On Nonparametric Hazard Estimation

Hobbs BP*
Department of Biostatistics, University of Texas M.D. Anderson Cancer Center, Houston, TX, USA

Abstract

The Nelson-Aalen estimator provides the basis for the ubiquitous Kaplan-Meier estimator, and therefore is an essential tool for nonparametric survival analysis. This article reviews martingale theory and its role in demonstrating that the Nelson-Aalen estimator is uniformly consistent for estimating the cumulative hazard function for right-censored continuous time-to-failure data.

Keywords: Asymptotic theory; Counting process; Hazard estimation; Martingale theory; Nelson-Aalen estimator; Time-to-failure analysis

Introduction

The Nelson-Aalen estimator Nelson [1,2] provides the foundation for the ubiquitous Kaplan-Meier survival estimator which consists of its product-integral. This article reviews martingale theory and its role in demonstrating that the Nelson-Aalen estimator is asymptotically uniformly consistent for the cumulative hazard function for right-censored continuous time-to-failure data and demonstrates its application using simulation.

Fundamental principles

Nonparametric statistical models for censored data were developed using counting processes which decompose into a martingales and integrated intensity processes. Martingale methods facilitate direct evaluation of small and large sample properties of hazard estimators for right censored failure time data. This section provides formal definition of a martingale process as well as reviews several intrinsic properties that are used to study small and large sample properties of hazard estimators for right censored failure time data.

Counting processes

A family of sub-σ-algebras \{F_t; t \geq 0\} of a σ-algebra F is called increasing if s \leq t implies F_s \subseteq F_t. An increasing family of sub-σ-algebras is called a filtration. When \{F_t; t \geq 0\} is a filtration, the σ-algebra \( \bigcap_{s,t} F_{s,t} \) is usually denoted by \( F_{\infty} \). The corresponding limit from the left, \( F_{\infty}^{-} \), is the smallest σ-algebra containing all the sets in \( \bigcup_{s \geq t} F_{s} \) and is written \( \sigma \left( \bigcup_{s \geq t} F_{s} \right) \). A filtration \( \{F_t; t \geq 0\} \) is right-continuous if, for any t, \( F_{t} = F_{t+} \). A stochastic basis is a probability space (Ω, F, P) with a right-continuous filtration {F_t: t ≥ 0}, and is denoted by (Ω, F, F_t). When \( \{F_t; t \geq 0\} \) is a filtration, the σ-algebra \( \bigcap_{s \geq t} F_{s} \) is usually denoted by \( F_{\infty}^{-} \). The corresponding limit from the left, \( F_{\infty}^{-} \), is the smallest σ-algebra containing all the sets in \( \bigcup_{s \geq t} F_{s} \) and is written \( \sigma \left( \bigcup_{s \geq t} F_{s} \right) \).

Martingales

Let \( X = \{X(t); t \geq 0\} \) denote a right-continuous stochastic process with left-hand limits and \( \{F_t; t \geq 0\} \) a filtration, defined on a common probability space. X is called a martingale with respect to \( \{F_t; t \geq 0\} \) if, X is adapted to \( \{F_t; t \geq 0\} \), \( E[X(t)] < \infty \) for all \( t < \infty \), and \( E[X(t + s)|F_t] = X(t) \) a.s. for all \( s \geq 0, t \geq 0 \). Thus, a martingale is essentially a process that has no drift and whose increments are uncorrelated. If \( E[X(t + s)|F_t] \geq X(t) \) a.s. X is a submartingale. Two fundamental properties of martingales are, for any \( h > 0 \),

\[ E[X(t)|F_{t-h}] = X(t-h), \quad \text{and} \quad E[X(t) - X(t-h)|F_{t-h}] = X(t) - X(t-h) - X(t-h) = 0. \]

Predictable processes

The stochastic process X is said to be predictable with respect to filtration \( F_t \) if for each t, the value of X(t) is specified by \( F_t \), and therefore is \( F_t \)-measurable [4]. Theorem A in Appendix A is the version of the Doob-Meyer Decomposition Theorem provided by Fleming and Harrington [3], which states that for any right-continuous nonnegative submartingale X there is a unique increasing right-continuous predictable process A, such that \( A(0) = 0 \) and A−X is a martingale. Also, there is a unique process A so that for any counting process N, with finite expectation, A−N is a martingale. This is shown in the Corollary A.2 [3]. The process A in Corollary A.2 is referred to as the compensator for the submartingale X.

