j(t) = \lim_{\Delta t \to 0} \frac{P(T \le t + \Delta t, C = j | T > t)}{\Delta t} = \frac{F_j'(t)}{P(T > t)} = \frac{\mathbf{p}e^{\mathbf{Q}t}\mathbf{L}\mathbf{v}_j}{\mathbf{p}e^{\mathbf{Q}t}\mathbf{1}_K} \tag{6}$$

(e.g., Braarud [7]). Here $\mathbf{1}_K$ is a K-vector of all 1s.

5 Statistical Inference in Coxian Phase-Type Models

In the present and next section we consider the problem of statistical inference in phase-type models, for data containing survival times only as well as for competing risks data.

Suppose one has a sample of n independent units, where for the ith unit one observes the lifetime T_i and, if applicable, a cause of failure, C_i. The task is to fit phase-type models to the data, where Coxian models will be considered first.

In practice, some of the lifetimes may be censored. Most of the methods to be considered are able to handle censorings, but this problem will not be pursued in the following. This also applies to the inclusion of covariates in the data.

5.1 Coxian Survival Models

Faddy, Graves and Pettitt [10] and McGrory, Pettitt and Faddy [15] considered, respectively, maximum likelihood estimation and Bayesian estimation for Coxian models. In the latter article, the authors used a reversible jump MCMC in their analysis, thus including also K as a parameter in the model. The authors analyze an example dataset comprising lengths of hospital stays of a sample of patients collected from two Australian hospitals to produce a model for a patient's expected length of stay. In particular, posterior distributions for the number of phases and the regression parameters were produced.

In the former article, Faddy et al. [10] considered different variations of Coxian models (Fig. 2), in particular an interesting model assuming $\mu_1 = \mu_2 = \cdots = \mu_{K-3} = 0$ and $\lambda_1 = \lambda_2 = \cdots \lambda_{n-3} = \mu_{n-2} + \lambda_{n-2} = \lambda$. Such a structure corresponds to a gamma-distributed component from the first $n - 2$ phases, which makes the model more flexible.

Slud and Suntornchost [20] advocated the use of parametric models based on phase-type distributions with a low number, say 3–8, of parameters. A main conclusion of [20] is that simple phase type models can do almost as well as nonparametric methods, where the latter are commonly the preferred choices in biostatistics and partly in reliability analysis.

Motivated by the above mentioned articles and correspondng conclusions, we shall next show by examples how corresponding statistical analyses can be made with Coxian competing risks models.

5.2 Model 1: Coxian Competing Risks Model with $K = 2$ Transient States and $m = 2$ Absorbing States

This model is illustrated in Fig. 4. The corresponding infinitesimal intensity matrix is

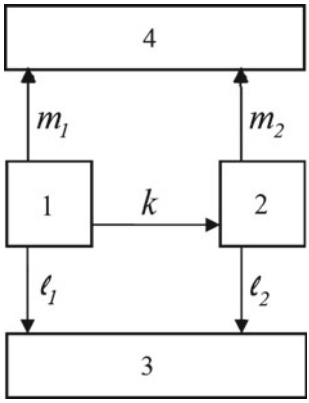

Fig. 4 Model 1

$$\mathbf{A} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} -a & k & l_1 & m_1 \\ 0 & -b & l_2 & m_2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array},$$

where $a = l_1 + m_1 + k$, $b = l_2 + m_2$. Hence

$$\mathbf{Q} = \begin{bmatrix} -a & k \\ 0 & -b \end{bmatrix}, \quad \mathbf{L} = \begin{bmatrix} l_1 & m_1 \\ l_2 & m_2 \end{bmatrix}.$$

