The Kaplan-Meier Estimator: New Insights and Applications in Multi-state Survival Analysis

Luís Meira-Machado(✉)

Centre of Mathematics and Department of Mathematics, University of Minho, Guimarães, Portugal
lmachado@math.uminho.pt
https://w3.math.uminho.pt/~lmachado/

Abstract. A topic that has received attention in statistical and medical literature is the estimation of survival for which the Kaplan-Meier product-limit estimator is the most commonly used estimator. This estimator is considered nonparametric because it does not rely on any assumptions about the probability distribution of the lifetime. The best known representation of the Kaplan-Meier estimator is based on a product of elementary probabilities whose underlying idea is the computation of conditional survival probabilities. The Kaplan-Meier estimator of survival can also be explained using the redistribution to the right algorithm, which removes the mass of a censored subject and redistributes this mass equally to all subjects who fail or are censored at later times. This paper presents additional alternative representations of this estimator, as well as applications and advantages of its use. One of these representations consists in defining the estimator as a sum of weights, which is a convenient form to estimate several quantities in the context of multi-state models. The estimator can also be represented as a weighted average of identically distributed terms, where the weights are obtained by using the inverse probability of censoring. The paper discusses how these formulations can be used to estimate several quantities in the context of multi-state models. Two real data examples are included for illustration of the methods.

Keywords: Censoring · Kaplan-Meier · Multi-state models · Survival Analysis

1 Introduction

Censored data analysis has several applications in longitudinal medical studies. In this type of data, it is frequent to have incomplete observations due to various factors such as loss of follow-up and dropout. The Kaplan-Meier product-limit estimator [1] has always been considered a standard method to obtain statistical summaries for such data. This method can handle right-censoring, which

Supported by Portuguese Foundation for Science and Technology, references UIDB/00013/2020, UIDP/00013/2020 and EXPL/MAT-STA/0956/2021.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
O. Gervasi et al. (Eds.): ICCSA 2023 Workshops, LNCS 14112, pp. 129–139, 2023.
https://doi.org/10.1007/978-3-031-37129-5_11

is the most common type of censoring in survival analysis. It is implemented in almost all statistical software, including R, SPSS, and STATA. The survival package of the statistical software R is the most used package to implement the Kaplan-Meier method. To deal with other forms of censorship beyond right-censoring, other nonparametric estimators of survival, such as the Lynden-Bell and Turnbull estimators, have been developed. The Lynden-Bell [2] and Turnbull [3] estimators can be seen as generalizations of the Kaplan-Meier estimator, with the Lynden-Bell estimator designed to handle left-truncated and right-censored data, and the Turnbull estimator capable of handling interval-censored data in addition to right-censored data.

The Kaplan-Meier estimator can be generalized to inhomogeneous Markov processes with a finite number of states. This generalization was considered by Aalen [4] for the competitive risk model and independently by Aalen and Johansen [5] and Fleming [6] for the general case. Aalen and Johansen [5] introduced the so-called Aalen-Johansen estimator of the transition probabilities which can be seen as a matrix version of the Kaplan-Meier estimator. This relationship makes it possible to present a weighted version of the Aalen-Johansen estimator.

This article is organized as follows. In Sect. 2, the Kaplan-Meier estimator is introduced, and different representations of the estimator are presented. In Sect. 3, a modification of the Kaplan-Meier estimator is introduced that provides an estimator with less variability. In Sect. 4, it is shown how the different representations can be useful to estimate particular quantities in multi-state models. In Sect. 5, the methods are illustrated using two real data sets. The main conclusions of this research are deferred to Sect. 6.

2 The Kaplan-Meier Estimator of Survival

Let T_i be a random variable denoting the lifetime for $i = 1, \cdots, n$ individuals from a homogeneous population. Consider also that C_i is a random variable denoting the censoring time for the i-th individual. Due to this censoring variable, in practice, what is observed is $Y_i = \min(T_i, C_i)$ and $\Delta_i = I(T_i \leq C_i)$. The censoring is assumed to be independent in the sense that the additional knowledge of censorings before any time t does not alter the risk of failure at time t. This section aims to introduce the Kaplan-Meier estimators of survival $S(t) = P(T > t)$ using different formulations.

