Bulletin of the *Transilvania* University of Braşov • Vol 3(52) - 2010
Series III: Mathematics, Informatics, Physics, 67-70

COMPETING RISKS AS PURGED MARKOV CHAINS

Mioara A. NICOLAIE[1,2]

Abstract

We present the competing risks model of Larson and Dinse ([4]) as a special case of the purged Markov chain.

2000 *Mathematics Subject Classification:* 62N01, 62N02.
Key words: competing risks, mixture model.

1 Introduction

Competing risks is a subdiscipline of survival analysis where, in addition to the time-to-event T, we observe the cause of failure $D \in \{1, \dots, J\}$. Interest focuses on the joint distribution of (T, D), denoted by $\mathcal{P}(T \le t, D = j)$ (see [1]). We shall discuss competing risks within the framework of Markov chain models in the spirit of [2] and introduce a nonparametric method in the context of the Larson and Dinse approach ([4]).

2 The Markov model

2.1 Notation

Let $(X_t)_{t\ge 0}$ be a nonhomogeneous Markov chain with one transient state "0 (alive)" and J absorbing states, $j = 1, \dots, J$ corresponding to "failure due to cause j". Here we call $(X_t)_{t\ge 0}$ a competing risks model. Denote by $S = \{0, 1, \dots, J\}$ the state space and by $\overline{P}(t) = (\overline{P}_{ij}(t))_{i,j\in S}$, $t \ge 0$, the transition matrix. Denote by

$$\overline{P}_{ij}(s, t) = \mathcal{P}(X(t) = j|X(s) = i),$$

the conditional probability of state j at time t given state i at time s, for $0 \le s < t \le \tau$. The transition intensities or the cause-specific hazards $\lambda_j(t)$ from state 0 to state j, $j \in \{1, \dots, J\}$ are

$$\lambda_j(t) = \lim_{\Delta t\to 0} \frac{\mathcal{P}(t \le T < t + \Delta t, D = j|T \ge t)}{\Delta t},\ j = 1, \dots, J. \tag{1}$$

[1]Faculty of Mathematics and Computer Science, *Transilvania* University of Braşov, Braşov – 500091, Romania, e-mail: alinanicolae@unitbv.ro
[2]Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, PO Box 9600, 2300 RC, Leiden, The Netherlands

The stochastic behaviour of $(X_t)_{t\geq 0}$ is completely determined by the $\lambda_j(t)$, $j = 1, \ldots, J$. In particular, the survival function $\overline{P}_{00}(0,t) = S(t) = \mathcal{P}(T > t)$ is given by

$$\overline{P}_{00}(0,t) = \exp(-\sum_{j=1}^{J} \int_0^t \lambda_j(s)ds), \tag{2}$$

and the cumulative incidence function $\overline{P}_{0j}(0,t) = \mathcal{P}(T \leq t, D = j)$ is given by

$$\overline{P}_{0j}(0,t) = \int_0^t \lambda_j(s)S(s-)ds, \tag{3}$$

for $j = 1, \ldots, J$. From estimates of the cause-specific hazards the transition probabilities $\overline{P}_{0j}(0,t)$, $j = 0, 1, \ldots, J$, can be estimated as plug-in estimates using (2) and (3).

2.2 The purged chain

Consider the random event

$$\mathcal{A}_j = \{\text{absorbtion takes place in state j}\} = \{X(\nu) = j\},\ \nu \geq 0, \tag{4}$$

where j is an absorbing state. We are interested in computing the conditional probability of the event $\{T > t\}$ given $\mathcal{A}_j$.

Proposition 1. *The sequence $(X_t)_{t\geq 0}$ is a Markov chain with respect to the conditional probability $\mathcal{P}_{\mathcal{A}_j}$, with state space $\{0, j\}$ (the only absorbing state being j).*

Proof. See [3], pp. 109-110.

