Section 12

Lecture 6

Features of *the martingale $I(t)$*

- The *predictable variation*

$$\langle I\rangle(t) = \langle \int GdM\rangle(t) = \int_0^t G^2(s)d\langle M\rangle(s).$$

- The *optional variation process*

$$[I](t) = \Big[\int GdM\Big](t) = \int_0^t G^2(s)d[M](s).$$

I might ask you to prove this in a future homework.

Nelson-Aalen's small sample behaviour

Define the indicator $J(t) = I(Z(t) > 0)$ and

$$W(t) = \begin{cases} \frac{J(t)}{Z(t)}, Z(t) > 0, \\ 0, Z(t) = 0. \end{cases}$$

Define $H^*(t) = \int_0^t J(s)\alpha(s)ds$.

Theorem (Small sample behaviour of Nelson-Aalen)

$\hat{H}(t)$ is an unbiased estimator of $H^*(t)$ with variance $Var(\hat{H}(t) - H^*(t)) = \mathbb{E}(\int_0^t W(s)^2 dN(s))$.

Argument for the Nelson-Aalen estimator

Remember that the Nelson-Aalen estimator is
$\hat{H}(t) = \sum_{T_j \leq t} \frac{1}{Z(T_j)} \equiv \sum_{T_j \leq t} \Delta \hat{H}(T_j)$,

- Suppose that our counting process N is an aggregated process obtained from observation of n individual processes.
- The Nelson-Aalen estimator $\hat{H}(t)$ can be re-written as a counting process integral

$$\begin{aligned}\int_0^t W(s)dN(s) &= \int_0^t W(s)\lambda(s)ds + \int_0^t W(s)dM(s) \\ &= \int_0^t J(s)\alpha(s)ds + \int_0^t W(s)dM(s)\end{aligned}$$

Argument for small sample behaviour of the Nelson-Aalen estimator

- We have defined $H^*(t) = \int_0^t J(s)\alpha(s)ds$, and thus

$$\hat{H}(t) - H^*(t) = \int_0^t W(s)dM(s).$$

  That is, a mean zero martingale.

- $\mathbb{E}\{\hat{H}(t) - H^*(t)\} = 0$ Thus, $\hat{H}(t)$ is an unbiased estimator of $H^*(t)$, but not necessarily of H. Remember that $\text{Var}(M(t)) = \mathbb{E}\{[M](t)\} = \mathbb{E}\{N(t)\}$, and using the result from Slide 155,

$$\text{Var}(\hat{H}(t) - H^*(t)) = \mathbb{E}\{[\hat{H} - H^*](t)\} = \mathbb{E}\{\int_0^t W(s)^2 dN(s)\}$$

Remark on $H(t)$ and $H^*(t)$

- Suppose we restrict ourselves to an interval $[0, \tau]$ such that $S(\tau) > 0$.
- Then $\lim_{n \to \infty} H^{*(n)}(t) = H(t)$.

Large sample properties

Suppose N is obtained by aggregating n individual processes. We will then use the Martingale central limit theorem.

Theorem (Rebolledo's martingale central limit theorem)

Let $V(t)$ be a strictly increasing continuous function with $V(0) = 0$. Let $\tilde{M}^{(n)}$, $n \geq 1$ be a sequence of mean zero martingales on $[0, \tau]$, and let $\tilde{M}_{\epsilon}^{(n)}$ be the martingale containing all the jumps of $\tilde{M}^{(n)}$ larger than a given $\epsilon > 0$. If

- $\langle \tilde{M}^{(n)} \rangle(t) \xrightarrow{P} V(t)$ *for all* $t \in [0, \tau]$ *as* $n \to \infty$
- $\langle \tilde{M}_{\epsilon}^{(n)} \rangle(t) \xrightarrow{P} 0$ *for all* $t \in [0, \tau]$ *and* $\epsilon > 0$ *as* $n \to \infty$,

then the sequence $\tilde{M}^{(n)}$ converges in distribution to the mean zero martingale U given by $U(t) = B(V(t))$, where $B(\cdot)$ is a Wiener process (Brownian motion).

What you must know about the Martingale Central Limit Theorem

Under fairly general assumptions, a sequence of martingales $\tilde{M}^{(n)}$ will converge in distribution to a mean zero Guassian martingale $U(t) = B(V(t))$, where $V(t)$ is a strictly increasing continuous function with $V(0) = 0$ and $B(t)$ is a Wiener process (formal definition in the next slide, as repetition). Thus, U inherits the properties of a Wiener process:

- $U(0) = 0$
- $U(t) - U(s)$ are normally distributed with mean 0 and variance $V(t) - V(s)$.
- independent increments.
- continuous sample paths.

Wiener process (Brownian motion, repetition)

Definition (Wiener process)

The $B = \{B(t) : t \in [0, \tau]\}$ is a process satisfying

- $B(0) = 0$,
- independent increments, that is, $B(t + u) - B(t)$ $u \geq 0$ are independent of $B(s)$, for all $s \leq t$,
- Gaussian increments, that is, $B(t + u) - B(t) \sim \mathcal{N}(0, u)$,
- continuous sample paths, that is, $B(t)$ is continuous in t.