2.1 The Product-Limit Formulation

The best-known representation of the Kaplan-Meier estimator is based on a product of elementary probabilities of the occurrence of an event at a certain point in time, providing estimates that decrease according to the observations (failures or events) detected. This representation is discussed in most books on survival analysis. Consider lifetime data with the event of interest being a failure. Then, the expression of the Kaplan-Meier estimator involves calculating

the number of individuals who fail at a given time, divided by the number of individuals at risk for failure at that same time. These successive probabilities are then multiplied by any earlier computed probabilities to get the final estimate.

Denote the J ordered failure times by $0 < t_1 < t_2 < \cdots < t_J$; and let d_j be the number of individuals who fail at time t_j, and r_j the number of individuals at risk just before that time. The idea behind the Kaplan-Meier estimator is to compute the conditional survival probabilities, $P(T \geq t_j | T \geq t_{j-1}) = 1 - d_j/r_j$. Then, the Kaplan-Meier survival estimator is the unconditional probability of survival, which is just the cumulative product of the conditional probabilities:

$$\widehat{S}^{KM}(t) = \prod_{j:t_j \leq t} \left(1 - \frac{d_j}{r_j}\right) \tag{1}$$

The expression for Eq. 1 can also be derived from maximum likelihood estimation of the hazard function [7,8].

The product-limit estimator is a nonparametric technique for estimating and plotting the survival probability as a function of time. The estimated survival plot is a step function that starts at the value 1 and has a series of decreasing horizontal steps that approach the shape of the population's true survival function. When there is no censoring, the Kaplan-Meier estimator is equivalent to the empirical estimator, for which every observation contributes equally to the size of the step. Confidence intervals are typically presented alongside the plot in order to demonstrate the level of uncertainty surrounding point estimates. One of the most common estimators that have been used to approximate its variance is Greenwood's formula. For more details on the calculation of the confidence intervals using the Greenwood method, please see Borgan [9].

2.2 The Weighted Version of the Kaplan-Meier Estimator

The Kaplan-Meier estimator can also be represented as a sum of weights, which is known as Kaplan-Meier weights. The weights correspond to the 'jump' of the estimator's distribution and are defined only for the event times. These weights can be obtained by the expression $W_j = \widehat{S}^{KM}(t_{j-1}) - \widehat{S}^{KM}(t_j)$, for $j = 2, \cdots, J$, with $W_1 = 1 - \widehat{S}^{KM}(t_1)$ and where $t_1 < \cdots < t_J$ are the ordered failure times. Then, the Kaplan-Meier estimator can be defined as follows:

$$\widehat{S}^{KM}(t) = 1 - \sum_{j=1}^{J} W_j I(t_j \leq t) \tag{2}$$

Next, an alternative notation is introduced that allows expressing how each failure time contributes to survival, even when there are tied failure times. For $i = 1, 2, \cdots, n$, let W_i be the Kaplan-Meier weight associated with Y_i when the marginal distribution of Y is estimated from pairs (Y_i, Δ_i)'s. That is,

$$W_i = \frac{\Delta_{[i]}}{n - i + 1} \prod_{j=1}^{i-1} \left[1 - \frac{\Delta_{[j]}}{n - j + 1}\right] \tag{3}$$

is the weight associated with $Y_{(i)}$, where $Y_{(1)} < \cdots < Y_{(n)}$ are the ordinal statistics obtained from Y_i's times and where $\Delta_{[i]}$ denotes the i-th concomitant. Ties between censored observations or between events are ordered arbitrarily. Ties between events and censoring times are treated so that the former (events) precede the latter (censoring).

For uncensored cases, the weight W_i is equal to $1/n$ for each i, while censored observations do not contribute to the weighting. The estimator of the Kaplan-Meier survival function can then be defined as follows:

$$\widehat{S}^{KM}(t) = 1 - \sum_{i=1}^{n} W_i I(Y_i \leq t) \tag{4}$$

As can easily be seen, the estimator is a step function with jump points located at the failure times. As will be shown below, the size of the weights (and the Kaplan-Meier 'jump') is a non-decreasing function. Later on, this form of representation of the Kaplan-Meier estimator will be used as a convenient way to introduce the idea of 'presmoothing'.