Square integrable martingales

A martingale, X(t) is called square integrable if \( E[X(t)]^2 = \text{var}[X(t)] < \infty \) for all \( t \leq \tau \), or equivalently, if \( E[X(t)^2] < \infty \) [4]. The variance of a square integrable martingale X(t) is estimated using the predictable variation process,

\[ \langle X \rangle(t) = \int_0^t \text{var}(dX(u)) |F_u \rangle. \]

For martingales \( X, Y \),

\[ \langle X, Y \rangle(t) = \int_0^t \text{cov}(dM(u), dM(u)) |F_u \rangle \]

is the predictable covariation process of \( X \) and \( Y \). If \( \langle X(t), X(t) \rangle \) is a martingale, then \( X(t) \) and \( Y(t) \) are orthogonal martingales [3]. Suppose \( X, Y \) are orthogonal martingales, for all \( t \neq j \). Then \( \sum_{t \geq 0} X(t)^2 = \sum_{t \geq 0} X(t)^2 \).

*Corresponding author: Hobbs BP, Department of Biostatistics, University of Texas M.D. Anderson Cancer Center, Houston, TX, USA, Tel: 713-794-4633; E-mail: bphobbs@mdanderson.org

Received May 29, 2015; Accepted June 10, 2015; Published June 18, 2015


Copyright: © 2015 Hobbs BP. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
Localizaton

A nonnegative random variable \( \tau \) is a stopping time with respect to filtration \( \{F_t\} \) if \( [\tau \leq t] \in F_t \) for all \( t \geq 0 \). An increasing sequence of random times \( \tau_m \), \( m = 1, 2, ... \) is a localizing sequence with respect to a filtration \( \{F_t\} \) if for each \( n \), \( X_n = [X(t \wedge \tau_m) : 0 \leq t \leq \infty] \) is a \( \mathbb{F}_m \)-martingale (submartingale) with respect to \( \{F_t\} \). Therefore, the change in \( M(t) = N(t) - A(t) \) over an interval \([s, t]\) is a martingale with respect to \( \{\mathbb{F}_t\} \).

Martingale approach to censored failure time data

Suppose \( T \) and \( U \) are nonnegative, independent random variables, and assume that the distribution of \( T \) has a density. Define a filtration \( \{\mathbb{F}_t\} \) by \( \mathbb{F}_t = \sigma\{N(u), I(X \leq u, \delta = 0) : 0 \leq u \leq t\} \), the indicator for the event of an uncensored observation. Therefore, \( M(t) = N(t) - A(t) \) is a martingale with respect to \( \{\mathbb{F}_t\} \) and \( U \) are independent.

Thus, we can conclude that \( A(t) < \infty \), and \( A(t) \) is locally bounded for all \( t \geq 0 \). Notice, \( \mathbb{E}[M(t)] = \mathbb{E}[N(t)] - \mathbb{E}[A(t)] \).