The subdistribution functions, subdensities and cause-specific hazard functions are found from (4) to (6), and for cause 3 they are, respectively,

$$\begin{aligned} F_3(t) &= \frac{(1-e^{-at})\,l_1 - \frac{k\left(e^{-bt}-e^{-at}\right)l_2}{a-b}}{a} - \frac{k\left(e^{-bt}-1\right)l_2}{ab} \\ f_3(t) &= \left[l_1 - \frac{kl_2}{a-b}\right]e^{-at} + \frac{kl_2}{a-b}e^{-bt} \\ \lambda_3(t) &= \frac{l_1(a-b)e^{-at} + k\left(e^{-bt}-e^{-at}\right)l_2}{(a-b)e^{-at} + k\left(e^{-bt}-e^{-at}\right)}. \end{aligned} \tag{7}$$

Note that the formulas are valid only when $a \neq b$. The case $a = b$ follows by taking limits as $b \to a$ (see next subsection). The corresponding formulas for cause 4 are similar, replacing the l_j by m_j.

5.3 Parametric Identifiability of Model 1

Suppose we will use Model 1 in a competing risks case where the pair (T, C) is observed, where $C = 3$ and $C = 4$ represent the absorbing states. In order to estimate the five parameters of the model consistently, we will need the model to be identifiable. This problem will next be considered in detail below, presumably indicating the flavor of the problem also for larger models.

The general identifiability problem of competing risks has been described in Sect. 3.2. The problem is that the underlying probability mechanism is not necessarily identifiable from observations of the pair (T, C). For the present model, the question of identifiability is the following: Does the distribution of the pair (T, C) determine the five parameters of the model, k, l_1, l_2, m_1, m_2?

The functions $f_3(t)$ and $f_4(t)$ are from (7) necessarily given on the form

$$f_3(t) = A_3 e^{-\lambda_1 t} + B_3 e^{-\lambda_2 t}, \quad f_4(t) = A_4 e^{-\lambda_1 t} + B_4 e^{-\lambda_2 t}.$$

Knowing the distribution of (T, C) means that $\lambda_1, \lambda_2, A_3, B_3, A_4, B_4$ are known to us. We may then without loss of generality assume that $\lambda_1 < \lambda_2$ (the case when they are equal will be treated separately). To identify the parameters of the model, we need to consider two cases, $a < b$ and $a > b$.

Consider first the case $a < b$. By (7) we must have

$$a = \lambda_1, \quad b = \lambda_2.$$

Further,

$$l_1 - \frac{kl_2}{a-b} = A_3, \quad \frac{kl_2}{a-b} = B_3, \quad m_1 - \frac{km_2}{a-b} = A_4, \quad \frac{km_2}{a-b} = B_4$$

From this it is straightforward to show that the five parameters of the model are uniquely given by

$$\begin{aligned} l_1 &= A_3 + B_3, \qquad l_2 = \frac{(\lambda_1 - \lambda_2)B_3}{\lambda_1 - A_3 - B_3 - A_4 - B_4}, \\ m_1 &= A_4 + B_4, \qquad m_2 = \frac{(\lambda_1 - \lambda_2)B_4}{\lambda_1 - A_3 - B_3 - A_4 - B_4}, \\ k &= \lambda_1 - A_3 - B_3 - A_4 - B_4. \end{aligned}$$

Suppose then that $a > b$. Then obviously $a = \lambda_2$ and $b = \lambda_1$ and it can be shown that the five parameters are uniquely given by

Table 1 Two different sets of parameter values corresponding to the subdensities (8)–(9)

Assumption	k	l_1	m_1	l_2	m_2
$a < b$	1	2	1	3	2
$a > b$	2	2	1	5/2	3/2

$$l_1 = A_3 + B_3, \quad l_2 = \frac{(\lambda_2 - \lambda_1)A_3}{\lambda_2 - A_3 - B_3 - A_4 - B_4},$$
$$m_1 = A_4 + B_4, \quad m_2 = \frac{(\lambda_2 - \lambda_1)A_4}{\lambda_2 - A_3 - B_3 - A_4 - B_4},$$
$$k = \lambda_2 - A_3 - B_3 - A_4 - B_4.$$

As an example, suppose we have "observed" that

$$f_3(t) = 5e^{-4t} - 3e^{-5t}, \tag{8}$$
$$f_4(t) = 3e^{-4t} - 2e^{-5t}. \tag{9}$$

By using the above results we conclude that there are exactly two different sets of parameters that give the above functions $f_3(t)$ and $f_4(t)$, see Table 1.