□

Definition 1. *(see [2])*
We call $(X_t)_{t\geq 0}$ under the conditional probability $\mathcal{P}_{\mathcal{A}_j}$ a purged chain corresponding to the original Markov chain $(X_t)_{t\geq 0}$.

Denote by $h_j(t)$ the corresponding transition intensity from state 0 to state j and by $P(t) = (P_{kl}(t))_{k,l\in\{0,j\}}$ the corresponding transition matrix in the purged chain.

The next theorem describes the relationship between the original Markov chain and the purged chain.

Theorem 1. *In the above notation, we have*

$$h_j(t) = \frac{\alpha_j(t)}{\overline{P}_{0j}(0,\tau) + \overline{P}_{00}(0,\tau)},\ 0 \leq t \leq \tau. \tag{5}$$

Moreover, for $0 \leq s < t \leq \tau$ we have

$$P_{00}(s,t) = \overline{P}_{00}(s,t)\frac{\overline{P}_{00}(t,\tau) + \overline{P}_{0j}(t,\tau)}{\overline{P}_{00}(s,\tau) + \overline{P}_{0j}(s,\tau)} \tag{6}$$

and

$$P_{0j}(s,t) = \frac{\overline{P}_{0j}(s,t)}{\overline{P}_{00}(s,\tau) + \overline{P}_{0j}(s,\tau)}. \tag{7}$$

Proof. We will prove (5) first. Starting from the definition of the intensity of the transition from 0 to j in the purged chain, we get

$$\begin{aligned}
h_j(t) &= \lim_{\Delta t\to 0} \frac{\mathcal{P}_{\mathcal{A}_j}(t \le T < t+\Delta t, D=j|T\ge t)}{\Delta t} \\
&= \lim_{\Delta t\to 0} \frac{1}{\Delta t}\frac{\mathcal{P}(\mathcal{A}_j \cap \{t\le T < t+\Delta_t, D=j\}\cap\{T\ge t\})}{\mathcal{P}(\mathcal{A}_j\cap\{T\ge t\})} \\
&= \lim_{\Delta t\to 0} \frac{1}{\Delta t}\frac{\mathcal{P}(\mathcal{A}_j \cap \{X(t)=j,\ t\le T < t+\Delta_t\}\cap\{X(t)=0,\ T\ge t\})}{\mathcal{P}(\mathcal{A}_j\cap\{X(t)=0,\ T\ge t\})}.
\end{aligned}$$

Note that the event $\{X(t)=0,\ T\ge t\}$ refers to the paths starting in state 0 which remain in this state by time t in the original chain (absorbtion might take place later). We gather them in $B_0=\{X(u)=0, 0\le u\le t\}$. We have

$$\begin{aligned}
&\mathcal{P}(\mathcal{A}_j \cap \{X(t)=j,\ t\le T<t+\Delta_t\}\cap B_0) \\
&= \mathcal{P}(B_0)\cdot\mathcal{P}(X(t)=j,\ t\le T<t+\Delta_t|B_0)\cdot\mathcal{P}(X(\nu)=j|\{X(t)=j,\ t\le T<t+\Delta_t\}\cap B_0) \\
&= \mathcal{P}(B_0)\cdot\mathcal{P}(X(t)=j,\ t\le T<t+\Delta_t|B_0).
\end{aligned}$$

On the other hand,

$$\begin{aligned}
&\mathcal{P}(\mathcal{A}_j\cap\{X(t)=0,\ T\ge t\}) = \mathcal{P}(\{X(\nu)=j\}\cap\{X(t)=0,\ T\ge t\}) \\
&= \mathcal{P}(\{X(\nu)=j\}\cap(\{\nu\le\tau\}\cup\{\nu>\tau\})\cap B_0) \\
&= \mathcal{P}(\{X(\nu)=j\}\cap\{\nu\le\tau\}\cap B_0)+\mathcal{P}(\{X(\nu)=j\}\cap\{\nu>\tau\}\cap B_0) \\
&= \mathcal{P}(B_0)\cdot\sum_{l\ge 0, t+l\le\tau}\mathcal{P}(\{X(t+l)=j\}\cap\{X(u)=0, t\le u\le t+l\}|B_0) \\
&\quad +\mathcal{P}(B_0)\cdot\sum_{l\ge 0, t+l\le\tau}\mathcal{P}(X(u)=0, t\le t+l|B_0) \\
&= \mathcal{P}(B_0)\cdot\overline{P}_{0j}(0,\tau)+\mathcal{P}(B_0)\cdot\overline{P}_{00}(0,\tau).
\end{aligned}$$