Nelson-Aalen estimator

Suppose time-to-failure data are collected over a finite interval \([0, \tau]\) on \( n \) subjects with independent failure times. The Nelson-Aalen estimator is a nonparametric estimator of their common cumulative hazard function, \( \hat{\Lambda}(t) \), in the presence of right censoring. Suppose \( T \) and \( U \), are the failure and censoring times and \( N_i(t) \) the observed counting process for the \( i \)th subject. Let \( (Y_i(t), t \geq 0) \) denote a process such that \( Y_i(t) = 1 \) if and only if the \( i \)th subject is uncensored and has survived to time \( t \). It is referred to as the at risk process and assumed left-continuous. For each \( t > 0 \), let \( F_{-t} = \sigma\{N(u), Y_i(u), i = 1, ..., n : 0 \leq u \leq t\} \) denote the filtration up to, but not including \( t \). Let \( N_i(t) = \sum_{s \leq t} Y_i(s) \) denote the aggregate processes that count the numbers of total failures and at risk in the interval \([0, t]\), respectively. Furthermore, let \( (Y_i(t), t \geq 0) \) indicate whether at least one subject is at risk at time \( t \). Thus, \( N_i(t) \) is given by \( \hat{\Lambda}(t) = \int_0^t \frac{J(u)}{1 - F(u)} \, du \), for continuous distribution function, \( F(t) = P[T \leq t] \). Note that by assuming continuous time we have assumed that no two of the counting processes \( N_i(t), i = 1, ..., n \), jump at the same moment.

Asymptotic uniform consistency of the Nelson-Aalen estimator

In this section, we demonstrate how to prove that \( \hat{\Lambda}(t) \) is an uniformly consistent estimator of \( \Lambda(t) = \int_0^t \frac{dF(u)}{1 - F(u)} \), for continuous distribution function, \( F(t) = P[T \leq t] \). First, note that by assuming continuous time we have assumed that no two of the counting processes \( N_i(t), i = 1, ..., n \), jump at the same moment.

Second, \( \Lambda(t) < \infty \), thus, \( F(t) < 1, \forall t \in [0, \tau] \).

First, notice, \( \hat{\Lambda}(t) - \Lambda(t) \leq \int_0^t \frac{J(u)}{1 - F(u)} \, du - \int_0^t \frac{dF(u)}{1 - F(u)} \, du = \int_0^t \frac{J(u)}{Y_i(u)} \, du - \int_0^t \frac{dF(u)}{1 - F(u)} \, du \).

Second, using the same argument as in Appendix B, we demonstrated in Section 2 that \( N_i(t) \) has compensator \( A(t) = \int_0^t \frac{dF(u)}{1 - F(u)} \). A simple alteration of (B.1) in Appendix B can be used to show that \( A(t) = \int_0^t \frac{dF(u)}{1 - F(u)} \).

Now, we must show that \( A(t) < \infty \) and \( A(t) \) is locally bounded \( \forall t \in [0, \tau] \). For all fixed \( n = 1, 2, ... \) and all \( t \in [0, \tau] \), it follows that

\[
\hat{\Lambda}(t) = \left( \frac{1}{Y_i(u)} - \frac{1}{F(u)} \right) \left( \int_0^t \frac{J(u)}{1 - F(u)} \, du - \int_0^t \frac{dF(u)}{1 - F(u)} \, du \right) \geq 0.
\]

The second regularity condition is necessary. First, \( \inf_{t \in [0, \tau]} \frac{1}{Y_i(u)} - \frac{1}{F(u)} \geq 0 \), implies that \( \hat{\Lambda}(t) > 0 \), \( \forall t \in [0, \tau] \).

Reference:

Using the variances of increments of $M(t)$, one can show that, 
\[
\mathbb{E}[dM(t)^2] = \mathbb{E}[dN(t)^2] = \mathbb{E}[dA(t)^2] = dt 
\]
Thus, taking the predictable variation process equal to its compensator: 
\[
\langle M(t) \rangle = A(t) = \int_0^t Y_u du
\]
It follows from (4) that 
\[
P \left[ \mathbb{E}[dN(t)^2] = \mathbb{E}[dM(t)^2] \right] > \mathbb{E}[dA(t)^2] = dt
\]
From our first regularity condition we have $Y(t) \to \infty$ in probability, as $n \to \infty$. Therefore, $P(\mathbb{E}[dM(t)^2] > \mathbb{E}[dA(t)^2]) \to 0$ for any $\eta > 0$. Thus, all that is left to show is that 
\[
sup_{\eta > 0} \mathbb{E}[\left| \int [1-J(u)] \frac{dF(u)}{1-F(u)} \right|] \to 0
\]
Noting that $\int [1-J(u)] \frac{dF(u)}{1-F(u)}$ is bounded by $\mathbb{E}(s) - 1 (Y(s) > 0) \mathbb{E}(s)$ e.g. 
\[
\int [1-J(u)] \frac{dF(u)}{1-F(u)} = \mathbb{E}(s) - 1 (Y(u) > 0) \mathbb{E}(s) > 0)
\]
From our first regularity condition we can conclude that 
\[
\lim_{x \to \infty} \mathbb{E}[\left| \int [1-J(u)] \frac{dF(u)}{1-F(u)} \right|] = 0
\]
Therefore, we have, 
\[
sup_{\eta > 0} \mathbb{E}(s) - 1 (Y(s) > 0) \mathbb{E}(s) \to 0 \text{ in probability as } n \to \infty. Therefore, the Nelson-Aalen estimator is asymptotically uniformly consistent for the cumulative hazard function under regularity conditions: $\inf_{t \geq s} Y(s) \to \infty$ in probability as $n \to \infty$ and $\Lambda(t) < \infty$.

**Simulation Study**

This section demonstrates the martingale approach for analysis of right-censored failure time data using simulation. Let $A_1, A_2, ...$ denote independent uniform $(0, \tau)$ random variables where $\tau$ is a known constant, and let $X_i = \left( A_i, t \right)$, $i=1, ..., n$. $N(t)$ counts observed failures for the ith subject observed over the interval $(0, \tau_i]$, where $\tau_i$ is a constant $0 < \tau_i \leq \tau$, $i=1, ..., n$. We assume that the intensity function $\lambda(t)$ of $N(t)$ with respect to the filtration $\mathcal{F}$ is $\sigma(N(u), Y(u'), X(u'))$, $i=1, ..., n; 0 \leq u \leq t \leq \tau$ is 
\[
\lambda(t) = Y(t) \exp[X(t) \beta] \alpha
\]
for $0 < t \leq \tau$, $i=1, ..., n$, and $\alpha > 0$. We simulated 30 realizations of the process $[N(t), 0 < t < \tau]$, and the corresponding martingale process $\lambda(t)$ and the Nelson-Aalen estimator $\Lambda(t)$ in the above example were computed for $n=10$. We approximated continuous time by partitioning $[0, 10]$ into disjoint intervals, $t_i$ of length $d_i=0.1$. Now it follows that $E[dN(t)] = \lambda(t) dt$ at each $t_i \in [0, \tau]$. For each subject, the process was simulated by iterating through each time interval $t_i$ within each subject. At each $t_i$ we draw a single sample, $dN(t)$, from the Poisson density with rate parameter $E[dN(t)]$. If $dN(t) = 0$, then we have observed the failure time for the ith subject to be $t_i$. Therefore, we set counting processes $N(k) = 1$ and $\lambda(k) = 0$, $0 \leq k \leq \tau$.

Figure 1 provides the resulting expected and simulated mean aggregated counting process $N(t)$ (right panel) as well as the simulated mean martingale (left panel). As expected, the simulated mean aggregated counting process $N(t)$ mirrors closely its expectation. Furthermore, the corresponding simulated mean aggregated martingale illustrates a process that has no random drift.

**Discussion**

Martingale characterizations play a critical role in the study of large sample properties of the Nelson-Aalen and thereby Kaplan-Meier estimators. This article reviewed martingale theory and its role in demonstrating that the Nelson-Aalen estimator is asymptotically uniformly consistent for the cumulative hazard function for right-censored continuous time-to-failure data and demonstrates its application using simulation.

**Acknowledgement**

The author was partially supported by NIH/NCI grants R01CA157458 and P03CA016672.

**References**


