Cumulative Incidence Curves

1.1 Setup

Consider time to event data $\{(T_i, \delta_i)\}$, where time T_i is absolutely continuous and the status δ_i is coded as:

$$\delta_i = \begin{cases} 0 & \text{censoring}, \\ 1 & \text{event}, \\ 2 & \text{competing risk}. \end{cases}$$

Definition 1.1.1. The **cumulative incidence curve** (CIC) for type-1 events $F_1(t)$ is the probability of experiencing the event before the competing risk by time t:

$$F_1(t) = \mathbb{P}(T \leq t, \delta = 1) = \int_0^t S(u)\lambda_1(u)du,$$

where S is the overall survival function:

$$S(t) = \mathbb{P}(T > t) = e^{-\Lambda_1(t) - \Lambda_2(t)}.$$

■

Discussion 1.1.1. The overall survival function $S(t)$ may be estimated from data of the form $\big\{(T_i, \delta_i^*)\big\}$, where:

$$\delta_i^* = \begin{cases} 0 & \text{censoring}, \\ 1 & \text{event or competing risk}. \end{cases}$$

Let $\hat{S}$ denote the Kaplan-Meier estimate of S. The Nelson-Aalen estimate of $\Lambda_1(u)$ is:

$$\hat{\Lambda}_1(t) = \int_0^t \frac{dN_1(u)}{Y(u)},$$

where $Y(u) = \sum_{i=1}^n \mathbb{I}(T_i \geq u)$ is the number at risk and $N_1(u)$ is the counting process for type-1 events:

$$N_1(t) = \sum_{i=1}^n \mathbb{I}(T_i \leq t, \delta_i = 1).$$

The standard estimator of F_1 is:

$$\hat{F}_1(t) = \int_0^t \hat{S}(u)d\hat{\Lambda}_1(u),$$

♠

1.2 Background

We make use of the following results proved elsewhere.

Proposition 1.2.1 (**Kaplan-Meier to Nelson-Aalen**). Let $S(t)$ denote the overall survival function and $\Lambda(t) = \Lambda_1(t) + \Lambda_2(t)$ the overall cumulative hazard. Then:

$$\frac{\sqrt{n}\big\{\hat{S}(t) - S(t)\big\}}{-S(t)} = \sqrt{n}\big\{\hat{\Lambda}(t) - \Lambda(t)\big\} + o_p(1), \tag{1.2.1}$$

where $\hat{S}$ is the KM estimator, and $\hat{\Lambda}$ the NA. ♦

Proposition 1.2.2 (**Martingale Expansion of Nelson-Aalen**). The NA estimator is expressible as:

$$\sqrt{n}\big\{\hat{\Lambda}(t) - \Lambda(t)\big\} = \int_0^t \frac{\sqrt{n}}{Y(u)} dM(u) + o_p(1),$$

where $M(u) = N(u) - \int_0^t Y(u)d\Lambda(u)$ is the counting process martingale.

Likewise, for the cause-specific cumulative hazard:

$$\sqrt{n}\big\{\hat{\Lambda}_1(t) - \Lambda_1(t)\big\} = \int_0^t \frac{\sqrt{n}}{Y(u)} dM_1(u) + o_p(1),$$

where $M_1(u) = N_1(u) - \int_0^t Y(u)d\Lambda_1(u)$ is the counting process martingale. ♦

Proposition 1.2.3. The predictable covariation is a bilinear form:

$$\langle M_1 + M_2, M_1 + M_2 \rangle = \langle M_1 \rangle + \langle M_2 \rangle + 2\langle M_1, M_2 \rangle.$$

The analogous result holds for the optional covariation. ♦

Proposition 1.2.4. If H_1, H_2 are predictable processes and M_1, M_2 are mean-zero martingales, then the predictable covariation of the corresponding stochastic integrals is:

$$\left\langle \int H_1(t)dM_1(t), \int H_2(t)dM_2(t) \right\rangle = \int H_1(t)H_2(t)d\langle M_1(t), M_2(t)\rangle.$$

The analogous result holds for the optional covariation. ♦

1.3 Martingale Representation

Proposition 1.3.5.

$$\begin{aligned}\sqrt{n}\big\{\hat{F}_1(t) - F_1(t)\big\} &= \int_0^t \sqrt{n}\big\{\hat{S}(u) - S(u)\big\} d\hat{\Lambda}_1(u) \\ &\quad + \int_0^t \sqrt{n} S(u)\big\{d\hat{\Lambda}_1(u) - d\Lambda_1(u)\big\}.\end{aligned}$$

♦

Proof.

$$\sqrt{n}\big\{\hat{F}_1(t) - F_1(t)\big\} = \int_0^t \sqrt{n}\hat{S}(u)d\hat{\Lambda}_1(u) - \int_0^t \sqrt{n}S(u)d\Lambda_1(u).$$

