Phase-Type Distributions for Competing Risks

Bo Henry Lindqvist
Department of Mathematical Sciences
Norwegian University of Science and Technology
Trondheim

https://www.ntnu.edu/employees/bo.lindqvist
bo.lindqvist@ntnu.no

ENBIS Spring Meeting
Grenoble, 19-20 May 2022

Contents of Talk

- Introduction to phase-type distributions
  - Coxian distributions
  - Extension to include competing risks
- Modeling degradation and maintenance
- Some theory for phase-type distributions
- Estimation and a numerical example

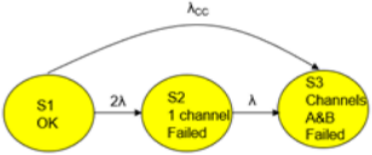
Main reference:

BHL: Phase-type models for competing risks, with emphasis on identifiability issues. *Lifetime Data Analysis 2022*.

Phase-type distributions represent the **time to absorption** for a finite state Markov chain $\{X(t); t \geq 0\}$ that moves through some or all of m transient states ("phases"), numbered $1, 2, \ldots, m$ before ending in a single absorbing state $m + 1$.

Simple examples:

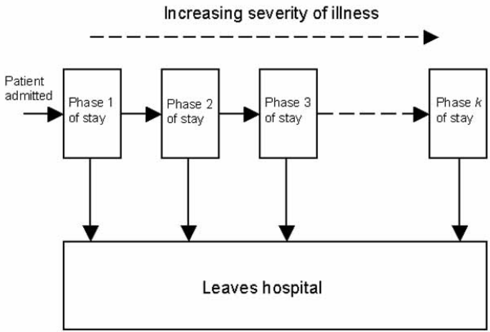
- Exponential distributions and their mixtures and convolutions distributions, Erlang distributions
- *Reliability theory:* failure time of systems or components



Coxian phase-type models

A **Coxian** phase-type distribution is obtained when all the transitions from the transient states are either to the next numbered transient state or to the absorbing state.

- Coxian phase-type models are popular in health studies, modeling *length of hospital stay*, going through "phases".
- The authors claim the superiority of Coxian phase-type models over common parametric models like *gamma* and *lognormal*.



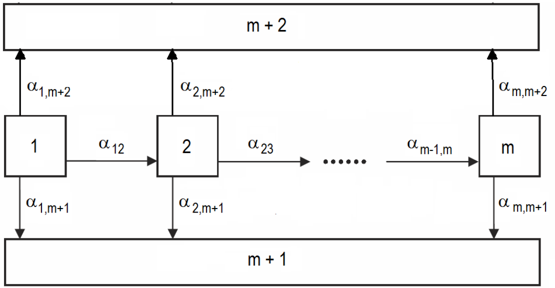
Phase-Type Modeling with multiple absorbing states

Let the Markov chain have m transient states and $K > 1$ absorbing states numbered $m+1, \ldots, m+K$. Define

$$T = \textit{absorption time}$$
$$C = \textit{type of event} \quad (\text{where } C = j \Leftrightarrow \text{absorption in } m+j)$$

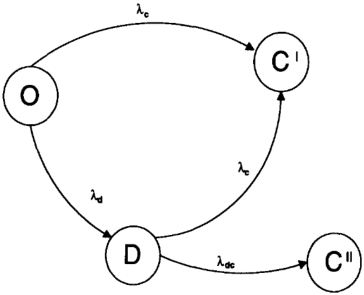
The pair (T, C) is then an observation from a standard **competing risks** model in survival analysis.

An extended *Coxian* model with $K = 2$:



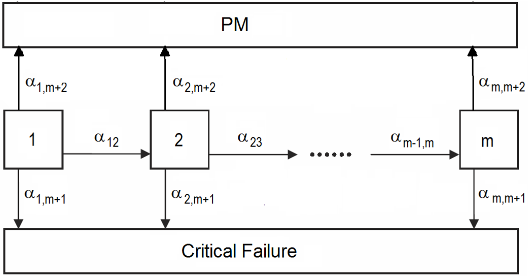
Example: Modeling of degraded and critical failures

Hokstad and Frøvig (RESS 1996): *The modelling of degraded and critical failures for components with dormant failures.*



- $m = K = 2$
- D denotes a *degradation* state
- C' denotes a *crtical failure* due to *shock*
- C" denotes a *critical failure* due to *degradation*.