2.3 The Redistribution to the Right Algorithm

An alternative and equivalent representation of the product-limit formulation considers a redistribution to the right algorithm in which the weight of the censored observations is distributed over all observations to the right. The algorithm has the following four steps:

Step 1: Arrange the data in increasing order, with censored observations to the right of uncensored observations in the case of ties.
Step 2: Put equal mass $1/n$ at each of the n observations.
Step 3: Start from the smallest observation and move to the 'right'. Each time a censored observation is reached, redistribute its mass evenly to all observations to the right.
Step 4: Repeat Step 3 until all censored observations (except largest observations) have no mass.

The algorithm then attributes mass (weight, W_i) to each failure time ($\geq 1/n$) and null mass (weight) to censored observations. The estimator can then be defined as in Eq. 4.

2.4 Inverse Probability of Censoring Weighting

The Kaplan-Meier estimator can also be represented as a weighted average of identically distributed terms, where the weights are obtained by using inverse probability censoring weighting (IPCW). Satten and Datta [10] showed that this representation, by means of a weighted average, is convenient for the development of asymptotic theory and leads to an interesting variance decomposition of the Kaplan-Meier estimator. The resulting estimator is written as follows:

$$\widehat{S}^{IPCW}(t) = 1 - \frac{1}{n}\sum_{i=1}^{n} \frac{I(Y_i \leq t)\Delta_i}{\widehat{G}(Y_i^-)} \tag{5}$$

where $P(C > y) = G(y^-)$ and where $\widehat{G}$ denotes the Kaplan-Meier estimator of the censoring variable C that is computed using the pairs $(Y_i, 1 - \Delta_i)$'s; that is, $\widehat{G}$ is the estimator of the survival function for censored data using $\widehat{S}^{KM}(\cdot)$ but considering failure times as 'censored' observations and censored times as 'failures'. The equivalence of the IPCW estimator and the KM estimator is shown in [10]. The IPCW formulation has been used to deal with censoring in several contexts, such as to introduce nonparametric estimators of time-dependent ROC curves conditional on covariates [11], conditional transition probabilities [12] or the conditional cumulative incidence functions [13].

3 Presmoothed Kaplan-Meier Estimators

Presmoothing was used by [14] and by several authors later (see, for example [15] and references therein). The underlying idea is to replace each censoring indicator in the weighted version of the Kaplan-Meier estimator with a smooth fit of a regression model with a binary response. This replacement reduces the variability of the estimator, with a possible but low cost of bias.

Let $m(y) = P(\Delta = 1 \mid Y = y)$ be the conditional probability of observing the event given that $Y = y$ is observed. This function can be estimated parametrically (e.g., using a logistic regression model) or nonparametrically (using kernel or spline-type functions). The presmoothed estimator of the survival function is simply a modified version of the weighted version of the Kaplan-Meier estimator with smoothed weights that are obtained by replacing the censoring indicator variables. The term 'presmoothing' comes from the fact that smoothing is only used to obtain a modified version of the Kaplan-Meier estimator; the estimate itself is not smoothed. The new estimator is given by as follows:

$$\widehat{S}^{PKM}(t) = 1 - \sum_{i=1}^{n} W_i^{PKM} I(Y_i \leq t) \tag{6}$$

where

$$W_i^{PKM} = \frac{m_n(Y_{(i)})}{n - i + 1} \prod_{j=1}^{i-1} \left[1 - \frac{m_n(Y_{(j)})}{n - j + 1}\right] \tag{7}$$

and where $m_n(y)$ is an estimator of $m(y)$ based on the (Y_i, Δ_i)'s. This function can be estimated using a parametric estimator such as a logistic regression model or using the Nadaraya-Watson kernel type estimator [15]. Recent contributions suggest that the resulting estimator has less variability in the estimation, in particular in the right-hand tail, when compared to the original (nonparametric) estimator.

