Recent Developments in Censored, Non-Markov Multi-State Models

Jacobo de Uña-Álvarez

Abstract. Nonparametric estimation of transition probabilities for a censored multi-state model is traditionally performed under a Markov assumption. However, this assumption may (and will) fail in some applications, leading to the inconsistency of the time-honoured Aalen-Johansen estimator. In such a case, alternative (non-Markov) estimators are needed. In this work we review some recent developments in this area. We also review the key problem of testing if a given (censored) multi-state model is Markov, giving modern ideas for the construction of an omnibus test statistic.

1 Introduction

A multi-state model is a model for a stochastic process $\{X(t), t \geq 0\}$ allowing individuals to move along a finite number of states. At each time point t, $X(t)$ denotes the state occupied by a representative individual in a homogeneous population, and let $\{X_r(t), t \geq 0\}$, $r = 1, ..., n$, be a collection of n trajectories (or histories) corresponding to n subjects randomly sampled from the target population. In this setup, much effort has been made to estimate the so-called transition probabilities

$$p_{ij}(s,t) = P(X(t) = j | X(s) = i) \tag{1}$$

where i and j are two states, and $0 \leq s < t$. The obvious estimator for $p_{ij}(s,t)$ is the empirical transition probability

$$p_{ij,n}(s,t) = \frac{\sum_{r=1}^{n} I(X_r(t) = j, X_r(s) = i)}{\sum_{r=1}^{n} I(X_r(s) = i)} \tag{2}$$

which is simply the proportion of observed transitions from i to j in the time interval $[s,t]$. In practice, this estimator is typically unavailable because of

Jacobo de Uña-Álvarez
Department of Statistics and OR, University of Vigo, Spain
e-mail: jacobo@uvigo.es

C. Borgelt et al. (Eds.): Combining Soft Computing & Stats. Methods, AISC 77, pp. 173–179.
springerlink.com

censoring. [1] introduced a nonparametric estimator for $p_{ij}(s,t)$ under censoring. The Aalen-Johansen estimator has become the standard tool for estimating the transition probabilities in a nonparametric way. However, it is constructed on the basis that the underlying process is Markovian, and its consistency can not be ensured in general.

To be more precise, applications of multi-state models to biomedical data have shown that the Markov assumption is sometimes violated. A stochastic process is said to be Markovian when, given the present, the future evolution does not depend on the past. Consider as an illustrative example the PROVA trial of bleeding episodes and mortality in liver cirrhosis in [2]. In this example, a three-state model is used to represent the individuals' histories; this model allows for three possible transitions in a progressive way: from healthy (no bleeding) to bleeding, from bleeding to death, and directly from healthy to death. This model is usually named 'illness-death model', and it is progressive in the sense that past states can not be visited again. [2] provided evidence (in agreement to previous studies) on the fact that the mortality is markedly increased shortly after the bleeding episode. This means that subjects in state 'bleeding' may (and will) have a different prognosis according to their entry times in that state (i.e., the past history is important for the future, so the process is not Markov).

There has been some investigation oriented to analyze the properties of the Aalen-Johansen estimator when the Markov assumption fails. For example, Aalen et al. [1] and Datta and Satten [3] established the consistency of the Aalen-Johansen estimator of the stage occupation probabilities $P_j(t) = P(X(t) = j)$ for a non-Markov process, while Glidden [4] developed confidence bands for such an estimator. More recently, however, Meira-Machado et al. [5] showed that, in general, the Aalen-Johansen estimator of $p_{ij}(s,t)$ may be dramatically biased if the Markov assumption is not fulfilled. The practical conclusion is that one should asses the Markovianity of the process before using the Aalen-Johansen estimator for estimation and inference purposes. And, if there is evidence of non-Markovianity, some alternative estimators should be used.

In this work we review some recent developments in nonparametric estimation of transition probabilities in non-Markov multi-state models. It is assumed that the available trajectories can be right-censored by a potential censoring time that is independent of the process. The available estimators are given in Section 2. In Section 3, we consider the problem of testing if a given process is Markov, reviewing the traditional approach and giving some modern alternative ideas too.