General Coxian model for degradation and PM



- One may include preventive maintenance (PM) as a risk, in addition to one or more *failure* states.
- The transient states $1, 2, \ldots, m$ then correspond to different levels of degradation.

Some theory for phase-type distributions

The infinitesimal transition matrix A that produces the phase-type distribution is

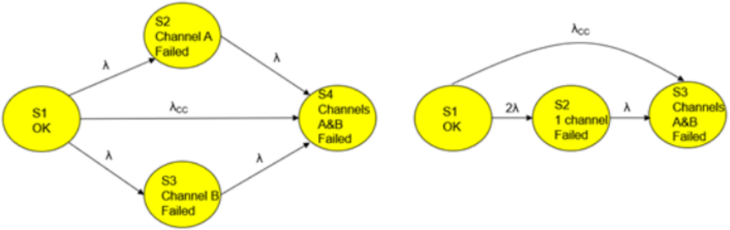
$$\mathrm{A} = \begin{bmatrix} \mathrm{Q} & \boldsymbol{\ell} \\ 0' & 0 \end{bmatrix}.$$

- Q is the $m \times m$ matrix corresponding to transitions between the transient states
- $\boldsymbol{\ell}$ is the m-vector defining direct transition intensities from the transient states to the absorbing state.

Finally, let p be the initial distribution with entries $p_i = P(X(0) = i)$ for $i = 1, \ldots, m$.

The pair (p, Q) is now a **representation** for the phase-type distribution.

Representations (p, Q) of phase-type distributions are not unique.



Two ways to represent the failure time of a two channel safety system using a Markov model (from EZ Blogs).

- The *order* of a phase-type distribution is the minimal number m of transient states of its representations.
- A phase-type distribution of order m can as a main rule be described by $2m - 1$ independent parameters (*while apparently* (p, Q) *would need* $m^2 + m - 1$ *entries...*)

Unique representation of phase type distributions is via Laplace transforms

The Laplace transform for the representation (p, Q) for T is

$$f^*(s) = E\left(e^{-sT}\right) = \mathrm{p}'(s\mathrm{I} - \mathrm{Q})^{-1}(-\mathrm{Q})\mathbf{1}$$

- This is a rational function of s, i.e., of the form $f^*(s) = N(s)/D(s)$ for polynomials $N(s)$ and $D(s)$
- **Remarkable results** by O'Cinneide (1990): *A distribution on $(0, \infty)$ is a phase-type distribution if and only if it has a* **strictly positive continuous density** *and a* **rational Laplace transform** *with a unique pole of maximal real part.* Moreover, **if all poles are** *real*, then the distribution can be represented as a Coxian distribution.

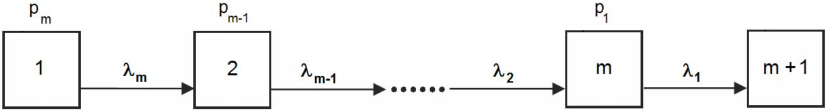
Canonical representation of upper triangular "Q"

Cumani (Microelectron Reliab 1982) showed that any phase-type distribution with *upper triangular* "Q" can be represented uniquely as a Coxian distribution (p, Q) where $\mathrm{p} = (1, 0, \ldots, 0)$ and

$$\mathrm{Q} = \begin{pmatrix} -\lambda_1 & \alpha_{12} & 0 & \cdots & 0 \\ 0 & -\lambda_2 & \alpha_{23} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -\lambda_m \end{pmatrix} \quad (1)$$

with $\lambda_1 \geq \lambda_2 \geq \ldots, \geq \lambda_m$.

Moreover, the following is another equivalent representation by Cumani (1982). Here the λ_i are as above and the probabilities above the squares are the probabilities of starting in that state.



Phase-Type Modeling of Competing Risks

Let A be the infinitesimal matrix of the Markov chain,

$$\mathrm{A} = \begin{bmatrix} \mathrm{Q} & \mathrm{L} \\ 0 & 0 \end{bmatrix}.$$

Then L defines transition intensities from transient to absorbing states.

The triple $(\mathrm{p}, \mathrm{Q}, \mathrm{L})$ represents the *phase-type competing risks* distribution. It has $(K+1)m - 1$ parameters.

