Setting

A **competing risk** is an event whose occurrence precludes observation of the event of interest. More generally, competing risks are present when a subject may experience one of J mutually exclusive types of event. Consider the transition model in figure (1.1):

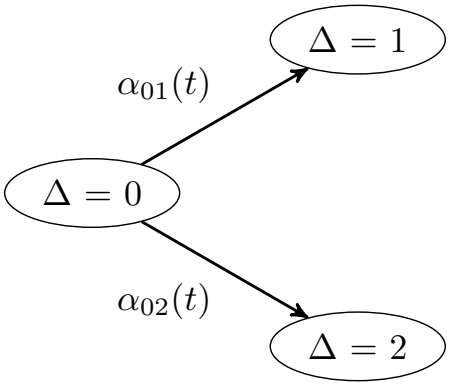

Figure 1.1: Transition Model for the Basic Competing Risk Setting

Here $\Delta = 0$ represents the initial state, $\Delta = 1$ is the event of interest, and $\Delta = 2$ is the competing event. Define the status process $\Delta_t \in \{0, 1, 2\}$, which tracks the state of the subject. The transition time T is defined as:

$$T = \inf\big\{t > 0 : \Delta_t \neq 0\big\}.$$

Letting C denote a random right censoring time, the observed data are $T \wedge C$, the minimum of the event and censoring times, and $\mathbb{I}(T \leq C) \cdot \Delta_T$, the status at the event time prefixed by an indicator that is 1 if the event is observed, and 0 otherwise. The non-informative censoring assumption requires that:

$$\mathbb{P}(t \leq T < T + dt, T \leq C | \mathscr{F}_t) = \mathbb{I}(t \leq T \wedge C) \cdot \alpha_{0\bullet}(t)dt.$$

Cause-specific Hazard

Definition 2.0.1. The **cause-specific hazard** [4] for competing risk k is:

$$\alpha_{0j}(t) \equiv \lim_{\epsilon \to 0} \frac{1}{\epsilon} \mathbb{P}\big(t \leq T < t + \epsilon, \Delta_T = j | T \geq t\big). \tag{2.0.1}$$

In differential notation:

$$\alpha_{0j}(t)dt = \mathbb{P}\big\{T \in [t, t + dt), \Delta_T = k | T \geq t\big\}.$$

■

Discussion 2.0.1. The risk set for the cause-specific hazard includes only those subjects who have not experienced any event by time t. Subjects who have experienced a competing risk j' are no longer at risk to experience event type j. The cause-specific hazard is interpreted as the risk of experiencing a type k event in the next short time interval dt given the patient has not experienced any type of event by time t. ♠

Definition 2.0.2. The **all-cause hazard** is the sum of the cause-specific hazards:

$$\alpha_{0\bullet}(t) = \sum_{j=1}^{J} \alpha_{0j}(t).$$

■

Discussion 2.0.2. The all-cause cumulative hazard is:

$$A_{0\bullet}(t) = \int_0^t \alpha_{0\bullet}(u)du.$$

The survival function for the random variable $T = \inf\{t > 0 : \Delta_t \neq 0\}$ is:

$$S(t) = \mathbb{P}(T > t) = e^{-A_{0\bullet}(t)}.$$

♠

Example 2.0.1. The cause-specific proportion hazards model is:

$$\alpha_{0j}(t; \boldsymbol{z}) = \alpha_{0j}(t)e^{\boldsymbol{z}'\boldsymbol{\beta}_j},$$

where $\alpha_{0j}(t)$ is an unspecified base cause-specific hazard, $\boldsymbol{z}$ is the set of covariates, and $\boldsymbol{\beta}_j$ is a vector of cause-specific log hazard ratios. The model is estimated by partial likelihood with competing events handled as for censoring:

```
coxph(Surv(time, status == j) ~ z, data = data)
```

Note that the justification for regarding competing events as censoring when estimating the cause-specific hazard is based on the partial likelihood, and does not require independence of latent failure times corresponding to type j and type $k \neq j$ events. Regarding competing events as censoring is *not* valid when estimating the cumulative incidence. ♠

Discussion 2.0.3. The cause-specific "survival" function

$$S_j(t) = \exp\left\{-\int_0^t \alpha_{0j}(u)du\right\}$$

represents the probability of being event-free by time t in a hypothetical setting where the competing events cannot occur before the event of interest [1]. ♠

Cumulative Incidence Function

Definition 3.0.1. In the presence of competing risks, the **cumulative incidence function** (CIF) for type j events is defined as:

$$\begin{aligned} F_j(t) &= \mathbb{P}(T \leq t, \Delta_T = j) \\ &= \int_0^t S(u-)\alpha_{0j}(u)du = \int_0^t S(u-)dA_j(u), \end{aligned}$$

where $S(t) = \mathbb{P}(T \geq t)$ is the overall survival function. Observe that the cumulative incidence function depends on each of the cause-specific hazards through $S(t)$. ■