2 Non-Markov Transition Probabilities

For the best of our knowledge, Meira-Machado et al. [5] proposed for the first time nonparametric estimators for the transition probabilities of a

censored non-Markov multi-state model. These authors considered the progressive illness-death (or disability) model, which consists in three different states (1='healthy', 2='diseased', and 3='dead') and the three possible transitions 1→2, 2→3, and 1→3. Put Z and T for the sojourn time in state 1 and the total survival time of the process (that is, the time up to reaching the absorbing state 3) respectively. It is seen in [5] that the transition probabilities are probabilities involving (Z,T), and hence the question becomes how the joint distribution function of (Z,T) can be consistently estimated under censoring. Here we briefly present their ideas, with a slightly different notation to simplify things.

The available information is represented by $\left(\widetilde{Z},\widetilde{T},\Delta_1,\Delta\right)$, where $\widetilde{Z}$ and $\widetilde{T}$ stand for the censored versions of Z and T, and Δ_1 and Δ are their respective censoring indicators. Note that the individual is observed to pass through state 2 if and only if $\widetilde{Z}<\widetilde{T}$, and in such a case $\widetilde{Z}$ is uncensored. Let $\left\{\left(\widetilde{Z}_i,\widetilde{T}_i,\Delta_{1i},\Delta_i\right),1\le i\le n\right\}$ be an iid sample of $\left(\widetilde{Z},\widetilde{T},\Delta_1,\Delta\right)$, and let $W_i=\frac{\Delta_i}{n-R_i+1}\prod_{R_j<R_i}\left[1-\frac{\Delta_j}{n-R_j+1}\right]$ be the Kaplan-Meier weight attached to $\widetilde{T}_i$ (here $R_i=Rank(\widetilde{T}_i)$). With this notation, any functional of the form $S(\varphi)=E\left[\varphi(Z,T)\right]$ is estimated by $S_n(\varphi)=\sum_{i=1}^n W_i\varphi\left(\widetilde{Z}_i,\widetilde{T}_i\right)$. By noting that

$$p_{11}(s,t)=\frac{P(t<Z)}{P(s<Z)},\quad p_{13}(s,t)=\frac{P(s<Z,T\le t)}{P(s<Z)}=\frac{E\left[\varphi_{s,t}(Z,T)\right]}{P(s<Z)},\qquad(3)$$

where $\varphi_{s,t}(u,v)=I(u>s,v\le t)$, and $p_{12}(s,t)=1-p_{11}(s,t)-p_{13}(s,t)$, we have the following estimators for $\left\{p_{1j}(s,t),j=1,2,3\right\}$:

$$\widehat{p}_{11}(s,t)=\frac{\widehat{S}_Z(t)}{\widehat{S}_Z(s)},\quad \widehat{p}_{13}(s,t)=\frac{1}{\widehat{S}_Z(s)}\sum_{i=1}^n W_i\varphi_{s,t}\left(\widetilde{Z}_i,\widetilde{T}_i\right),\qquad(4)$$

and $\widehat{p}_{12}(s,t)=1-\widehat{p}_{11}(s,t)-\widehat{p}_{13}(s,t)$, where $\widehat{S}_Z(.)$ is the Kaplan-Meier estimator of the survival function of Z. Similarly, since

$$p_{23}(s,t)=\frac{P(Z\le s,s<T\le t)}{P(Z\le s<T)}=\frac{E\left[\widetilde{\varphi}_{s,t}(Z,T)\right]}{P(T>s)-P(s<Z)}\qquad(5)$$

where $\widetilde{\varphi}_{s,t}(u,v)=I(u\le s,s<v\le t)$, and $p_{22}(s,t)=1-p_{23}(s,t)$, one can introduce the following estimators for $\left\{p_{2j}(s,t),j=2,3\right\}$:

$$\widehat{p}_{23}(s,t)=\frac{1}{\widehat{S}_T(s)-\widehat{S}_Z(s)}\sum_{i=1}^n W_i\widetilde{\varphi}_{s,t}\left(\widetilde{Z}_i,\widetilde{T}_i\right)\quad\text{and}\quad\widehat{p}_{22}(s,t)=1-\widehat{p}_{23}(s,t),\ (6)$$

where $\widehat{S}_T(.)$ is the Kaplan-Meier estimator of the survival function of T.