Sufficient conditions for the Martingale CLT

We will study the limiting behaviour of sequences of stochastic integrals on the form

$$\int_0^t G^{(n)}(s)dM^{(n)}(s),$$

where G is a predictable process and M is counting process martingale. More generally, we will study sums of $1, \dots, k$ such integrals

$$\sum_{j=1}^{k} \int_0^t G_j^{(n)}(s)dM_j^{(n)}(s).$$

In this setting, Let $V(t) = \int_o^t v(s)ds$. For the martingale CLT to hold it is sufficient that the following two conditions hold (under regularity conditions):

- $\sum_{j=1}^{k} \left(G_j^{(n)}(s)\right)^2 \lambda_j^{(n)}(s) \xrightarrow{P} v(s) > 0$ for all $j = 1, \dots, k$ $s \in [0, \tau]$, as $n \to \infty$.
- $G_j^{(n)}(s) \xrightarrow{P} 0$ for all $j = 1, \dots, k$ amd $s \in [0, \tau]$, as $n \to \infty$.

Properties of the Nelson-Aalen estimator

Theorem (Large sample porperties of Nelson-Aalen)

*$\sqrt{n}(\hat{H}(t) - H^*(t))$ converges in distribution to a mean zero martingale with variance $\sigma^2(t) = \int_0^t \frac{\alpha(s)}{z(s)} ds$.
At a particular t, the Nelson-Aalen estimator is approximately normal, and a consistent estimator of the variance of $\hat{H}(t)$ is*

$$\hat{\sigma}^2(t) = \int_0^t W(s)^2 dN(s).$$

Note that we can also the survival function from the cumulative hazard function using the estimator

$$\hat{\hat{S}}(t) = \exp\{-\hat{H}(t)\},$$

which is consistent by the continuous mapping theorem: for a continuous function g, if $X^{(n)} \xrightarrow{D} X$ then $g(X^{(n)}) \xrightarrow{D} g(X)$.

Large Sample Properties of Nelson-Aalen

Remember that $W(s) = \frac{J(t)}{Z(t)}$ when $Z(t) > 0$. Consider the martingale

$$\sqrt{n}(\hat{H}(t) - H^*(t)) = \int_0^t \sqrt{n} W(s) dM(s).$$

Assume that $\frac{Z(t)}{n} \xrightarrow{P} z(t) > 0$ for all $t \in [0, \tau]$ as $n \to \infty$.
When $J(t) = 1$, we have that for $G(t) = \sqrt{n} J(t)/Z(t)$ and $\lambda(t) = Z(t)\alpha(t)$,

$$G(t)^2 \lambda(t) = \frac{J(t)\alpha(t)}{Z(t)/n} \xrightarrow{P} \frac{\alpha(t)}{z(t)}$$

$$G(t) = \frac{1}{\sqrt{n}} \frac{J(t)}{Z(t)/n} \xrightarrow{P} 0,$$

for all $t \in [0, \tau]$ as $n \to \infty$.
So, here we can use the martingale clt...

How to obtain confidence intervals

We can therefore write standard confidence intervals at level η by $\hat{H}(t) \pm z_{1-\eta/2}\hat{\sigma}(t)$, where $z_{1-\eta/2}$ is the $1 - \eta/2$ fractile of the standard normal distribution.

Time between first and second birth

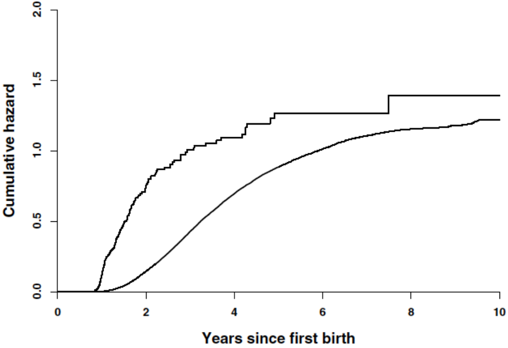


Fig. 3.1 *Nelson-Aalen estimates for the time between first and second births. Lower curve: first child survived one year; upper curve: first child died within one year.*

Time between first and second birth

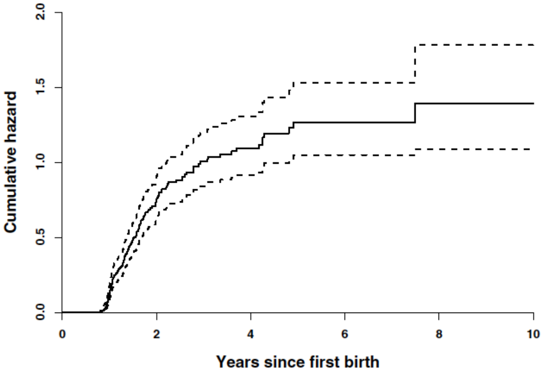


Fig. 3.2 *Nelson-Aalen estimates for the time between first and second births with log-transformed 95% confidence intervals for women who lost their first child within one year of its birth.*

Example Divorce Rates[21]