If $a = b$ in Model 1, then it follows by taking the limit as $b \to a$ in (7) that

$$f_3(t) = l_1 e^{-at} + kl_2 te^{-at}, \quad f_4(t) = m_1 e^{-at} + km_2 te^{-at}.$$

The observed $f_2(t)$ and $f_3(t)$ are hence necessarily of the form

$$f_3(t) = C_3 e^{-\lambda t} + D_3 te^{-\lambda t}, \quad f_4(t) = C_4 e^{-\lambda t} + D_4 te^{-\lambda t}.$$

It follows immediately that $l_1 = C_3$ and $m_1 = C_4$ and from this that $l_2, m_2.k$ are uniquely given as well.

5.4 Identifiability of Coxian Phase-Type Models

It follows from the above that the parameters of Model 1 are identifiable in the case when $a = b$, but that, in the case where $a \neq b$, an additional assumption on the relative size of a and b has to be made as part of the prior specification of the model.

In a practical application of Model 1 we might assume that state 2 involves a more severe condition for the unit than state 1. As a result of this, the transition rates to the absorbing states are expected to be higher from state 2 than from state 1. In this case it might be reasonable to assume that $a < b$, in which case we have an identifiable model.

For Model 1, it is seen that the eigenvalues of Q determine the exponents of the exponentials in the expressions for the subdensity functions $f_3(t)$ and $f_4(t)$. More generally, considering the general Coxian competing risks model in Fig. 3, it is seen that the eigenvalues of $\mathbf{Q}$ are

$$\rho_i = \mu_{i1} + \mu_{i2} + \lambda_i$$

for $i = 1, 2, \ldots, K$, where by convention we put $\lambda_K = 0$. It follows that the subdensity function f_{K+1} is of the form

$$f_{K+1}(t) = \sum_{i=1}^{K} A_i e^{-\rho_i t}$$

provided the ρ_i are all different, and similarly for $f_{K+2}(t)$. If, say, ρ_i has multiplicity $r > 1$, then terms of the form $Ct^j e^{-\rho_i t}$ for $j = 1, 2, \ldots, r-1$ are included in the functions. Motivated by the study of Model 1, we may in general have to consider all permutations of the ρ_i in order to check identifiability.

We close the discussion on identifiability by noting that identifiability problems may also occur in ordinary Coxian models with a single absorbing state. To be explicit, consider Fig. 4 and assume that $m_1 = m_2 = 0$. Then $f_3(t)$ in (7) is simply the density function of the time T to absorption in state 3. We will now show by an example that two different parameterizations can give rise to the same density function. Namely, let (l_1, l_2, k) be given by either $(2, 4, 1)$ or $(2, 3, 2)$. In both cases we obtain the density function of T equal to $f(t) = 6e^{-3t} - 4e^{-4t}$. Hence the model is not identifiable from data on lifetimes T. Again, a prior choice of whether $l_1 + k$ is greater or smaller than l_2 has to be made.

5.5 *Case Study: Pneumonia on Admission to Intensive Care Unit ([4])*

Beyersmann, Allignol and Schuhmacher [4] present data for 747 patients at an intensive care unit, where the purpose is to examine the effect of hospital-acquired infections. The data set contains information on pneumonia status on admission, time of intensive care unit stay and 'intensive care unit outcome', either hospital death or alive discharge.

In order to increase flexibility compared to Model 1 we build on the earlier mentioned idea of Faddy et al. [10] by adding two states. The model is presented in Fig. 5, and we shall denote it by Model 2. We may think of the extended model as adding to the waiting time in state 3 a gamma-distributed length of time. Note that the approach of [10] would make the additional assumption that $l_1 + m_1 + k_1 = k_0$. In this case the total waiting time in state 3 is gamma-distributed.