4 Application to Multi-state Models

In longitudinal medical studies, patients may observe multiple events in a given follow-up period. The analysis of these studies can be successfully performed by multi-state models [13,16,17]. A multi-state model is a stochastic process $(X(t), t \in \mathcal{T})$ with a finite state space, where $X(t)$ represents the state occupied by the process at time $t \geq 0$. One such model is the progressive illness-death model, shown in Fig. 1, which is fully characterized by three states and three transitions between them. The stochastic process of this model is characterized by its three potential transition times T_{hj}, from State h to State j. In this model, there are two competitive transitions leaving State 1, but only the transition with the lower time to occurrence is observed. Accordingly, denote by $\rho = I(T_{12} \leq T_{13})$ the indicator function of transition to State 2 in a given time and consider $U = \min(T_{12}, T_{13})$ the length of stay in State 1. Finally, denote by $T = U + \rho T_{23}$ the total time of the process. However, the vector (Z, Y, Δ_1, Δ) is observed in practice due to limitations in follow-up, follow-up losses, and other factors. Here, $Z = \min(U, C)$ denotes the length of stay in State 1, while $\Delta_1 = I(Z \leq C)$ is the corresponding indicator; $Y = \min(T, C)$ denotes the overall lifetime and $\Delta_1 = I(Z \leq C)$ the respective censoring indicator. Patients that are not in the absorbing state at the end of follow-up contribute with right censored observations. Here C denotes the potential censoring time that is assumed independent of the process, i.e. that C and (Z, Y) are independent.

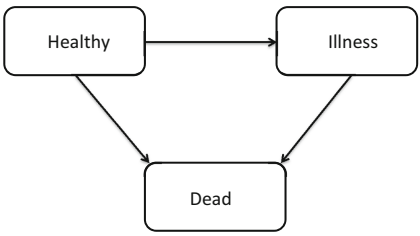

Fig. 1. The progressive illness-death model.

One of the main goals in clinical applications of multi-state models is the estimation of transition probabilities. These quantities have provided increasing interest as they allow for long-term predictions of the process. Several applications of the alternative representations of the Kaplan-Meier estimator are illustrated below to introduce estimators with lower variability for the transition probabilities.

4.1 The Aalen-Johansen Estimator

For two states h, j and $s < t$, introduce the so-called transition probabilities $p_{hj}(s,t) = P(X(t) = j \mid X(s) = h)$. Aalen and Johansen [5] introduced a

nonparametric estimator of the transition probabilities for Markovian models. Their proposal can be thought of as the generalization of the Kaplan-Meier approach for the simple mortality model (model with two states and a single transition). Explicit formulae of the Aalen-Johansen estimator for the illness-death model are available [18].

Let $Z_{(i)}$ denote the ordinal statistics obtained from the Z_i's and let $\Delta_{1[i]}$ be the i-th concomitant. Similarly, introduce the $Y_{(i)}$ ordinal statistics obtained from the Y_i's and $\Delta_{[i]}$ be its i-th concomitant. According to this notation, the Aalen-Johansen estimator for the transition probability $p_{11}(s,t)$ is the Kaplan-Meier estimator defined as:

$$\widehat{p}_{11}^{AJ}(t) = \prod_{s<Z_{(i)}\leq t} \left(1 - \frac{\Delta_{1[i]}}{nM_{0n}(Y_{(i)})}\right) \tag{8}$$

and

$$\widehat{p}_{22}^{AJ}(t) = \prod_{s<Z_{[i]}\leq t, Z_{(i)}\leq Y_{(i)}} \left(1 - \frac{\Delta_{[i]}}{nM_{1n}(Y_{(i)})}\right) \tag{9}$$

where $M_{0n}(y) = \frac{1}{n}\sum_{i=1}^{n} I(Z_i \geq y)$ and $M_{1n}(y) = \frac{1}{n}\sum_{i=1}^{n} I(Z_i < y \leq Y_i)$. Finally, the transition probability $p_{12}(s,t)$ can be estimated by plug-in methods using the formulae in [18].

The introduction of this representation of the Aalen-Johansen estimator allows the use of presmoothing in the Aalen-Johansen estimator. As usual, the new estimators are obtained by replacing the censoring indicator variables with smooth functions that can be estimated parametrically or nonparametrically. Further details can be seen in [19]. These estimators are implemented in the statistical software R. The **mstate** package can be used to obtain estimates of the transition probabilities in any multi-state model while the **survidm** package [20] can only be used in the illness-death model.