Discussion 3.0.1. In the absence of competing events, the survival function $S(t)$ is typically estimated via the Kaplan-Meier (KM) product integral estimator $\hat{S}(t)$, and $1 - \hat{S}(t)$ provides a consistent estimate of the all cause cumulative incidence function $F(t) = 1 - S(t)$. In the presence of competing events, $1 - \hat{S}(t)$ systematically *overestimates* the CIF. This is because $1 - \hat{S}(t)$ estimates a true distribution function, which approaches 1 as $t \to \infty$. By contrast, in the limit as $t \to \infty$, the CIF converges to the marginal probability of experiencing a type j event:

$$\lim_{t \to \infty} F_j(t) = \lim_{t \to \infty} \mathbb{P}(T \leq t, \Delta_T = j) = \mathbb{P}(\Delta_T = j).$$

If competing events are in fact present, then $\mathbb{P}(\Delta_T = j) < 1$.

♠

Discussion 3.0.2. CIFs are additive. Thus, if $F_1(t)$ is the CIF for type 1 events and $F_2(t)$ is the CIF for type 2 events, $F_1(t) + F_2(t)$ is the CIF for the composite outcome of type 1 or 2 events:

$$F_1(t) + F_2(t) = \mathbb{P}\big(T \leq t, \Delta_T = 1\big) + \mathbb{P}\big(T \leq t, \Delta_T = 2\big) = \mathbb{P}\big(T \leq t, \Delta_T \in \{1, 2\}\big).$$

Additive stems from $(T \leq, \Delta_T = 1)$ and $(T \leq, \Delta_T = 2)$ being disjoint events. Graphically, the individual CIFs may be stacked such that the overall height represents the cumulative incidence of the composite event. In the case of $J = 2$ events, $F_1(t) + F_2(t)$ is, by law of total probability, $F(t) = 1 - S(t)$:

$$F(t) = \mathbb{P}(T \leq t) = \sum_{j=1}^{J} \mathbb{P}(T \leq t, \Delta_T = j).$$

♠

Example 3.0.1. The cumulative incidence functions of all competing events may be consistently estimated using the matrix-valued generalized of the KM product integral estimator. For example, in the case of $J = 2$ competing events (1.1):

$$\prod_{0<u\leq t} \{\boldsymbol{I} + d\boldsymbol{A}\} = \prod_{0<u\leq t} \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} -A_{0\bullet}(u) & A_{01}(u) & A_{02}(u) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right\}$$

Notice that rows 2 and 3 of $d\boldsymbol{A}$ are all zero, reflecting that $\Delta = 1$ and $\Delta = 2$ are absorbing states. Entry $(1, 1)$ of the product integral is the overall survival function $S(T)$, and is identical to the standard KM estimator. Entry $(1, 2)$ is the CIF for type 1 events $F_1(t)$, and entry $(1, 3)$ is the CIF for type 2 events $F_2(t)$.

♠

3.1 Simulation

Example 3.1.2. Competing risks data may be simulated as follows.

i. Specify the cause-specific hazards $\alpha_{01}(t), \cdots, \alpha_{0J}(t)$.

ii. Simulate event times from the overall hazard $\alpha_{0\bullet}(t) = \sum_{j=1}^{J} \alpha_{0j}(t)$.

iii. For a given event time T, draw the event type Δ_T from a multinomial distribution, with the probability of event j being $\alpha_{0j}(T)/\alpha_{0\bullet}(T)$.

iv. Simulate the censoring time C. The final data are $\big(T \wedge C,\ \mathbb{I}(T \leq C) \cdot \Delta_T\big)$.

In cases where the overall hazard $\alpha_{0\bullet}(t)$ has a non-standard form, the probability integral theorem may be used to generate the event time T. Recall that if U is a standard uniform $(0, 1)$ random variable, then the distribution of $F^{-1}(U)$ is F. Letting $F = \exp\{A_{0\bullet}(t)\}$, the distribution corresponding to the cumulative overall hazard:

$$T \sim A_{0\bullet}^{-1}\big\{ -\ln(1 - U)\big\}.$$

When the cumulative overall hazard lacks an analytical inverse, T may be obtained from U by numerically solving:

$$A_0(T) + \ln(1 - U) \overset{\text{Set}}{=} 0.$$

♠

3.2 Estimation

To estimate the CIF, define the sample counting process for type j events:

$$N_j(t) = \sum_{i=1}^{n} \mathbb{I}(T_i \leq t, T_i \leq C_i, \Delta_{T_i} = j),$$

and the sample at-risk process:

