DSC 2003 Working Papers
(Draft Versions)
http://www.ci.tuwien.ac.at/Conferences/DSC-2003/

Estimating event histories of Markov chains using covariate information in R and S-PLUS

Harald Fekjær[1,2], **Odd O. Aalen**[2], **Ørnulf Borgan**[3]

[1]Cancer Registry of Norway; [2]Section of Medical Statistics, University of Oslo, Norway;
[3]Institute of Mathematics, University of Oslo, Norway

Correspondence to:
Harald Fekjær
Cancer Registry of Norway (Kreftregiteret)
Montebello, N-0310 Oslo, Norway
E-mail: harald.fekjaer@kreftregisteret.no

An S based software package for estimating event histories of Markov chains using covariate information and Aalen's addtive hazard regression model. See:
http://www.med.uio.no/imb/stat/addreg/

1 Background:

In clinical trials, epidemiology (see: figure 1 and figure 2) and many other fields, one often observes a number of individuals passing through several states. Each individual may have a set of covariates measured, and one may want to estimate probabilities of interesting transitions within the state space, and how these transitions depend on the covariates.

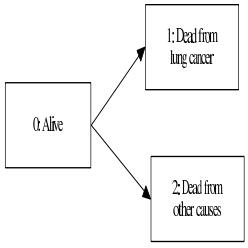

Figure 1: A model for competing risks

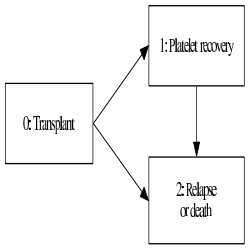

Figure 2: A Markov model for patients who have got a bone marrow transplantation

2 Method:

When estimating transition probabilities in Markov chains using covariate information, there are two steps involved. The first one is to estimate how the transition rates between the various states depend on the covariates. The second one is to combine the estimated rates for a given set of covariates into appropriate transition probabilities. In Aalen et. al. (2001)[1], several already well-established methods are combined for these steps, and a method for estimating transition probabilities in Markov chains using covariate information is presented.

An empirical transition matrix is used to calculate the transition probabilities, while Aalen's additive hazard regression model[2,3] is used for estimating the transition rates between the various states. Aalen's additive model is an alternative to Cox's proportional hazard regression model[4,5], and gives a very flexible modeling of time-varying effects, something that is especially useful when estimating transition probabilities. In addition, the method is based on martingale and counting process theory[6], which is crucial when it is later combined with the empirical transition matrix.

3 Implementation

Using the Addreg software[7], Aalen's additive hazard regression model can be fitted on each transition. After choosing suitable covariates for each transition, the result-

ing "addreg"-objects are marked with a "from" and "to" state, and given as a list to the addregmc-function together with covariates. The transition probabilities are then calculated (together with confidence intervals) for the given set of covariates, and can easily be plotted or extracted at certain time-points of interest. This list notation with "from/to addreg"-objects, makes a very flexible way of specifying Markov chains, so that any Markov chain can be specified (with only the available computer time and memory as a limit) and not only certain special cases.

4 Results:

Both theory and software are now available for estimation transition probabilities in "any" Markov chain using covarite information. The software is freely available (GNU license) on the Addreg webpage[7], and runs on both the R[8] and S-PLUS[9] software packages. To download a copy, go to:

http://www.med.uio.no/imb/stat/addreg/

Figure 3 gives an example of the results.

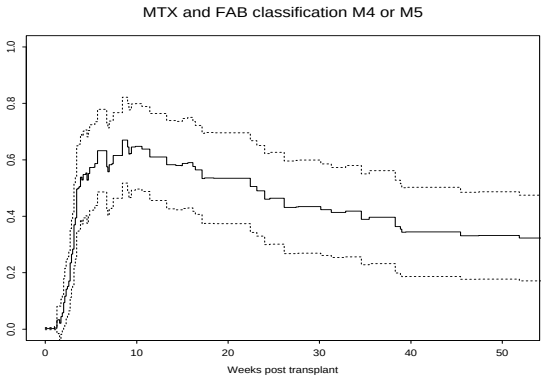

Figure 3: Estimated probability of being in response ($P_{01}(0,t)$) with pointwise 95% confidence intervals for an AML high risk patients with FAB classification M4 or M5 using MTX

5 Funding:

The project recieved financial suport from the Norwegian Cancer Society and University of Oslo, Norway. At the University of Oslo, support is given though "Section of Medical Statistics" and the thematic research area "Norevent"[10] at the Medical faculty.

6 Reference:

1. Aalen, O., Borgan, Ø., Fekjær, H. (2001). *Covariate adjustment of event histories estimated from Markov chains: The additive approach*, Biometrics, vol. 57, 108-116.[1]

2. Aalen, O. (1980). *A model for nonparametric regression analysis of counting processes*, Lecture Notes in Statistics, Vol. 2, Springer Verlag, New York, 1-25.

3. Aalen, O.O. (1989). *A linear regression model for the analysis of life times*, Statistics in Medicine, 8: 907-925.

4. Cox, D.R. (1972). *Regression models and life-tables.* Journal of the Royal Statistical Society, Series B 34:187-202.

5. Klein, J.P. and Moeschberger, M.L. (1997). *Survival analysis: techniques for censored and truncated data.* Springer-Verlag, New York.

6. Andersen, P.K., Borgan, Ø., Gill, R.D. and Keiding, N. (1993). *Statistical Models Based on Counting Processes.* Springer-Verlag, New York.

7. http://www.med.uio.no/imb/stat/addreg/

8. http://www.R-project.org/

9. http://www.insigthful.com/

10. http://www.med.uio.no/imb/stat/Norevent/

[1] A preprint version of this paper is available from
http://www.math.uio.no/eprint/stat_report/2001/02-01.html