4.2 Landmark Markov-Free Estimators

The Aalen-Johansen estimator is the usual estimator of the transition probabilities providing consistent estimates if the multi-state model is Markov. When the multi-state model is non-Markov, this is no longer the case. The Markov assumption states that future and past are independent given the present state. This means that the future evolution of the process is independent of the states already visited and the transition times between them. This simplifying assumption allows the construction of simple estimators but is violated in some applications [21]. This is a relevant observation, since the Aalen-Johansen estimator may be inconsistent if the process is non-Markovian [22]. This section briefly describes two alternative methods for estimating transition probabilities that do not rely on the Markov assumption. Both methods involve the concept of subsampling. The first estimator is due to de Uña-Álvarez and Meira-Machado [23] and uses a procedure based on (differences between) Kaplan-Meier estimators derived from a subset of the data consisting of all subjects observed to be in the

given state at the given time. The second method, known as landmark Aalen-Johansen, is due to Putter and Spitoni [24] and is obtained by combining the Aalen-Johansen estimate of the state occupation probabilities derived from the same subset. Both approaches provide consistent estimators for the target that are free of the Markov condition.

In the illness-death model, the five transition probabilities can be expressed as depending on the joint distribution of (Z, Y). For example, $p_{11}(s,t) = P(U > t \mid U > s)$, $p_{12}(s,t) = P(U \leq t, Y > t \mid U > s)$ and $p_{22}(s,t) = P(U \leq t, Y > t \mid U \leq s, Y > s)$, whereas the two remaining transition probabilities ($p_{13}(s,t)$ and $p_{23}(s,t)$) can be obtained from those ones. To estimate $p_{1j}(s,t)$, $j = 1, 2, 3$, [23] consider the subsample of all subjects observed to be in State 1 at time s (i.e., those with a sojourn time in State 1 greater than s; $U > s$), whereas $p_{2j}(s,t)$, $j = 2, 3$, are estimated using the subsample of all subjects observed to be in State 2 at time s (i.e., those that have observed a transition to State 2 before time s and are still in that same state at time s; $U \leq s, Y > s$). Then, $\widehat{p}_{11}^{LM}(s,t) = \widehat{S}_U^{(s)}(t)$, $\widehat{p}_{12}^{LM}(s,t) = \widehat{S}_T^{(s)}(t) - \widehat{S}_U^{(s)}(t)$ and $\widehat{p}_{22}^{LM}(s,t) = \widehat{S}_T^{[s]}(t)$ where $\widehat{S}_U^{(s)}(t)$ is the Kaplan-Meier estimator of survival of the sojourn time in State 1 computed over the subsample of all subjects observed to be in State 1 at time s; $\widehat{S}_T^{(s)}(t)$ is the Kaplan-Meier estimator of survival of the total time computed over the same subsample, and $\widehat{S}_T^{[s]}(t)$ is the Kaplan-Meier estimator of survival of the total time computed over the subsample of all subjects observed to be in State 2 at time s.

The Landmark Aalen-Johansen approach introduced in [24] also uses a procedure that is based on the same subsamples of data. Assuming those subsamples, the transition probabilities are then estimated using the so-called Aalen-Johansen estimator, which in turn provides now consistent estimates since they can be seen as an occupation probability (consistency follows from [25]).

Both landmark approaches can use the presmoothing ideas described above to reduce the variability of the estimators. Presmoothing is highly recommended in landmark estimators since these methods lead to estimators that have greater variability due to their reliance on smaller datasets (subsampling).

5 Real Data Illustration

In this section, proposed methods are applied to two real data sets. The first comes from a bladder cancer study, whereas the second comes from a colon cancer study. Both data sets are available as part of the R **survival** package.

The present study aims to investigate multiple recurrence times in 85 patients with bladder tumors. To accomplish this, the estimated Kaplan-Meier survival curve (labeled as KM) and its presmoothed version (labeled as PKM) for the first and fourth recurrence times were obtained. Figure 2 shows the correspond-

ing survival plots. As expected, the presmoothed method has less variability than the KM estimator, which has fewer jump points as t increases. The extra jump points of the presmoothed estimator (PKM) correspond to patients with censored values of the first recurrence (left hand plot) and fourth recurrence (right hand plot). Presmoothing was obtained using a logistic regression model.