$$Y(t) = \sum_{i=1}^{n} \mathbb{I}(T_i \wedge C_i \geq t).$$

The **Nelson-Aalen** estimator of the cumulative hazard for type k events is:

$$\hat{A}_j(t) = \int_0^t \frac{dN_j(u)}{Y(u)}.$$

The asymptotic sampling variance of $\hat{A}_j(t)$ may be estimated via:

$$\hat{\mathbb{V}}\{\hat{A}_j(t)\} = \int_0^t \frac{dN_j(u)}{Y^2(u)}.$$

Let $\hat{S}(\cdot)$ denote the KM estimator of the survival function:

$$\hat{S}(t) = \prod_{0<u\leq t} \big\{1 - d\hat{A}_{0\bullet}(u)\big\}.$$

The asymptotic sampling variance of $\hat{S}(t)$ is:

$$\hat{\mathbb{V}}\{\hat{S}(t)\} = \big\{\hat{S}(t)\big\}^2 \int_0^t \frac{dN_j(u)}{Y^2(u)}.$$

The CIF is estimated by:

$$\hat{F}_j(t) = \int_0^t \hat{S}(u-) d\hat{A}_j(t).$$

In the case of $J = 2$ competing events, the asymptotic sampling variance of $\hat{S}(t)$ is:

$$\begin{aligned}
\hat{\mathbb{V}}\{\hat{F}_1(t)\} = F_1^2(t) \int_0^t \frac{dN_1(u)}{Y^2(u)} + \int_0^t \frac{\big\{1 - F_2(u)\big\}^2 dN_1(u)}{Y^2(u)} \\
+ F_1^2(t) \int_0^t \frac{dN_2(u)}{Y^2(u)} + \int_0^t \frac{F_1^2(u) dN_2(u)}{Y^2(u)} \\
- 2F_1(t) \int_0^t \frac{\big\{1 - F_2(u)\big\} dN_1(u)}{Y^2(u)} \\
- 2F_1(t) \int_0^t \frac{F_1(u) dN_2(u)}{Y^2(u)}.
\end{aligned}$$

3.3 Subdistribution Hazard

Example 3.3.3. Define the *subdistribution process* [2] for state j as:

$$\xi(t) = \mathbb{I}\big\{\Delta(t) = j\big\}.$$

Note that ξ_t may be zero either because Δ_t remains in the initial state, or because Δ_t has transitioned into a competing state. The *subdistribution time* $\tilde{T}$ is defined as:

$$\tilde{T}_j = \inf\big\{t > 0 | \xi(t) \neq 0\big\} = \begin{cases} T & \Delta_T = j, \\ \infty & \Delta_T \neq j. \end{cases}$$

The distribution function for the subdistribution time $\tilde{T}$ is identically the cumulative incidence function for type j events:

$$\mathbb{P}(\tilde{T}_j \leq t) = \mathbb{P}(T \leq t, \Delta_T = j) = F_1(t),$$

and $\mathbb{P}(\tilde{T}_j = \infty) = \mathbb{P}(X_T \neq j)$. The counting process corresponding to $\tilde{T}_j$:

$$\tilde{N}_j(t) = \mathbb{I}\big\{\tilde{T}_j \wedge C_i \leq t, \tilde{T}_j \leq C_i\big\}$$

is generally observable. However, the at-risk process is problematic:

$$\tilde{Y}_j(t) = \mathbb{I}\big(\tilde{T}_j \wedge C \geq t\big).$$

If the status process Δ_t enters state $k \neq j$, then $\tilde{T}_j$ becomes ∞, and $\tilde{T}_j \wedge C = C$. However, observation of the process may cease at $T < C$, the time at which Δ_t entered state k. When this is so, the censoring time C must be inferred. Fine and Gray [3] propose to address the missing censoring time by estimating the risk set:

$$\hat{\tilde{Y}}_j(t) = \mathbb{I}\big(T \wedge C \geq t\big) + \frac{\hat{G}(t-)}{\hat{G}(T-)}\mathbb{I}(C \geq T)\mathbb{I}(T < t, \Delta_T \neq j).$$

Here $\hat{G}(\cdot)$ is the Kaplan-Meier estimate for the distribution of the censoring time C, and:

$$\frac{\hat{G}(t-)}{\hat{G}(T-)} = \frac{\mathbb{P}(C \geq t)}{\mathbb{P}(C \geq T)} = \mathbb{P}(C \geq t | C \geq T).$$

Estimation of the risk set via $\hat{\tilde{Y}}_j(t)$ is the approach taking by the `cmprsk` package. ♠

Definition 3.3.1. The **subdistribution hazard** of Fine and Gray [3] is the hazard corresponding to $\tilde{T}_j$:

$$\lambda_j(t) = \lim_{\epsilon \to 0} \frac{1}{\epsilon}\mathbb{P}\big(t \leq \tilde{T}_j < t + \epsilon | \tilde{T}_j \geq t\big).$$