Subdistribution function ($C = j \Leftrightarrow$ absorption in $m + j$):

$$F_j(t) = P(T \leq t, C = j) = P(X(t) = m + j) = \mathrm{p}\mathrm{Q}^{-1}(e^{\mathrm{Q}t} - \mathrm{I})\boldsymbol{\ell}_j$$

Here $\boldsymbol{\ell}_j$ is the jth column in L.

Cause-specific hazard function:

$$\lambda_j(t) = \lim_{\Delta t \to 0} \frac{P(T \leq t + \Delta t, C = j | T > t)}{\Delta t} = \frac{F_j'(t)}{P(T > t)} = \frac{\mathrm{p}e^{\mathrm{Q}t}\boldsymbol{\ell}_j}{\mathrm{p}e^{\mathrm{Q}t}\mathbf{1}}$$

Are phase-type competing risks models unique?

Consider the two Coxian models with $p^{(a)} = p^{(b)} = (1,0)'$ and

$$Q^{(a)} = \begin{pmatrix} -4 & 1 \\ 0 & -5 \end{pmatrix}, \quad L^{(a)} = \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}$$

and

$$Q^{(b)} = \begin{pmatrix} -5 & 2 \\ 0 & -4 \end{pmatrix}, \quad L^{(b)} = \begin{pmatrix} 2 & 1 \\ 5/2 & 3/2 \end{pmatrix}$$

It turns out that these representations lead to the same subdistribution functions $F_j^{(a)}(t)$ and $F_j^{(b)}(t)$, and hence the same joint distribution for (T, C).

CONSEQUENCE FOR STATISTICAL INFERENCE: Parameter estimates from maximum likelihood estimation may converge to different sets of values for the parameters.

How to ensure identifiability of Coxian representations

THEOREM (Rizk et al. arXiv 2019, BHL LIDA 2022)

Consider two non-redundant Coxian phase-type distributions for competing risks, given by $(\mathrm{p}^{(a)}, \mathrm{Q}^{(a)}, \mathrm{L}^{(a)})$ *and* $(\mathrm{p}^{(b)}, \mathrm{Q}^{(b)}, \mathrm{L}^{(b)})$*, where* $\mathrm{p}^{(a)} = \mathrm{p}^{(b)} = (1, 0, \ldots, 0)$*.*

Assume further that the diagonals of $\mathrm{Q}^{(a)}$ *and* $\mathrm{Q}^{(b)}$ *are ordered in the same way.*

Then if $F_j^{(a)}(t) = F_j^{(b)}(t)$ *for all* t, j *we have*

$$\begin{aligned} \mathrm{Q}^{(b)} &= \mathrm{Q}^{(a)} \\ \mathrm{L}^{(b)} &= \mathrm{L}^{(a)} \end{aligned}$$

Example on previous slide: We would have uniqueness if order of the diagonal elements was decided apriori.

A general identifiability result for competing risks

THEOREM (Telek and Horvath 2007, BHL LIDA 2022)

Let $(\mathrm{p}^{(a)}, \mathrm{Q}^{(a)}, \mathrm{L}^{(a)})$ *and* $(\mathrm{p}^{(b)}, \mathrm{Q}^{(b)}, \mathrm{L}^{(b)})$ *be two nonredundant phase-type representations for competing risks, having subdistribution functions* $F_j^{(a)}(t)$ *and* $F_j^{(b)}(t)$*, respectively.*

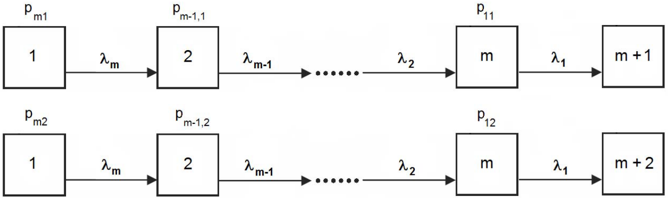
Then $F_j^{(a)}(t) = F_j^{(b)}(t)$ *for all* t *and* j *if and only if there exists a nonsingular* $m \times m$ *matrix* B *with* $\mathrm{B}1 = 1$ *such that* $\mathrm{p}^{(b)\prime} = \mathrm{p}^{(a)\prime}\mathrm{B}$, $\mathrm{Q}^{(b)} = \mathrm{B}^{-1}\mathrm{Q}^{(a)}\mathrm{B}$ *and* $\mathrm{L}^{(b)} = \mathrm{B}^{-1}\mathrm{L}^{(a)}$.