Adding and subtracting $S(u)$ from the first integrand:

$$\begin{aligned}\int_0^t \sqrt{n}\hat{S}(u)d\hat{\Lambda}_1(u) &= \int_0^t \sqrt{n}\big\{\hat{S}(u) - S(u) + S(u)\big\}d\hat{\Lambda}_1(u)\\ &= \int_0^t \sqrt{n}\big\{\hat{S}(u) - S(u)\big\}d\hat{\Lambda}_1(u) + \int_0^t \sqrt{n}S(u)d\hat{\Lambda}_1(u).\end{aligned}$$

■

Proposition 1.3.6.

$$\begin{aligned}&\int_0^t \sqrt{n}\big\{\hat{S}(u) - S(u)\big\}d\hat{\Lambda}_1(u)\\ &\qquad = -\int_0^t \sqrt{n}\big\{\hat{\Lambda}(u) - \Lambda(u)\big\}\cdot dF_1(u) + o_p(1)\\ &\qquad = -\sqrt{n}\big\{\hat{\Lambda}(t) - \Lambda(t)\big\}F_1(t) + \int_0^t \sqrt{n}F_1(u)d\big\{\hat{\Lambda}(u) - \Lambda(u)\big\} + o_p(1).\end{aligned}$$

♦

Proof. Since $\hat{\Lambda}_1(\cdot)$ is consistent for $\Lambda_1(\cdot)$:

$$\int_0^t \sqrt{n}\big\{\hat{S}(u) - S(u)\big\}d\hat{\Lambda}_1(u) = \int_0^t \sqrt{n}\big\{\hat{S}(u) - S(u)\big\}d\Lambda_1(u) + o_p(1).$$

Applying (1.2.1):

$$\begin{aligned}\int_0^t \sqrt{n}\big\{\hat{S}(u) - S(u)\big\}d\Lambda_1(u) &= -\int_0^t \frac{\sqrt{n}\big\{\hat{S}(u) - S(u)\big\}}{-S(u)}\cdot S(u)d\Lambda_1(u)\\ &= -\int_0^t \sqrt{n}\big\{\hat{\Lambda}(u) - \Lambda(u)\big\}\cdot dF_1(u) + o_p(1).\end{aligned}$$

Integrating by parts:

$$\begin{aligned}&\int_0^t \sqrt{n}\big\{\hat{\Lambda}(u) - \Lambda(u)\big\}\cdot dF_1(u)\\ &\qquad = \sqrt{n}\Big[\big\{\hat{\Lambda}(u) - \Lambda(u)\big\}F_1(u)\Big]_{u=0}^{u=t} - \int_0^t \sqrt{n}F_1(u)d\big\{\hat{\Lambda}(u) - \Lambda(u)\big\}\\ &\qquad = \sqrt{n}\big\{\hat{\Lambda}(t) - \Lambda(t)\big\}F_1(t) - \int_0^t \sqrt{n}F_1(u)d\big\{\hat{\Lambda}(u) - \Lambda(u)\big\}.\end{aligned}$$

■

Proposition 1.3.7 (**Martingale Representation**).

$$\begin{aligned}\sqrt{n}\big\{\hat{F}_1(t) - F_1(t)\big\} =& - F_1(t)\int_0^t \frac{\sqrt{n}\cdot dM(u)}{Y(u)} + \int_0^t \frac{\sqrt{n}\cdot F_1(u)}{Y(u)}dM(u) \\ &+ \int_0^t \frac{\sqrt{n}\cdot S(u)}{Y(u)}dM_1(u) + o_p(1).\end{aligned}$$

♦

Proof. From the previous propositions:

$$\begin{aligned}\sqrt{n}\big\{\hat{F}_1(t) - F_1(t)\big\} =& - \sqrt{n}\big\{\hat{\Lambda}(t) - \Lambda(t)\big\}F_1(t) + \int_0^t \sqrt{n}F_1(u)d\big\{\hat{\Lambda}(u) - \Lambda(u)\big\} \\ &+ \int_0^t \sqrt{n}S(u)\big\{d\hat{\Lambda}_1(u) - d\Lambda_1(u)\big\} + o_p(1).\end{aligned}$$

The conclusion follows from applying the martingale expansion of the Nelson-Aalen estimator to $\sqrt{n}\big\{\hat{\Lambda}(t) - \Lambda(t)\big\}$, $\sqrt{n}d\big\{\hat{\Lambda}(t) - \Lambda(t)\big\}$, and $\sqrt{n}d\big\{\hat{\Lambda}_1(t) - \Lambda_1(t)\big\}$.