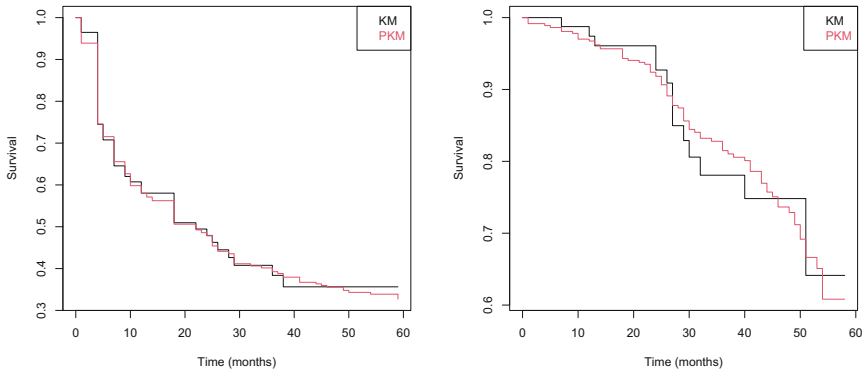

Fig. 2. Estimates of the survival for the first recurrence (left) and forth recurrence (right) using the original Kaplan-Meier (KM) estimator and the presmoothed Kaplan-Meier estimator. Bladder recurrence cancer data.

Data from a colon cancer clinical trial is also used, where patients were monitored from the date of cancer diagnosis until censoring or death from colon cancer. From the total of 929 patients, 468 developed a recurrence and among these 414 died. Thirty eight patients have died of causes unrelated to their disease and without evidence of recurrence. In addition to the two event times (recurrence and death) and corresponding censoring indicator functions, a vector of covariates are available.

The aim with this application is to illustrate the differences between the estimated transition probabilities from the Aalen-Johansen estimator (AJ), the Landmark estimator (LM), and the semiparametric estimators based on presmoothing (PAJ, PLM). Again, parametric presmoothing based on the logistic model is used. The left hand side of Fig. 3 depicts the estimates of $p_{12}(s = 1825; t)$ as functions of the time (for a fixed value of $s = 1825$, i.e., 5 years). Since the recurrence state (State 2) is transient, this curve is first increasing and then decreasing. The estimated curves reveal the lower variability of the presmoothed estimators. The PLM estimator seems to be the one that obtains more reliable curves. Similar conclusions can be obtained from the plots depicted on the right hand side of Fig. 3, in which the transition probability $p_{23}(s, t)$ is estimated through the four methods again for $s = 1825$.

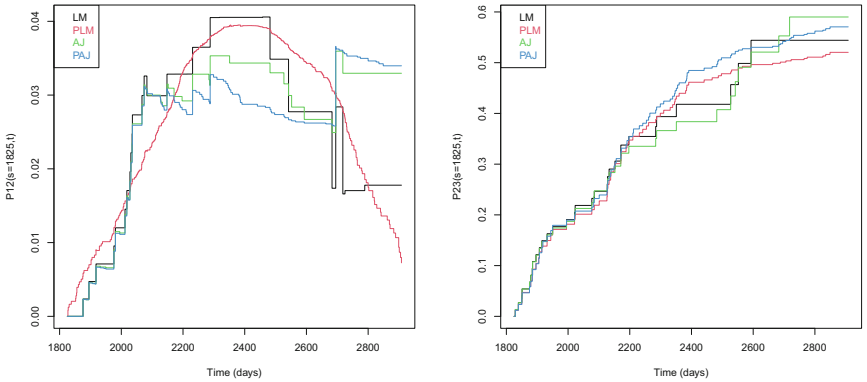

Fig. 3. Estimated transition probabilities for $p_{12}(s,t)$, (left), and for $p_{23}(s,t)$ for $s = 1825$. Colon cancer data.

6 Discussion

In this work, different representations of the Kaplan-Meier estimator were introduced. Based on one of these representations, new estimators were introduced for quantities of interest in the context of multi-state survival analysis. Specifically, estimators for the transition probabilities were introduced, although the same approach can be used to estimate other relevant quantities in medical applications, such as the bivariate distribution function for censored gap times and the cumulative incidence function.

The proposed estimators can be extended to multi-state models beyond the progressive illness-death model. Specifically, it is possible to apply the new method based on the computation of Kaplan-Meier curves in subsamples to any progressive multi-state model.

