Restricted Mean Survival Time

1.1 Setup

Let T denote the time to event and C a random censoring time. Define:

$$U = \min(T, C), \qquad \delta = \begin{cases} 1 & \text{if observed,} \\ 0 & \text{if censored.} \end{cases}$$

The **restricted mean survival time** (RMST) $\theta(\tau)$ is the area under the survival curve $S(t) = \mathbb{P}(T > t)$ up to time τ:

$$\theta = \theta(\tau) = \int_0^\tau S(t)dt.$$

An estimator for $\theta(\tau)$ is given by:

$$\theta(\tau) = \int_0^\tau \hat{S}(t)dt,$$

where $\hat{S}(t)$ is the Kaplan-Meier (KM) estimator of the survival function.

1.2 Asymptotics

Proposition 1.1. Define $\mu(t, \tau)$ as the area under the survival curve from time t to the truncation time τ:

$$\mu(t, \tau) = \int_t^\tau S(u)du.$$

The standardized process $\sqrt{n}\big\{\hat{\theta}(\tau) - \theta(\tau)\big\}$ converges weakly to a Brownian motion:

$$\sqrt{n}\big\{\hat{\theta}(\tau) - \theta(\tau)\big\} \rightsquigarrow W\big\{\sigma^2_{\text{RMST}}(\tau)\big\},$$

with variance function:

$$\sigma^2_{\text{RMST}}(\tau) = \int_0^\tau \frac{\mu^2(t, \tau)\alpha(t)}{y(t)}dt.$$

Here $\alpha(t)$ is the (instantaneous) hazard at time t, and $y(t)$ is the probability of being at risk at time t:

$$y(t) = \operatorname*{plim}_{n\to\infty} n^{-1}Y(t) = \mathbb{P}(U > t).$$

♦

Proof. Note first that:

$$\frac{d\mu(t,\tau)}{dt} = -S(t)dt,$$

or $d\mu(t,\tau) = -S(t)dt$. Recall that the Kaplan-Meier and Nelson-Aalen estimators are asymptotically linked via:

$$\frac{\sqrt{n}\big\{\hat{S}(u) - S(u)\big\}}{-S(u)} = \sqrt{n}\big\{\hat{A}(u) - A(u)\big\} + o_p(1).$$

The standardized RMST process is:

$$\begin{aligned}
\sqrt{n}\big\{\hat{\theta}(\tau) - \theta(\tau)\big\} &= \int_0^\tau \sqrt{n}\big\{\hat{S}(u) - S(u)\big\}du \\
&= \int_0^\tau \frac{\sqrt{n}\big\{\hat{S}(u) - S(u)\big\}}{-S(u)} \cdot \{-S(u)du\} \\
&= \int_0^\tau \sqrt{n}\big\{\hat{A}(u) - A(u)\big\}d\mu(t,\tau) + o_p(1).
\end{aligned}$$

Integrating by parts:

$$\int_0^\tau \sqrt{n}\big\{\hat{A}(u) - A(u)\big\}d\mu(t,\tau) =$$

$$\cancelto{0}{\Big[\sqrt{n}\big\{\hat{A}(u) - A(u)\big\}\mu(t,\tau)\Big]_{t=0}^{t=\tau}} - \int_0^\tau \mu(t,\tau) \cdot \sqrt{n}d\big\{\hat{A}(u) - A(u)\big\}.$$

The first term on the right vanishes because the cumulative hazard is zero at time zero $\hat{A}(0) = A(0) = 0$, and $\mu(\tau,\tau) = 0$. Thus:

$$\sqrt{n}\big\{\hat{\theta}(\tau) - \theta(\tau)\big\} = -\int_0^\tau \mu(t,\tau) \cdot \sqrt{n}d\big\{\hat{A}(u) - A(u)\big\} + o_p(1). \tag{1.1}$$