■

Corollary 1.3.1.

$$\begin{aligned}\sqrt{n}\big\{\hat{F}_1(t) - F_1(t)\big\} =& - F_1(t)\int_0^t \frac{\sqrt{n}}{Y(u)}dM_1(u) + \int_0^t \frac{\sqrt{n}\cdot\{1 - F_2(u)\}}{Y(u)}dM_1(u) \\ &- F_1(t)\int_0^t \frac{\sqrt{n}}{Y(u)}dM_2(u) + \int_0^t \frac{\sqrt{n}\cdot F_1(u)}{Y(u)}dM_2(u) + o_p(1).\end{aligned}$$

♣

Proof. Substituting $M(u) = M_1(u) + M_2(u)$ and $S(u) = 1 - F_1(u) - F_2(u)$ gives the result. ■

1.4 Predictable and Optional Variations

Proposition 1.4.8. Suppose $n^{-1}Y(t) \xrightarrow{p} y(t)$, then as $n \to \infty$:

$$\begin{aligned}\big\langle\sqrt{n}\big\{\hat{F}_1(t) - F_1(t)\big\}\big\rangle \xrightarrow{p}& F_1^2(t)\int_0^t \frac{\lambda_1(u)}{y(u)}du + \int_0^t \frac{\big\{1 - F_2(u)\big\}^2\lambda_1(u)}{y(u)}du \\ &+ F_1^2(t)\int_0^t \frac{\lambda_2(u)}{y(u)}du + \int_0^t \frac{F_1^2(u)\lambda_2(u)}{y(u)}du \\ &- 2F_1(t)\int_0^t \frac{\big\{1 - F_1(u)\big\}\lambda_1(u)}{y(u)}du \\ &- 2F_1(u)\int_0^t \frac{F_1(u)\lambda_2(u)}{y(u)}du \end{aligned} \tag{1.4.2}$$

♦

Proof. Finding the optional variation:

$$
\begin{aligned}
\left\langle\sqrt{n}\left\{\hat{F}_1(t)-F_1(t)\right\}\right\rangle = F_1^2(t)\int_0^t \frac{n}{Y^2(u)}d\langle M_1(u)\rangle + \int_0^t \frac{n\left\{1-F_2(u)\right\}^2}{Y^2(u)}d\langle M_1(u)\rangle \\
+ F_1^2(t)\int_0^t \frac{n}{Y^2(u)}d\langle M_2(u)\rangle + \int_0^t \frac{nF_1^2(u)}{Y^2(u)}d\langle M_2(u)\rangle \\
- 2F_1(t)\int_0^t \frac{n\left\{1-F_1(u)\right\}}{Y^2(u)}d\langle M_1(u)\rangle \\
- 2F_1(t)\int_0^t \frac{nF_1(u)}{Y^2(u)}d\langle M_2(u)\rangle + o_p(1).
\end{aligned}
$$

Substituting $d\langle M_j(u)\rangle = Y(u)\lambda_j(u)du$:

$$
\begin{aligned}
\left\langle\sqrt{n}\left\{\hat{F}_1(t)-F_1(t)\right\}\right\rangle = F_1^2(t)\int_0^t \frac{\lambda_1(u)}{n^{-1}Y(u)}du + \int_0^t \frac{\left\{1-F_2(u)\right\}^2\lambda_1(u)}{n^{-1}Y(u)}du \\
+ F_1^2(t)\int_0^t \frac{\lambda_2(u)}{n^{-1}Y(u)}du + \int_0^t \frac{F_1^2(u)\lambda_2(u)}{n^{-1}Y(u)}du \\
- 2F_1(t)\int_0^t \frac{\left\{1-F_2(u)\right\}\lambda_1(u)}{n^{-1}Y(u)}du \\
- 2F_1(u)\int_0^t \frac{F_1(u)\lambda_2(u)}{n^{-1}Y(u)}du + o_p(1).
\end{aligned}
$$

Letting $n^{-1}Y(u) \xrightarrow{p} y(u)$ gives the result. ■

Discussion 1.4.1. By the martingale central limit theorem:

$$
\sqrt{n}\left\{\hat{F}_1(t)-F_1(t)\right\} \rightsquigarrow W\left\{\sigma^2_{\text{CIC}}(t)\right\},
$$

where $\sigma^2_{\text{CIC}}(t)$ is the RHS of (1.4.2). An estimate of the variance is obtained by finding the optional variation:

$$
\begin{aligned}
\hat{\sigma}^2_{\text{CIC}}(t) = \left[\sqrt{n}\left\{\hat{F}_1(t)-F_1(t)\right\}\right] = F_1^2(t)\int_0^t \frac{dN_1(u)}{n^{-1}Y^2(u)} + \int_0^t \frac{\left\{1-F_2(u)\right\}^2 dN_1(u)}{n^{-1}Y^2(u)} \\
+ F_1^2(t)\int_0^t \frac{dN_2(u)}{n^{-1}Y^2(u)} + \int_0^t \frac{F_1^2(u)dN_2(u)}{Y^2(u)} \\
- 2F_1(t)\int_0^t \frac{\left\{1-F_2(u)\right\}dN_1(u)}{n^{-1}Y^2(u)} \\
- 2F_1(t)\int_0^t \frac{F_1(u)dN_2(u)}{n^{-1}Y^2(u)}.
\end{aligned}
$$