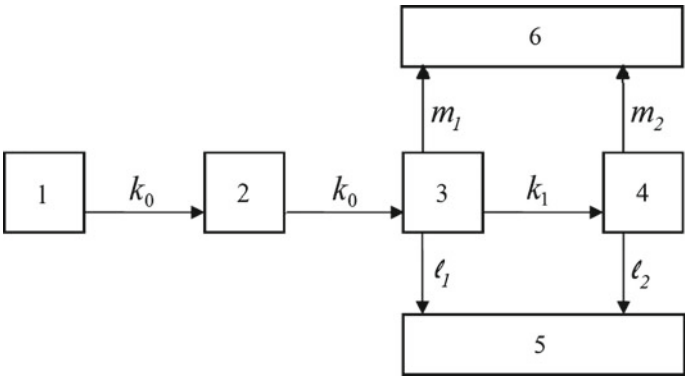

Fig. 5 Model 2

Table 2 Estimated parameters from Model 2 for the pneumonia data of [4]

	k_0	k_1	l_1	l_2	m_1	m_2
Pneumonia	0.079642	0.064075	0.102990	0.015845	0.479752	0.000569
No pneumonia	0.749985	0.072420	0.009248	0.008576	0.155579	0.047151

We made separate analyses for patients with and without pneumonia at admission. The absorbing states 5 and 6 correspond to, respectively, hospital death and discharge from hospital. The estimates of the parameters are given with one line for each analysis in Table 2.

In order to evaluate the results, we present in Fig. 6 plots of the cumulative incidence functions obtained from Model 2, together with nonparametric estimates found by using the Aalen-Johansen estimators (see, e.g., Borgan [6]). The parametric estimates seem to perform very well, a conclusion which confirms the findings and suggestions of [20] as reported earlier.

6 Statistical Inference for General Phase-Type Distributions

Asmussen, Nerman and Olsson [3] presented a general approach to estimation of phase-type distributions from lifetime data. Their idea was to consider the class of phase-type distributions, for a fixed K, as a multi-parameter exponential family. Since one then obviously is in the setting of incomplete observations, they suggested to implement the EM algorithm. Lindqvist [14] gave some details on how to extend the approach of [3] to the competing risks case.

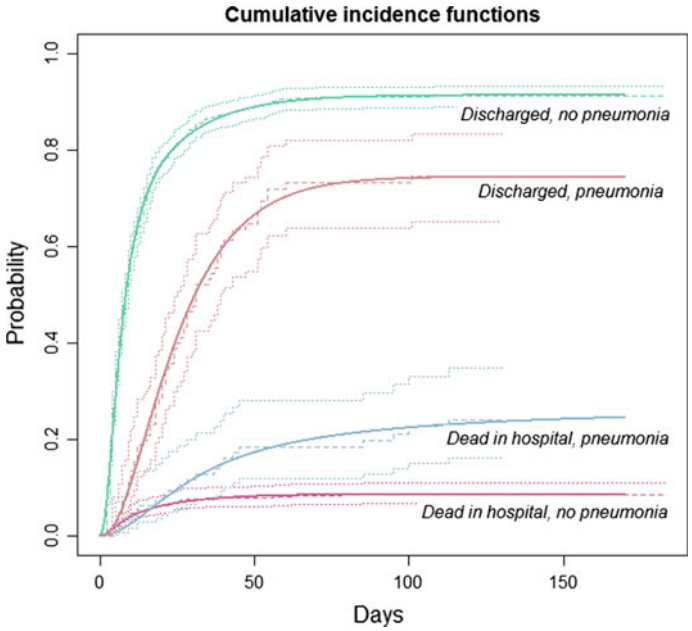

Fig. 6 Estimated cumulative incidence functions for the pneumonia-data of [4]. The smooth estimates are based on model 2 and are compared to nonparametric estimates with corresponding pointwise confidence intervals for each curve

Bladt, Gonzalez and Lauritzen [5] considered Bayesian estimation of phase-type distributions, constructing a Gibbs sampler which draws phase-type parameters from their posterior distribution. They reported as a main advantage of their method, that the uncertainty of estimates of complex functionals of the phase-type distributions could easily be obtained. It is not so clear, on the other hand, how to do this for the EM-algorithm approach. Aslett and Wilson [2] have improved on the method of Bladt et al. [5], and also provide an R-package for practical computation. The approach of Bladt et al. [5] was extended to the competing risks case by Laache [12].

