Heckman and Honoré (1989):
The Identifiability of the Competing Risks Model

Jason Blevins

March 5, 2008

The Classical Competing Risks Model

- Suppose there are J competing risks $\{1, 2, \ldots, J\}$.
- Associated with each risk is a stochastic arrival time T_j.
- We observe only the distribution of the *identified minimum*:
  - The arrival time $T = \min_j T_j$.
  - The associated risk $I = \arg\min_j T_j$.
- Goal: Identify the joint distribution of the latent failure times given that we only observe the distribution of the identified minimum.
- Note that we aren't considering regressors yet.

Cox and Tsiatis Nonidentification Theorem

- For any joint distribution of latent failure times, there exists another such distribution with *independent failure times* that yields the same distribution of the minimum (Cox, 1959, 1962; Tsiatis, 1975).

- That is, given *r.v.*'s $(T_1, T_2, \ldots, T_J)$ there exist $(S_1, S_2, \ldots, S_J)$ with $S_i \perp\!\!\!\perp S_j$ for all $i \neq j$ such that (T, I_T) and (S, I_S) are observationally equivalent.

- In light of this result, any empirical work needed to proceed by placing some structure on the form of dependence across risks, for example, by assuming independence.

Importance of Dependence

- We are concerned with *conditional independence*—independence of the risks $T_1, \ldots, T_J$ conditional on X.

- Even conditional independence may not hold if, for example, we are studying an individual whose behavior may affect all of the risks.

- Yashin, Manton, and Stallard (1986): How do smoking, blood pressure, and body weight (regressors) affect time of death from cancer, heart disease, etc. (risks).

Overview

- Establish an identification theorem for a general class of competing risks models with regressors.

- This class includes models with marginal distributions that follow:
  - Proportional hazards.
  - Mixed proportional hazards.
  - Accelerated hazards.

- Results are presented for only two competing risks but generalize to any arbitrary finite number of risks.

Proportional Hazards Model

- We want to model the time of death T conditional on some covariates X.
- Conditional on X, T has cdf $F(t|x)$ and pdf $f(t|x)$.
- Hazard function: $\lambda(t|x) = \frac{f(t|x)}{1-F(t|x)}$.
- Integrated hazard: $\Lambda(t|x) = \int_0^t \lambda(s|x)\, ds$.
- If $\lambda(t|x) = z(t)\phi(x)$ then $\Lambda(t|x) = Z(t)\phi(x)$ with $Z(t) = \int_0^t z(s)\, ds$.
- We can also work with the survivor function: $S(t|x) = \Pr(T > t|x)$.
- Since $\Lambda(t|x) = -\ln[1 - F(t|x)]$, we have

$$S(t|x) = \Pr(T > t|x) = \exp\left[-Z(t)\phi(x)\right].$$

Proportional Hazards and Competing Risks

- Assuming for the moment that failure times are independent, we can easily generalize this to model competing risks.

- The distribution of each failure time has a proportional hazard specification.

- $Z(t)$ and ϕ may differ across risks.

- The joint survivor function is

$$S(t_1, t_2|x) = 1 - (1 - \exp[-Z_1(t_1)\phi_1(x)])\,(1 - \exp[-Z_2(t_2)\phi_2(x)])\,.$$

Introducing Dependence

- We could draw two *independent* failure times T_1 and T_2 by drawing (independently) $U_j \sim U(0,1)$ and solving for T_j:

$$U_j = S_j(t|x) = \exp\{-Z_j(t)\phi_j(x)\}.$$

- If $K(u_1, u_2) = u_1 u_2$ is the CDF of U_1 and U_2, the joint survivor function is

$$S(t_1, t_2|x) = K\left[\exp\{-Z_1(t)\phi_1(x)\}, \exp\{-Z_2(t)\phi_2(x)\}\right].$$

- We can introduce dependence in T_1 and T_2 by introducing dependence in U_1 and U_2 via K.

- For a general K on $[0,1]^2$ the joint survivor function is

$$S(t_1, t_2|x) = K\left[\exp\left\{-Z_1(t_1)\phi_1(x)\right\}, \exp\left\{-Z_2(t_2)\phi_2(x)\right\}\right]. \tag{1}$$

Generalization: Mixed Proportional Hazards

- Suppose that the competing risks are independent, $\phi_j(x) = e^{x\beta_j}$, and that one of the covariates, ω, is not observed:

$$S(t_1, t_2|x) = \int_\Omega \exp\left[-Z_1(t_1)e^{x\beta_1+c_1\omega}\right]\exp\left[-Z_2(t_2)e^{x\beta_2+c_2\omega}\right]\, dG(\omega).$$

- We can arrive at this model by choosing K such that:

$$K(\eta_1, \eta_2) = \int_\Omega \eta_1^{\exp(c_1\omega)}\eta_2^{\exp(c_2\omega)}\, dG(\omega).$$

Generalization: Accelerated hazards

$$S(t|x) = \exp\left[-Z\{t\phi(x)\}\right]$$

- Joint survivor with dependent competing risks:

$$S(t_1, t_2|x) = K\left(\exp\left[-Z_1\{t_1\phi_1(x)\}\right], \exp\left[-Z_2\{t_2\phi_2(x)\}\right]\right).$$

- For any K, the marginal distributions give rise to univariate accelerated hazard models.

