Cumulative Incidence

Germán Rodríguez
grodri@princeton.edu

Spring 2012

We continue our treatment of competing risks by considering estimation of the cumulative incidence function and the Fine and Gray competing risks regression model.

1 The Cumulative Incidence Function

In our earlier discussion we introduced the cause-specific densities

$$f_j(t) = \lim_{dt \downarrow 0} \Pr\{T \in (t, t+dt) \text{ and } J = j\}/dt$$

which have the property of summing to the overall density $f(t) = \sum_j f_j(t)$.

The integral

$$I_j(t) = \int_0^t f_j(u)du = \Pr\{T \le t \text{ and } J = j\}$$

is called the *cumulative incidence function* (CIF), and represents the probability that an event of type j has occurred by time t.

Earlier we also introduced the cause-specific hazards

$$\lambda_j(t) = \lim_{dt \downarrow 0} \Pr\{T \in (t, t+dt) \text{ and } J = j | T > t\}/dt$$

representing the (conditional) rate of occurrence of events of type j at time t among survivors to that time. The cause-specific density can be written as

$$f_j(t) = S(t)\lambda_j(t)$$

reflecting the fact that to experience an event of type j at time t you first have to survive to time t, and then experience an event of type j conditional on having survived to t.

This representation leads to a non-parametric estimator of the cumulative incidence function which extends the Kaplan-Meier estimator. With distinct failure times $0 < t_{(1)} < \cdots < t_{(m)} < \infty$, the estimator is

$$\hat{I}_j(t) = \sum_{i:t_{(i)} \leq t} \hat{S}(t_{(i)}) \frac{d_{ij}}{n_i}$$

where d_{ij} is the number of events of type j at time $t_{(i)}$, n_i is the total number of observations at risk at time $t_{(i)}$, and $\hat{S}(t_{(i)}$ is the standard Kaplan-Meier estimator of survival to time $t_{(i)}$.

This is a step function with increments every time a failure of type j occurs. An interesting feature of this function is that if we add the cumulative incidence of all types of failure we obtain the complement of the Kaplan-Meier estimator:

$$\sum_j \hat{I}_j(t) = 1 - \hat{S}(t)$$

In words, at any time t the observations are either still at risk with probability $S(t)$, or have experienced an event of type j with probability $I_j(t)$ for some j. In the case of mortality you are either alive or have succumbed to cause of death j.

Standard errors for the cumulative incidence function can be obtained using the delta method, although the derivation is a bit more complicated that in the case of Greenwood's formula.

In the Stata logs we study how long U.S. Supreme Court Justices serve on the court, treating death and retirement as competing risks, with the nine justices currently serving treated as censored observations. We find, for example, that averaging over the existence of court, the probability that a justice will die on the job is 48% and the probability of retiring is 52%.

2 The Fine-Gray Model

How do you introduce covariates in the context of competing risks? There are essentially two approaches:

1. You can apply a Cox proportional-hazards model to the cause-specific hazards introduced earlier, or

2. You can use a model due to Fine and Gray that focuses on the cumulative incidence function.

The first approach is more structural, focusing on the covariates of the risk of each type of event. The second approach is more descriptive, focusing on the probability of each event type.

To understand the difference in approaches note that a covariate may appear to increase the incident of events of a certain type simply by lowering the rate of occurrence of events of other types, even if it has no effect on the rate of occurrence of the event in question.

We now describe the Fine and Gray model. Let $I_j(t,x)$ denote the cumulative incidence function for events of type j given a vector of covariates x. We can formally treat the complement of the CIF as a survival function and calculate the underlying hazard. To avoid confusion with the cause-specific and overall hazards we follow Fine and Gray in calling this a *sub*-hazard for cause j and denote it with a bar

$$\bar{\lambda}_j(t,x) = -\frac{d}{dt}\log(1 - I_j(t,x)) = \frac{f_j(t)}{1 - I_j(t)}$$

They then propose a proportional hazards model for the sub-hazard associated with type j, effectively writing it as

$$\bar{\lambda}_j(t,x) = \bar{\lambda}_{j0}(t)\exp\{x'\beta_j\}$$

where $\bar{\lambda}_{j0}(t)$ is the baseline sub-hazard for events of type j and $\exp\{x'\beta_j\}$ is the relative risk associated with covariates x.