In the uncensored case, all the involved Kaplan-Meier weights reduce to $1/n$, and hence the estimators introduced along (4) and (6) collapse to (2).

Under censoring, the consistency of $\widehat{S}_Z(.)$ and $\widehat{S}_T(.)$ follows because the censoring is independent of the process; on the other hand, Theorem 1 in [5] ensures the consistency of $\widehat{p}_{13}(s,t)$ and $\widehat{p}_{23}(s,t)$, which involve multivariate Kaplan-Meier integrals in the sense of [7]. When the process is Markov, the estimators (4)-(6) are less efficient than their Aalen-Johansen counterparts. However, when the Markov condition is violated, these estimators are preferred since Aalen-Johansen can be systematically biased. Both estimators (Markov, non-Markov) were compared in simulated and practical settings; [5] used the PROVA trial data as an illustrative example, while [6] provided a nice comparison of the methods in a trial on breast cancer.

The idea in [5] for the construction of non-Markov transition probabilities in the illness-death model can be generalized to any other progressive multistate model. Certainly, since each given transition probability is a function of the sojourn times in the several existing states, the Kaplan-Meier weights pertaining to the time up to reaching the final absorbing state can be used to construct consistent empirical weighted averages. Details on this are provided in [8].

Doubtless, the main drawback of the non-Markov estimators is their large variance, mainly in the heavily censored case. In order to mitigate this problem, some presmoothing of the censoring indicators Δ_i can be performed. By 'presmoothing' it is meant that each Δ_i is replaced by some fit to the binary regression $m(z,t) = P\left(\Delta = 1 | \widetilde{Z} = z, \widetilde{T} = t\right)$ before the Kaplan-Meier weights W_i are computed. In [10] this idea was applied in the scope of the three-state progressive model (which is just an illness-death model with forbidden transition $1 \rightarrow 3$) to introduce a new estimator of the joint distribution of the sojourn times. Also, [11] consider presmoothed non-Markov transition probabilities for the illness-death model. The main consequence of this approach is that non-Markov estimators with improved variance can be obtained. In practice, the presmoothing function $m(z,t)$ is estimated by fitting some parametric model or via nonparametric regression methods.

3 Testing the Markov Assumption

In this section, we consider the problem of checking the Markov assumption in practice. Note that this issue is relevant, since (as discussed above) the Aalen-Johansen estimator may be inconsistent when the Markov condition is violated. For simplicity of exposure, consider the progressive illness-death model and let $\lambda(t|s)$ be the hazard rate of T at time $t \geq s$ conditionally on $Z = s$ and $Z < T$. Note that the Markov assumption states that the value of $\lambda(t|s)$ does not depend on s. This is typically tested via a proportional hazards specification $\lambda(t|s) = \lambda_0(t)e^{\beta s}$. Then, the null hypothesis representing the Markov condition is $H_0 : \beta = 0$. The model can be fitted (and a test performed) by standard methods from the cases with an uncensored Z. For the PROVA trial data, the estimated coefficient was $\widehat{\beta} = 0.00526$ (s.e. $= 0.00167$),

and the likelihood ratio test gave a p-value of $p = 0.000434$ thus rejecting the Markov condition.