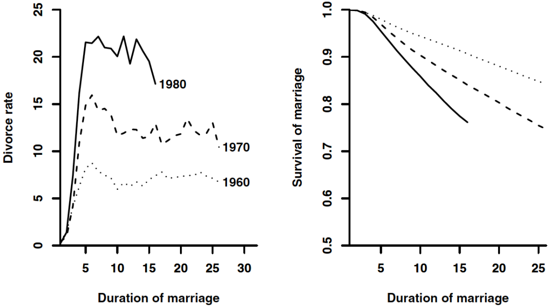
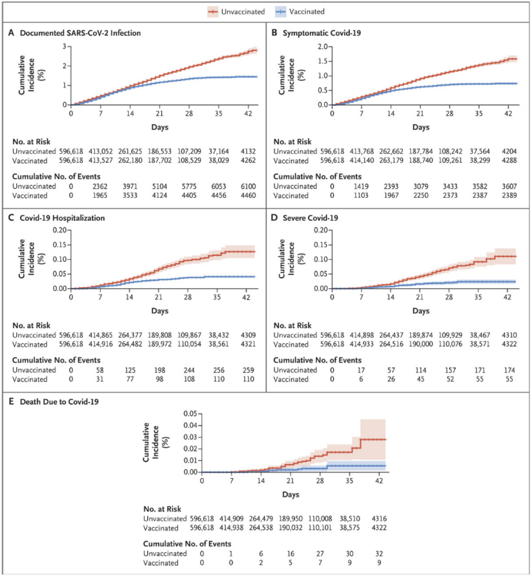


Fig. 1.4 *Rates of divorce per 1000 marriages per year (left panel) and empirical survival curves (right panel) for marriages contracted in 1960, 1970, and 1980. (Based on data from Statistics Norway.)*

[21] from Aalen et al (2008)

Example COVID in real-life



Angioplasty is standard of care

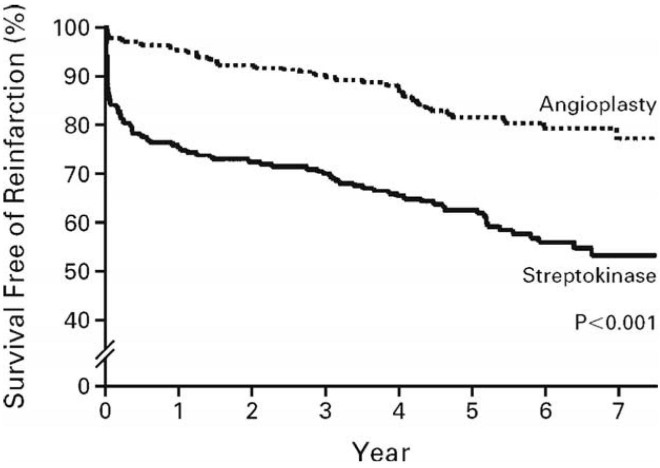


Fig. 1.5 *Survival curves for two treatments of myocardial infarction. Figure reproduced with permission from Zijlstra et al. (1999). Copyright The New England Journal of Medicine.*

Kaplan-Meier estimator

The most common estimator in the survival analysis literature is the Kaplan-Meier estimator.
Suppose we partition $[0, t]$ into small intervals defined by $0 < t_0, t_1, \cdots t_K = t$, and use conditional probabilities to express

$$S(t) = \prod_{k=1}^{K} S(t_k \mid t_{k-1}), \quad S(v \mid u) = \frac{S(v)}{S(u)}, v > u.$$

Then we estimate $S(t_k \mid t_{k-1})$ by 1 if no event happened in $(t_{k-1}, t_k]$, and if an event happened, and if $T_j \in (t_{k-1}, t_k]$ we estimate it by
$1 - 1/Z(t_{k-1}) = 1 - 1/Z(T_j)$, where $Z(t) = \sum_{i=1}^{n} I(\tilde{T}_i \geq t)$.

Definition (The Kaplan-Meier estimator)

$$\hat{S}(t) = \prod_{T_j \leq t} \left\{ 1 - \frac{1}{Z(T_j)} \right\},$$

where $T_j \in \{T_j : D_j = 1\}$.

Relation between Nelson-Aalen and Kaplan-Meier

We shall see that the Nelson-Aalen estimator and the Kaplan-Meier estimator are closely related.

Let $T > 0$ be a random survival time. Hitherto, we have assumed that $S(t)$ is absolutely continuous. Now, we relax this assumption: suppose $S(t)$ is *cadlàg*. Then we can define the cumulative hazard as

$$H(t) = -\int_0^t \frac{dS(u)}{S(u-)}$$

where $S(u-)$ is the left limit of $S(u)$ and the right hand side is a Stieltjes integral (Now, however $\alpha(t) = dH(t)/dt$ does not necessarily exist). In differential form $dS(t) = -S(t-)dH(t)$ or more formally in integral form $S(t) = 1 - \int_0^t S(u-)dH(u)$.

- If $S(t)$ is absolutely continuous, then $dS(u) = -f(u)du$ and $S(u-) = S(u)$, and $H(t) = \int_0^t \frac{f(u)}{S(u)}du$.

Informally, think about $dS(t)$ as the increment of S in $[t, t + dt)$.