While the formulation looks very similar to Cox regression, the present model applies to the sub-hazard underlying the cumulative incidence function, not the cause-specific hazards. One problem with this approach is that the sub-hazard is hard to interpret. From the Fine and Gray definition,

$$\begin{aligned}\bar{\lambda}_j(t) = \quad & \lim_{dt\downarrow 0}\Pr\{T \in (t, t+dt) \text{ and } J = j| \\ & T > t \textbf{ or } T \le t \text{ and } J \neq j\}/dt\end{aligned}$$

In other words, we count events of type j in a small interval $(t, t+dt)$ but treat as the risk set those alive at t *and* those who failed before t due to causes other than j.

The authors themselves recognize that this is an "un-natural" hazard because units who experienced an event of some other type before time t are not really at risk of experiencing an event of type j at t. One way to derive the sub-hazard as a standard hazard is to imagine a random variable T^* which equals T_j if an event of type j occurs and equals ∞ if an event of another type occurs, but this is also artificial.

In the end, the authors argue that their formulation is just a convenient way to model the incidence function. I agree, and tend to view their model as just a binary outcome model for the cumulative incidence function using the complementary log-log link. This is because under their model

$$\log(-\log(1 - I_j(t, x))) = \log(-\log(1 - I_{j0}(t))) + x'\beta_j$$

Thus, the effect of the covariates is to shift the transformed CIF up or down by an amount depending on the coefficients. Because the transformation is monotonic we know that positive coefficients indicate increases in the CIF and negative coefficients indicate decreases, but quantifying the effect requires conducting illustrative calculations.

In the Stata logs we study the length of service of U.S. Supreme Court justices treating death and retirement as competing risks and age at appointment and calendar year of appointment as predictors. (This is one case where estimating anything at zero values of the covariates is fraught with peril, as the court was founded in 1789 and age at appointment goes from 33 to 66.)

Fitting a Cox model to the hazard of death gives hazard ratios of 1.07 for age and 0.99 for calendar year, so the risk of death increases 7% per year of age at appointment and declines about one per cent per calendar year of appointment.

Fitting a similar Cox model to the risk of retirement gives hazard ratios of 1.10 for age and 1.00 for year, so the risk of retiring increases 10% per year of age at appointment and does not depend on the calendar year of appointment.

These two models give us a good understanding of the underlying process, and they can be used to estimate overall survival, cause-specific densities and hazards, and even the CIFs of death and retirement from their definitions.

Alternatively, we can fit a Fine-Gray model directly to the CIF for death or for retirement. Fitting a model to the CIF of death gives sub-hazard ratios (called SHR in Stata) of 1.01 for age at appointment (not significant) and 0.99 for calendar year of appointment (highly significant).

The first finding is that the probability of dying while serving on the court does *not* depend on age at appointment. You may find this result a bit surprising, as I did, but note that justices who are appointed at an older age have a higher risk of death than those appointed at younger ages in the same period, but they also have a higher risk of retirement, and these two forces are about equal so they balance out.

The second finding is that the probability of dying in the court has declined with calendar year of appointment, so justices appointed more recently are less likely to die and hence more likely to retire. The sub-hazard ratio of 0.99, however, is hard to interpret in terms other than the sign and significance without additional calculations.

The best bet here is to compute illustrative values of the CIF. In the Stata logs we show that the probability that a justice appointed at age 55 will leave the court by death is 29.7% if appointed in 1950 and 20.2% if appointed in 2000 (both figures lower than the overall mean of 48%). Note that

$$\log(-\log(1-.202))-\log(-\log(1-.297))=-0.446,$$

and $-0.446/50=-0.009$. the coefficient of year, which Stata reports as an SHR of $\exp\{-0.009\}=0.99$. Thus, the transformed CIF is declining 0.009 per calendar year.