In practice, lifetime data will typically include measured covariates. Most of the methods considered above can be extended to this case, and in some cases this is already a feature of the methods (e.g., [10] and [15]). Lindqvist [14] presented some ideas on a general approach for inclusion of covariates.

References

1. Aalen OO (1995) Phase type distributions in survival analysis. Scandinavian J Statistics 22(4):447–463
2. Aslett, LJM, Wilson SP (2011) Markov chain Monte Carlo for inference on phasetype models. ISI

3. Asmussen S, Nerman O, Olsson M (1996) Fitting phase-type distributions via the EM algorithm. Scandinavian J Statistics 23(4):419–441
4. Beyersmann J, Allignol A, Schumacher M (2011) Competing risks and multistate models with R. Springer Science & Business Media
5. Bladt M, Gonzalez A, Lauritzen SL (2003) The estimation of phase-type related functionals using Markov chain Monte Carlo methods. Scandinavian Actuarial J 2003(4):280–300
6. Borgan O (1998) Three contributions to the encyclopedia of biostatistics: The Nelson-Aalen. Kaplan-Meier and Aalen-Johansen estimators, Encyclopedia of Biostatistics
7. Braarud IH (2012) Phase type and competing risk models in survival analysis. Department of Mathematical Sciences, Norwegian University of Science and Technology, Project work
8. Cox DR (1955) A use of complex probabilities in the theory of stochastic processes. In: Mathematical proceedings of the Cambridge philosophical society vol 51, No 02. Cambridge University Press, pp 313–319
9. Cox DR (1959) The analysis of exponentially distributed life-times with two types of failure. J Royal Stat Soc Series B (Methodological) 21:411–421
10. Faddy MJ, Graves N, Pettitt AN (2009) Modeling length of stay in hospital and other right skewed data: Comparison of phase-type, gamma and lognormal distributions. Value Health 12:309–314
11. Gillespie J, McClean S, Scotney B, Garg L, Barton M, Fullerton K (2011) Costing hospital resources for stroke patients using phase-type models. Health Care Manag Sci 14(3):279–291
12. Laache CH (2014) Phase-type inference on competing risks models with covariates, using MCMC methods. Master's thesis, Department of Mathematical Sciences, Norwegian University of Science and Technology
13. Lawless JF (2003) Statistical models and methods for lifetime data. Wiley-Interscience, Hoboken, NJ
14. Lindqvist, B. (2016). Phase-type models for competing risks. In: Frenkel I, Lisnianski A (eds) Proceedings of the second international symposium on stochastic models in reliability engineering, life science and operations management (SMRLO16). Beer Sheva, Israel, pp 37–40, 15–18 February 2016, IEEE CPS, 978-1-4673-9941-8/16. doi:10.1109/SMRLO.2016.17
15. McGrory CA, Pettitt AN, Faddy MJ (2009) A fully Bayesian approach to inference for Coxian phase-type distributions with covariate dependent mean. Comput Stat Data Anal 53(12):4311–4321
16. Neuts MF (1981) Matrix-geometric solutions in stochastic models. Johns Hopkins University Press, Baltimore/London
17. O'Cinneide CA (1989) On non-uniqueness of representations of phase type distributions. Commun Stat—Stoch Models 5:247–259
18. Payne K, Marshall AH, Cairns KJ (2011) Investigating the efficiency of fitting Coxian phase-type distributions to health care data. IMA J Manag Math. doi:10.1093/imaman/dpr008
19. Ross SM (1983) Stochastic processes. Wiley, New York
20. Slud EV, Suntornchost J (2014) Parametric survival densities from phase-type models. Lifetime Data Anal 20(3):459–480
21. Tsiatis A (1975) A nonidentifiability aspect of the problem of competing risks. Proc Natl Acad Sci 72(1):20–22