The standardized Nelson-Aalen process is expressible as a zero-mean martingale:

$$\sqrt{n}\big\{\hat{A}(u) - A(u)\big\} = \int_0^u \frac{\sqrt{n}}{Y(t)}dM(t).$$

Here $dM(u) = dN(u) - \alpha(u)Y(u)du$ is the increment in the counting process martingale for $N(u)$, the **total** number of events by time u. Substituting

$$\sqrt{n}d\big\{\hat{A}(u) - A(u)\big\} = \frac{\sqrt{n}}{Y(u)}dM(u).$$

into (1.1) gives:

$$\sqrt{n}\big\{\hat{\theta}(\tau) - \theta(\tau)\big\} = -\int_0^\tau \frac{\mu(u,\tau)\sqrt{n}}{Y(u)}dM(u) + o_p(1).$$

Finally, taking the predictable variation:

$$
\begin{aligned}
\left\langle \sqrt{n}\big\{\hat{\theta}(\tau) - \theta(\tau)\big\} \right\rangle &= \int_0^\tau \frac{n\mu^2(t,\tau)}{Y^2(t)} d\langle M(t)\rangle + o_p(1) \\
&= \int_0^\tau \frac{n\mu^2(t,\tau)}{Y^2(t)} Y(t)\alpha(t)dt + o_p(1) \\
&= \int_0^\tau \frac{\mu^2(t,\tau)\alpha(t)}{n^{-1}Y(t)} dt + o_p(1) \\
&\xrightarrow{p} \int_0^\tau \frac{\mu^2(t,\tau)\alpha(t)}{y(t)} dt,
\end{aligned}
$$

■

1.3 Influence Function

Taking τ as fixed, the influence function expansion for the RMST is:

$$
\begin{aligned}
\sqrt{n}(\hat{\theta} - \theta) &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \psi_i + o_p(1), \\
\psi_i &= -\int_0^\tau \frac{\mu(t,\tau)}{y(t)} dM(t)
\end{aligned}
\tag{1.2}
$$

Here $\mu(t,\tau) = \int_t^\tau S(u)du$, $y(u) = \mathbb{P}(U > t)$, and:

$$dM(t) = dN(t) - Y(t)dA(t) = \sum_{i=1}^{n} dN_i(t) - Y_i(t)dA(t).$$

The influence function contributions can be estimated by:

$$\hat{\psi}_i = -\int_0^\tau \frac{\hat{\mu}(t,\tau)}{n^{-1}Y(t)} \left\{ dN_i(t) - Y_i(t)d\hat{A}(t) \right\},$$

where $\hat{\mu}(t,\tau) = \int_0^\tau \hat{S}(t)dt$ is the integral of the Kaplan-Meier curve $\hat{S}(t)$, and $d\hat{A}(t)$ is the increment of the Nelson-Aalen estimator:

$$d\hat{A}(t) = \frac{\sum_{i=1}^{n} dN_i(t)}{Y(t)}$$

1.4 Perturbation

The variance of $\hat{\theta}(\tau)$ can be estimated analytically via:

$$\hat{\mathbb{V}}\{\hat{\theta}(\tau)\} = \hat{\sigma}^2_{\text{RMST}}(\tau) = \int_0^\tau \frac{\hat{\mu}^2(t,\tau)}{n^{-1}Y(t)} d\hat{A}(t).$$

Perturbation provides an alternative applicable to more complex estimators. To estimate the sampling variance via perturbation:

1. Precompute the influence function contributions ψ_i for all subjects.
2. For B iterations, repeat the following:
   i. Draw n perturbation weights $w_i^{(b)} \overset{\text{IID}}{\sim} N(0,1)$.
   ii. Calculate and store:
   $$D_b = \frac{1}{n}\sum_{i=1}^{n} w_i^{(b)} \psi_i.$$
3. Estimate $\mathbb{V}(\hat{\theta})$ by taking the variance of the B realizations $\{D_1, \cdots, D_B\}$.