Identification Theorem

Assume that (T_1, T_2) has joint distribution (1). Then Z_1, Z_2, ϕ_1, ϕ_2, and K are identified from the minimum of (T_1, T_2) under the following assumptions:

1. K is continuously differentiable with partial derivatives K_1 and K_2 and for $i = 1, 2$, $\lim_{n\to\infty} K_i(\eta_{1n}, \eta_{2n})$ is finite for all sequences η_{1n}, η_{2n} for which $\eta_{1n} \to 1$ and $\eta_{2n} \to 1$ for $n \to \infty$. We also assume that K is strictly increasing in each of its arguments.

2. $Z_1(1) = Z_2(1) = 1$ and $\phi_1(x_0) = \phi_2(x_0) = 1$ for some x_0.

3. The support of $\{\phi_1(x), \phi_2(x)\}$ is $(0, \infty) \times (0, \infty)$.

4. Z_1 and Z_2 are nonnegative, differentiable, strictly increasing functions, except that we allow them to be infinite for finite t.

Mapping Observables to Unobservables

Observed distributions:

$$Q_1(t|x) = \Pr(T_1 \geq t, T_2 \geq T_1|x) \quad Q_2(t|x) = \Pr(T_2 \geq t, T_1 \geq T_2|x).$$

Tsiatis (1975) establishes the following mappings:

$$\frac{\partial Q_1}{\partial t}(t|x) = \left[\frac{\partial S}{\partial t_1}\right]_{t_1=t_2=t} \quad \frac{\partial Q_2}{\partial t}(t|x) = \left[\frac{\partial S}{\partial t_2}\right]_{t_1=t_2=t}.$$

We have

$$\begin{aligned}\frac{\partial Q_1}{\partial t}(t|x) = -K_1\left[\exp\left\{-Z_1(t)\phi_1(x)\right\}, \exp\left\{-Z_2(t)\phi_2(x)\right\}\right]\\ \times \exp\left\{-Z_1(t)\phi_1(x)\right\} Z_1'(t)\phi_1(x).\end{aligned}$$

Identification of ϕ_j

Taking the ratio of $\frac{\partial Q_1(t|x)}{\partial t}$ at x and x_0 yields

$$\frac{K_1\left[\exp\{-Z_1(t)\phi_1(x)\},\exp\{-Z_2(t)\phi_2(x)\}\right]\exp\{-Z_1(t)\phi_1(x)\}\,Z_1'(t)\phi_1(x)}{K_1\left[\exp\{-Z_1(t)\phi_1(x_0)\},\exp\{-Z_2(t)\phi_2(x_0)\}\right]\exp\{-Z_1(t)\phi_1(x_0)\}\,Z_1'(t)\phi_1(x_0)}.$$

Taking $t \to 0$ and using the normalization yields $\phi_1(x)$. Our choice of x was arbitrary so $\phi_1(x)$ is identified on the entire support of X. Similarly for $\phi_2(x)$.

Identification of K

We know $S(t,t|x)$ since $S(t,t|x) = Q_1(t|x) + Q_2(t|x)$. Furthermore,

$$S(t,t|x) = K\left(\exp\left[-Z_1(t)\phi_1(x)\right], \exp\left[-Z_2(t)\phi_2(x)\right]\right).$$

Setting $t = 1$ gives

$$S(1,1|x) = K\left(\exp[-\phi_1(x)], \exp[-\phi_2(x)]\right).$$

and letting $\phi_1(x)$ and $\phi_2(x)$ vary over $(0,\infty)^2$ (by Assumption 3) yields K.

Identification of Z_j

$$S(t, t|x_n) = K\left(\exp[-Z_1(t)\phi_1(x_n)], \exp[-Z_2(t)\phi_2(x_n)]\right)$$

- Let $\phi_2(x) \to 0$ while holding $\phi_1(x)$ fixed.
- Then
$$S(t, t|x) \to K\left(\exp[-Z_1(t)\phi_1(x)], 1\right).$$
- Since K and ϕ_1 are known and K is strictly increasing in both arguments, we have $Z_1(t)$ for any t.
- Similarly for $Z_2(t)$.

Conclusion

Identification argument:

- Given the distribution of (T, I) and exploiting multiplicative separability gives us $\phi_j(x)$ for $j = 1, 2$.
- Using the full range of $\phi_j(x)$ on $(0, \infty)$ yields K.
- Using K, ϕ_j, and related properties gives us $Z_j(t)$.

Implications of Nonparametric Identification:

- Identification does not depend on parametric functional forms or assumed forms of risk dependence (modulo separability of the hazard).
- Highlights the role of regressors in identification in contrast to the Cox-Tsiatis nonidentification result.
- Suggests the possibility of a nonparametric estimator.