In terms of the original event-time and status processes:

$$\lambda_j(t) = \lim_{\epsilon \to 0} \frac{1}{\epsilon}\mathbb{P}\big\{t \leq T < t + \epsilon, \Delta_T = j | (T \geq t) \cup (T < t \cap \Delta_T \neq j)\big\}. \tag{3.3.2}$$

■

Discussion 3.3.1. The risk set of the cause-specific hazard (2.0.1) includes only those who have not yet experienced an event of any time. By contrast, the risk set of the subdistribution hazard (3.3.2) includes both those who have not experienced an event of any type ($T \geq t$) *and* those who have experienced a competing event ($T < t \cap D \neq k$). ♠

Proposition 3.3.1. The subdistribution hazard is connected with the CIF via:

$$\lambda_j(t) = -\frac{d \ln \{1 - F_j(t-)\}}{dt}.$$

♦

Proof. From the definition of the CIF:

$$F_j(t-) = \mathbb{P}(T < t, \Delta_T = j).$$

Taking the complement:

$$1 - F_j(t-) = \mathbb{P}\{(T \geq t) \cup (T < t \cap \Delta_T \neq j)\}. \tag{3.3.3}$$

The logarithmic derivative of $1 - F_j(t-)$ is:

$$-\frac{d}{dt} \ln \{1 - F_j(t-)\} = \frac{\frac{d}{dt} F_j(t-)}{1 - F_j(t-)}.$$

The denominator is (3.3.3). The numerator is:

$$\frac{d}{dt} F_j(t-) = \mathbb{P}(T = t, \Delta_T = j).$$

Since the event $\{T = t, \Delta_T = j\}$ is contained within $\{(T \geq t) \cup (T < t \cap \Delta_T \neq j)\}$, their intersection is the former event. Therefore:

$$\begin{aligned} \frac{\frac{d}{dt} F_j(t-)}{1 - F_j(t-)} &= \frac{\mathbb{P}(T = t, \Delta_T = j)}{\mathbb{P}\{(T \geq t) \cup (T < t \cap \Delta_T \neq j)\}} \\ &= \mathbb{P}\big\{T = t, \Delta_T = j | (T \geq t) \cup (T < t \cap \Delta_T \neq j)\big\}. \end{aligned}$$

■

Corollary 3.3.1. The subdistribution hazard λ_j is defined such that:

$$F_j(t) = \mathbb{P}(T \leq t, \Delta_T = j) = 1 - e^{\int_0^t \lambda_j(u) du}.$$

♣

Discussion 3.3.2. The proportional hazards model for the subdistribution hazard is:

$$\lambda_j t) = \lambda_{0j}(t) e^{\boldsymbol{z}' \boldsymbol{\beta}_j},$$

where $\lambda_{0j}(t)$ is an unspecified baseline subdistribution hazard, $\boldsymbol{z}$ is a set of covariates, and $\boldsymbol{\beta}_j$ is a vector of log subdistribution hazard ratios. This model may be estimated using the `crr` function in the `cmprsk` package:

```
crr(ftime = data$time, fstatus = data$status, cov1 = z)
```

A covariate can have different associations with the cause-specific and subdistribution hazards. For example, suppose a treatment reduces the mortality rate, but has no effect on the rate of event j. Since the treatment does not increase the event rate, it will have no association with the cause-specific hazard. However, the cumulative incidence of the event is expected to increase, since patients survive longer, and thereby spend more time at risk for the event [5]. Conversely, a treatment may have an effect on the cause-specific hazard without having an effect on the cumulative incidence [3]. Defining a composite endpoint between the event of interest and mortality may be appropriate since this approach adjusts for mortality rate differences between the two arms. ♠

References

[1] Austin, PC and Lee, DS and Fine, JP. "Introduction to the Analysis of Survival Data in the Presence of Competing Risks". In: *Circulation* 133.6 (2016), pp. 601–609.

[2] J Beyersmann, M Schumacher, and A Allignol. *Competing Risk and Multistate Models with R*. Springer-Verlag, 2012.

[3] Fine, JP and Gray, RJ. "A Proportional Hazards Model for the Subdistribution of a Competing Risk". In: *Journal of the American Statistical Association* 94.446 (1999), pp. 496–509.

[4] Prentice, RL and Kalbfleisch, JD and Peterson, AV and others. "The Analysis of Failure Times in the Presence of Competing Risks". In: *Biometrics* 34.4 (1978), pp. 541–554.

[5] Wolbers, M and Koller, MT and Stel, VS and others. "Competing risks analyses: objectives and approaches". In: *European Heart Journal* 35 (2014), pp. 2936–2941.