♠

Area Under the Cumulative Incidence Curve

Example 2.0.1. Consider the area under the cumulative incidence curve (AUCIC):

$$U(\tau) = \int_0^\tau F_1(t)dt = \int_0^\tau \int_0^t S(u)d\Lambda_1(u).$$

and let $\hat{U}(\tau)$ denote the estimator:

$$\hat{U}(\tau) = \int_0^\tau \hat{F}_1(t)dt = \int_0^\tau \int_0^t \hat{S}(u)d\hat{\Lambda}_1(u).$$

Using the martingale representation of $\sqrt{n}\big\{\hat{F}_1(t) - F_1(t)\big\}$, the standardized difference:

$$\sqrt{n}\big\{\hat{U}(\tau) - U(\tau)\big\} = \int_0^\tau \sqrt{n}\big\{\hat{F}_1(t) - F_1(t)\big\}dt$$

is expressible as:

$$\begin{aligned}\int_0^\tau \sqrt{n}\big\{\hat{F}_1(t) - F_1(t)\big\}dt = &- \int_0^\tau F_1(t)\int_0^t \frac{\sqrt{n}}{Y(u)}dM_1(u)dt + \int_0^\tau\int_0^t \frac{\sqrt{n}\big\{1-F_2(u)\big\}}{Y(u)}dM_1(u)dt \\ &- \int_0^\tau F_1(t)\int_0^t \frac{\sqrt{n}}{Y(u)}dM_2(u)dt + \int_0^\tau\int_0^t \frac{\sqrt{n}F_1(u)}{Y(u)}dM_2(u)dt + o_p(1).\end{aligned}$$

The predictable variation is:

$$\begin{aligned}\big\langle \sqrt{n}\big\{\hat{U}(\tau) - U(\tau)\big\}\big\rangle = &\int_0^\tau F_1^2(t)\int_0^t \frac{\lambda_1(u)}{n^{-1}Y(u)}dudt + \int_0^\tau\int_0^t \frac{\big\{1-F_2(u)\big\}^2\lambda_1(u)}{n^{-1}Y(u)}dudt \\ &+ \int_0^\tau F_1^2(t)\int_0^t \frac{\lambda_2(u)}{n^{-1}Y(u)}dudt + \int_0^\tau\int_0^t \frac{F_1^2(u)\lambda_2(u)}{n^{-1}Y(u)}dudt \\ &- 2\int_0^\tau F_1(t)\int_0^t \frac{\big\{1-F_2(u)\big\}\lambda_1(u)}{n^{-1}Y(u)}dudt \\ &- 2\int_0^\tau F_1(t)\int_0^t \frac{F_1(u)\lambda_2(u)}{n^{-1}Y(u)}dudt + o_p(1).\end{aligned}$$

Let $\sigma^2_{\text{AUCIC}}(\tau)$ denote the limit in probability of the predictable variation as $n \to \infty$. By the martingale central limit theorem:

$$\sqrt{n}\big\{\hat{U}(\tau) - U(\tau)\big\} \rightsquigarrow W\big\{\sigma^2_{\text{AUCIC}}(t)\big\}.$$

An estimate of the asymptotic variance is obtained from the optional variation:

$$\begin{aligned}\hat{\sigma}^2_{\text{AUCIC}}(\tau) = &\int_0^\tau \hat{F}_1^2(t)\int_0^t \frac{dN_1(u)}{n^{-1}Y^2(u)}dt + \int_0^\tau\int_0^t \frac{\big\{1-\hat{F}_2(u)\big\}^2 dN_1(u)}{n^{-1}Y^2(u)}dt \\ &+ \int_0^\tau \hat{F}_1^2(t)\int_0^t \frac{dN_2(u)}{n^{-1}Y^2(u)}dt + \int_0^\tau\int_0^t \frac{\hat{F}_1^2(u)dN_2(u)}{n^{-1}Y^2(u)}dt \\ &- 2\int_0^\tau \hat{F}_1(t)\int_0^t \frac{\big\{1-\hat{F}_2(u)\big\}dN_1(u)}{n^{-1}Y^2(u)}dt \\ &- 2\int_0^\tau \hat{F}_1(t)\int_0^t \frac{\hat{F}_1(u)dN_2(u)}{n^{-1}Y^2(u)}dt.\end{aligned}$$

♠