In practice, the model $\lambda(t|s) = \lambda_0(t)e^{\beta s}$ may not be appropriated due to several reasons. First, the linear predictor βs can not cope in general with other type of effects. To illustrate this, consider the three-state progressive model with $\log(T_2) = f(Z) + \varepsilon$ where $T_2 = T - Z$ and where ε is an error term independent of Z. In this case we get $\lambda(t|s) = \lambda_0((t-s)e^{-f(s)})e^{-f(s)}$ where λ_0 stands for the hazard of $W = e^{\varepsilon}$. Take a extreme-value distribution for ε (so λ_0 becomes constant), $Z \sim U[0,2]$ and $f(s) = (s-1)^2$. Then, the test for $\beta = 0$ under the linear specification $\lambda(t|s) = \lambda_0(t)e^{\beta s}$ is expected to have low (or even no) power. Second, the proportional hazards assumption may fail; this is the case, for example, when $\log(T_2) = f(Z) + \varepsilon$ and $W = e^{\varepsilon}$ does not follow a Weibull distribution. Of course, this may influence the performance of the test. In Table 1 we report the proportion of rejection of this test at level $\alpha = 0.05$ for several sample sizes n among 1,000 Monte Carlo trials in these two situations (we take $f(s) = 0$ and $\varepsilon \sim N(0,1)$ in the second case, labeled as No PH; in this case, the Markov assumption does not hold because T_2 is not exponentially distributed). We see that, in these two cases, the classical test exhibits a very poor power. Of course, in the first simulated scenario (labeled as PH) power could be increased through a more flexible specification of the predictor; in the second model, however, there is not a clear solution to the lack of power of the test.

In Table 1, we have also reported the results pertaining to a new method of testing. The new method is based on the fact that, under the Markov assumption, the variables T and Z are independent conditionally on the event $A_t = \{Z \leq t < T\}$, for each given $t > 0$. More specifically, we have performed a test of no correlation between T and Z conditionally on A_t with $t = 2$. The choice $t = 2$ is interesting because it guarantees the maximum expected sample size (i.e. the largest $P(A_t)$) under the two simulated models. We see in Table 1 that this new idea may lead to a more powerful test. Besides, we have seen in the simulations that there is some negative correlation between the p-values of the classical test and those of the new approach, indicating that both testing procedures are able to detect different type of alternatives. Hence, a complementary use of both approaches could be recommended in practice.

Clearly, the combination of several t values should help to increase the power of the new method of testing. Moreover, by considering a whole set of t-values one could explore the variation of the pertaining p-values. To illustrate this, consider the significance trace $\{p(t) : t \in [t_0, t_1]\}$, where $p(t)$ stands for the p-value of the suggested correlation test, when conditioning on A_t. In Figure 1 we depict this curve for the simulated model No PH in Table 1 with $n = 500$, $t_0 = 1$ and $t_1 = 3$; the given curve is indeed the first quartile of the p-values along the 1,000 Monte Carlo simulations. Roughly speaking, the information in this Figure is that (a) in more than 25% of the cases the trace is able to reject the Markovianity of the process (recall however the

Table 1 Proportion of rejection at level $\alpha = 0.05$ along 1,000 Monte Carlo simulations of sample size n for two non-Markov three-state progressive models: classical method vs. new approach ($t = 2$).

	PH		No PH	
n	Classical	New	Classical	New
100	.069	.119	.073	.097
250	.096	.159	.089	.157
500	.091	.193	.099	.248
1000	.122	.267	.148	.378