References

1. Kaplan, E.L., Meier, P.: Nonparametric estimation from incomplete observations. J. Am. Stat. Assoc. **53**, 457–481 (1958)
2. Lynden-Bell, D.: A method of allowing for known observational selection in small samples applied to 3CR quasars. Monthly Notices Royal Astron. Soc. **2**, 95–118 (1971)
3. Turnbull, B.W.: Nonparametric estimation of a survivorship function with doubly censored data. J. Am. Stat. Assoc. **69**, 169–173 (1974)
4. Aalen, O.O.: Nonparametric estimation of partial transition probabilities in multiple decrement models. Ann. Stat. **6**, 534–545 (1978)
5. Aalen, O.O., Johansen, S.: An empirical transition matrix for nonhomogeneous Markov chains based on censored observations. Scand. J. Stat. **5**, 141–150 (1978)
6. Fleming, T.R.: Nonparametric estimation for nonhomogeneous Markov processes in the problem of competing risks. Ann. Stat. **6**, 1057–1070 (1978)
7. Johansen, S.: The product limit estimator as maximum likelihood estimator. Scand. J. Stat. **5**(4), 195–199 (1978)

8. Feltz, C.J., Dykstra, R.L.: Maximum likelihood estimation of the survival functions of N stochastically ordered random variables. J. Am. Stat. Assoc. **80**(392), 1012–1019 (1985)
9. Borgan, Ø.: The Kaplan-Meier estimator. In: Armitage P, Colton T, eds. Encyclopedia of biostatistics, pp. 2154–2160. Wiley, Chichester (1998)
10. Satten, G.A., Datta, S.: The Kaplan-Meier estimator as an inverse-probability-of-censoring weighted average. Am. Stat. **55**(3), 207–210 (2001)
11. Rodríguez-Álvarez, M.J., Meira-Machado, L., Abu-Assi, E., Raposeiras-Roubín, S. (2016). Nonparametric estimation of time-dependent ROC curves conditional on a continuous covariate. Stat. Med. **35**(7), 1090–1102 (2016)
12. Meira-Machado, L., de Uña-Álvarez, J., Datta, S.: Nonparametric estimation of conditional transition probabilities in a non-Markov illness-death model. Comput. Stat. **30**, 377–397 (2015)
13. Meira-Machado, L., Sestelo, M.: Estimation in the progressive illness-death model: a nonexhaustive review. Biom. J. **61**(2), 245–263 (2019)
14. Dikta, G.: On semiparametric random censorship models. J. Stat. Plann. Inf. **66**, 253–279 (1998)
15. Cao, R., López, de Ullibarri, I., Janssen, P., Veraverbeke, N.J.: Presmoothed Kaplan-Meier and Nelson-Aalen estimators: Nonparamet. Stat **17**, 31–56 (2005)
16. Andersen, P.K., Borgan, O., Gill, R.D., Keiding, N.: Statistical Models Based on Counting Processes. Springer, New York (1993). https://doi.org/10.1007/978-1-4612-4348-9
17. Meira-Machado, L., de Uña-Álvarez, J., Cadarso-Suárez, C., Andersen, P.K.: Multi-state models for the analysis of time to event data. Stat. Methods Med. Res. **18**, 195–222 (2009)
18. Borgan, Ø.: The Kaplan-Meier estimator. In: Armitage P, Colton T. (eds.) Encyclopedia of biostatistics, pp. 5–10. Wiley, Chichester (1998)
19. Moreira, A., de Uña-Álvarez, J., Meira-Machado, L.: Presmoothing the Aalen-Johansen estimator in the illness-death model. Electr. J. Stat. **7**, 1491–1516 (2013)
20. Soutinho, G., Sestelo, M., Meira-Machado, L.: survidm: an R package for inference and prediction in an illness-death model. R J. **13**(2), 70–89 (2021)
21. Soutinho, G., Meira-Machado, L.: Methods for checking the Markov condition in multi-state survival data. Comput. Stati. **37**(2), 751–780 (2022)
22. 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)
23. de Uña-Álvarez, J., Meira-Machado, L.: Nonparametric estimation of transition probabilities in the non-Markov illness-death model: a comparative study. Biometrics **71**, 141–150 (2015)
24. Putter, H., Spitoni, C.: Non-parametric estimation of transition probabilities in non-Markov multi-state models: the landmark Aalen-Johansen estimator. Stat. Methods Med. Res. **27**(7), 2081–2092 (2018)
25. Datta, S., Satten, G.A.: Validity of the Aalen-Johansen estimators of stage occupation probabilities and Nelson-Aalen estimators of integrated transition hazards for non-Markov models. Statist. Probab. Lett. **55**, 403–411 (2001)