Remark: In non-uniqueness example can equivalence of subdistribution functions be verified by using

$$\mathrm{B} = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}.$$

Canonical model for Coxian competing risks (Rizk et al. Stat Med 2021, BHL LIDA 2022)

By extending Cumani's canonical model we obtain the following picture for a unique *canonical model for phase-type competing risks* with $K = 2$ where original "Q" is upper triangular. Here, the p_{ij} are the probabilities of starting in one of the $2m$ states of the picture, where $\sum_{i=1}^{2} \sum_{k=1}^{K} p_{ij} = 1$. Also, $\lambda_1 \geq \lambda_2 \geq \ldots, \geq \lambda_m$.

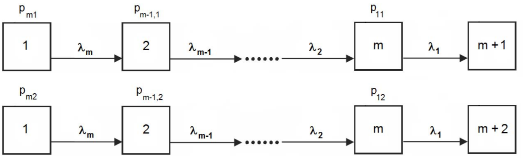


Generalization to general K is essentially straightforward.

Estimation in canonical model using EM algorithm

The EM-algorithm is often useful for maximum likelihood estimaton in phase-type models (e.g. Asmussen et al. Scand J Stat 1996), letting the *latent* observations be the set of states visited by the Markov chain.

The canonical representation of Cumani allows a simpler EM-algorithm:



- The *latent* observation for an item is the starting state.
- The E-step calculates the expected starting state given the data (T, C).
- The M-step is straightforward, but involves ordered λ_i.

Example: VHF data (Mendenhall and Hader, 1958)

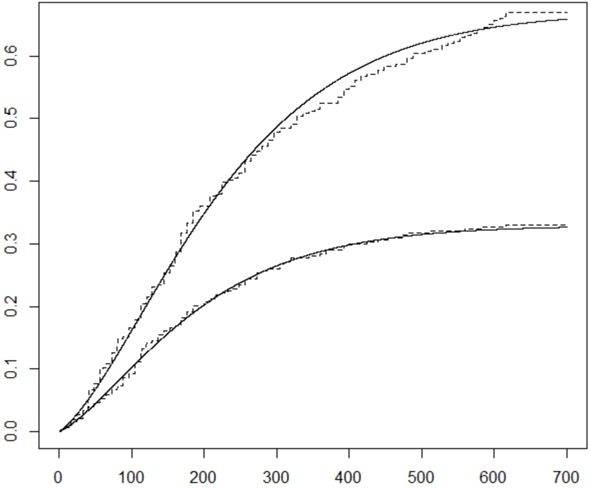
Times to failure for ARC-1 VHF communication transmitter-receivers of a single commercial airline.

Competing risks:

- Confirmed failure (critical failure, 218 cases)
- Unconfirmed failure (PM, 107 cases)

Estimated Coxian competing risks model with $m = 3$, $K = 2$ by EM algorithm.

Example: VHF data (Mendenhall and Hader, 1958)



Estimated subdistribution functions for VHF data.

Upper: Critical failure. Lower: PM. Solid curves: ML estimates from phase-type model. Dashed curves: Aalen-Johansen estimates.

Concluding remarks

- The main object has been to show how phase-type methodology can can be extended to cover competing risks modeling in survival analysis.
- This is part of various other new theoretical developments and extensions of the classical models. For example using nonhomogenous Markov chains, semi-Markov models, and models including unobserved heterogeneity
- In the presented work, special interest has been in identifiability issues. A basic problem is to characterize representations $(\mathrm{p}^{(a)}, \mathrm{Q}^{(a)}, \mathrm{L}^{(a)})$ and $(\mathrm{p}^{(b)}, \mathrm{Q}^{(b)}, \mathrm{L}^{(b)})$ which lead to the same phase-type competing risks distribution.
- For this, it has been necessary to go deeper into some of the theoretical properties underlying the phase-type methodology. The key literature here is a series of papers by O'Cinneide from around 1990 and the intriguing reliability motivated paper by Cumani 1982.