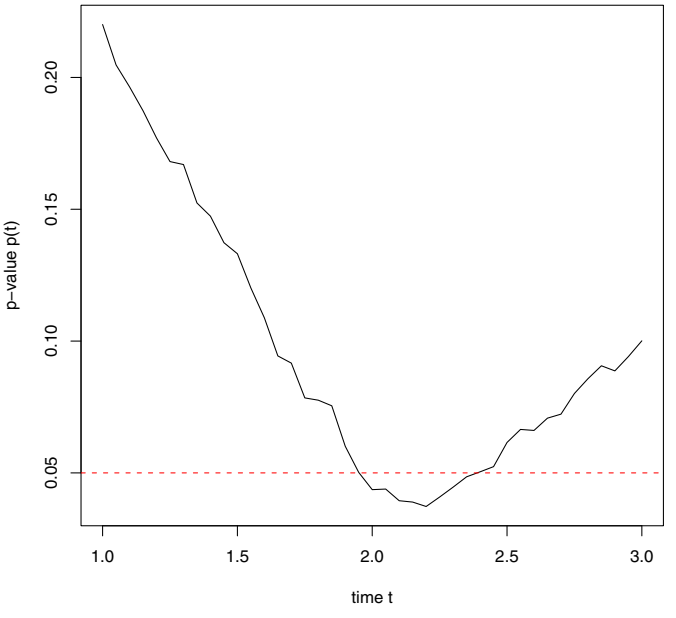

Fig. 1 Significance trace for No PH model with $n = 500$: first quartile along 1,000 Monte Carlo simulations.

poor power of the classical method in this case, see Table 1), and that (b) the greatest evidence against the null is achieved around $t = 2.2$.

Of course, although we have used the Pearson correlation coefficient to implement the new method, it can be adapted in an obvious manner to be based on other measures of association too. A key issue here is how to incorporate the censoring effects in the definition of the test statistic. This is not obvious at all. In [9], an omnibus test statistic which compares the joint distribution function of (Z, T) to the product of marginals conditioning on each A_t was introduced, accounting for censoring effects. However, the performance of this

test in practice is still unexplored, and this seems to be a very promising field of research.

The problem of testing the Markov assumption has been discussed here for the illness-death model (and for the three-state progressive model) for the sake of conciseness. In general, one will be interested in testing that the entry time to the present state (and other measured covariates in the individual's history) is unrelated to the future hazard. At the end, this type of assumptions can be formalized in a simple manner so the methods reviewed here (or obvious modifications) still apply.

Acknowledgements. The author thanks P.K. Andersen for providing the PROVA trial data. Research funded by projects MTM2008-03129, PGIDIT07PXIB300191PR, and by the INBIOMED project (DXPCTSUG, Ref. 2009/063).

References

1. Aalen, O., Johansen, S.: An empirical transition matrix for nonhomogeneous Markov chains based on censored observations. Scand. J. Stat. 5, 141–150 (1978)
2. Andersen, P.K., Esbjerj, S., Sorensen, T.I.A.: Multistate models for bleeding episodes and mortality in liver cirrhosis. Stat. Med. 19, 587–599 (2000)
3. Datta, S., Satten, G.A.: Validity of the AalenJohansen estimators of stage occupation probabilities and Nelson Aalen integrated transition hazards for non-Markov models. Stat. Probab. Lett. 55, 403–411 (2001)
4. Glidden, D.: Robust inference for event probabilities with non-Markov event data. Biometrics 58, 361–368 (2002)
5. Meira-Machado, L., de Uña-Álvarez, J., Cadarso-Suárez, C.: Nonparametric estimation of transition probabilities in a non-Markov illness-death model. Lifetime Data Anal. 12, 325–344 (2006)
6. Meira-Machado, L., de Uña-Álvarez, J., Cadarso-Suárez, C., Andersen, P.K.: Multi-state models for the analysis of time-to-event data. Statist. Meth. Med. Research 18, 195–222 (2009)
7. Stute, W.: Consistent estimation under random censorship when covariables are present. J. Multivariate Anal. 45, 89–103 (1993)
8. de Uña-Álvarez, J.: Estimación no paramétrica en modelos multi-estado no-Markovianos. In: Actas del XXX Congreso Nacional de la SEIO, Valladolid, Spain (2007a)
9. de Uña-Álvarez, J.: Testing that a multi-state model is Markov: new methods. In: Gomes, M.I., Pestana, D., Silva, P. (eds.) Abstracts of the 56th Session of the ISI, Lisbon, Portugal (2007b)
10. de Uña-Álvarez, J., Amorim, A.P.: A semiparametric estimator of the bivariate distribution function for censored gap times. Discussion Papers in Statistics and OR 09/03, University of Vigo (2009)
11. de Uña-Álvarez, J., Amorim, A.P., Meira-Machado, L.: Presmoothing the transition probabilities in the illness-death model (in preparation, 2010)