Phase type distributions for competing risks

Bo Henry Lindqvist*
Norwegian University of Science and Technology, Trondheim, Norway bo@math.ntnu.no

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. The class of phase-type distributions is both flexible and conceptually simple to work with. Further, it is essentially no loss of generality to work with phase-type distributions, since the class of phase-type distributions (with the number of states taking any finite value) is dense. Hence any lifetime distribution can, at least in principle, be approximated arbitrarily close by a phase-type distribution. Multivariate phase-type distributions can likewise be defined. They possess many of the properties of the univariate phase-type distributions. Phase-type distributions have received much attention in applied probability, in particular queuing theory. Here they generalize the celebrated Erlang distribution. Phase-type distributions are also much used in reliability theory and also in medical statistics. The problem of fitting phase-type distributions to lifetime data has been considered by several authors. One approach is via the EM algorithm, with the possibility of having right-censored or interval-censored observations. There are also Bayesian approaches in the literature, based on Markov chain Monte Carlo methods.

In the present paper we extend the phase-type methodology to the case of competing risks. 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. Of particular interest is model estimation from (possibly censored) competing risks data. It will in the paper be shown how estimation via the EM algorithm, known from the literature for ordinary phase-type models, can be performed also in the case of several absorbing states. Likewise, methods for Bayesian estimation may be extended. An attempt to treat covariates in the competing risks data will be discussed.

Key Words: Markov chain, censored data, EM algorithm, Markov chain Monte Carlo