Therefore, we get

$$\begin{aligned}
h_j(t) &= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\frac{\mathcal{P}(B_0)\cdot\mathcal{P}(X(t)=j,\ t\le T<t+\Delta_t|B_0)}{\mathcal{P}(B_0)\cdot\overline{P}_{0j}(0,\tau)+\mathcal{P}(B_0)\cdot\overline{P}_{00}(0,\tau)} \\
&= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\frac{\mathcal{P}(X(t)=j,\ t\le T<t+\Delta_t|B_0)}{\overline{P}_{0j}(0,\tau)+\overline{P}_{00}(0,\tau)}
\end{aligned}$$

which yields (5). The last part of the statement of the theorem can be proved similarly. □

2.3 The Larson and Dinse approach

There is a close connection between the purged chains of Section 2.2 and the mixture model for competing risks of Larson and Dinse. In the Larson and Dinse approach one considers that

$$\mathcal{P}(T\le t, D=j)=\mathcal{P}(T\le t|D=j)\cdot\mathcal{P}(D=j) \tag{8}$$

which corresponds to

$$\overline{P}_{0j}(0,t) = \big[1 - P_{00}(0,t)\big] \cdot \mathcal{P}(D = j) = \big[1 - \exp\big(-\int_0^t h_j(s)ds\big)\big] \cdot \mathcal{P}(D = j), \qquad (9)$$

respectively. In other words, the original competing risks model $(X_t)_{t\leq 0}$ is subject to decomposition into J purged Markov chains. Therefore, the driving forces of this approach are $h_j(t)$ and $\mathcal{P}(D = j)$, $j = 1, \ldots, J$, which can be estimated through maximum likelihood estimation by means of some specific semi-parametric and/or parametric models (see [4]). Note that because the random variable D is not observable on its own, $\mathcal{P}(D = j)$ cannot be estimated model-free.

It is straightforward to show that $h_j(t) \geq \alpha_j(t)$, $j = 1, \ldots, J$. This overestimation of the failure rate due to cause j in the purged chain is intuitively clear too, because of the selection (mortality selection) made in the purged chain. Therefore, caution should be taken when we want to transfer conclusions from the the original chain to the purged chains, according to Theorem 1. Moreover, because the original chain is completely determined by the cause-specific hazards, which are the "natural" observable quantities in competing risks, non-parametric estimation in this chain could be "transferred" via Theorem 1 to the Larson and Dinse approach.

3 Acknowledgement

The author gratefully acknowledges support by CNCSIS-UEFISCSU research grant PNII - IDEI 209/2007 and by Netherlands Organization for Scientific Research Grant ZONMW-912-07-018.

References

[1] Putter, H., Fiocco, M. and Geskus, R.B.: *Tutorial in biostatistics: Competing risks and multi-state models*. In: Stat. Med. **26** (2007) No. 11, p. 2389 – 2430.

[2] Hoem, J. M.: *Purged and partial Markov chains*. In: Skand. Aktuar Tidskr. **52** (1969), p. 147 – 155.

[3] Iosifescu, M.: *Finite Markov Processes and Their Applications*. New York. John Wiley & Sons, (1980).

[4] Larson, M. G., Dinse, G. E.: *A mixture model for the regression-analysis of competing risks data*. In: J. R. Statist. Soc. C **34** (1985) No. 3, p. 201 – 211.