Survival Analysis

6. Multivariate Survival

Germán Rodríguez

Princeton University

March 12, 2018

Introduction

Our final topic is multivariate survival analysis, where we have multiple observable outcomes. Areas of application include

- Series of events, such as birth intervals or spells of unemployment, where each individual can experience one or more events in succession
- Kindred lifetimes, such as survival of husband and wife, or survival of children in the same family, where we have related individuals experiencing events
- Competing risks, where each individual can experience one of several types of events, although the models here are more of conceptual than practical interest
- Event history models, involving transitions among different states, for example from single to cohabiting or married, from cohabiting to married or separated, and so on.

We provide some basic definitions and discuss shared frailty models.

Bivariate Survival

We start with two survival times T_1 and T_2. The *joint* survival is

$$S_{12}(t_1, t_2) = \Pr\{T_1 \geq t_1, T_2 \geq t_2\}$$

Here $S_{12}(t, t)$ is the probability that neither unit has failed by t.

The *conditional* survival comes in two variants

$$S_{1|2}(t_1|T_2{=}t_2) = \Pr\{T_1 \geq t_1|T_2 = t_2\}$$

which conditions on unit 2 failing at t_2, and

$$S_{1|2}(t_1|T_2{\geq}t_2) = \Pr\{T_1 \geq t_1|T_2 \geq t_2\}$$

which conditions on unit 2 surviving to just before t_2.

We also have the *marginal* survival functions we already know. If T_1 and T_2 are independent then the joint survival is the product of the marginals.

Bivariate Hazards

The *joint* hazard function is defined as

$$\lambda_{12}(t_1, t_2) = \lim \Pr\{T_1 \in [t_1, t_1+dt), T_2 \in [t_2, t_2+dt) | T_1 \geq t_1, T_2 \geq t_2\}/dt^2$$

the instantaneous rate of failures at t_1 and t_2 given that the units had survived to just before t_1 and t_2.

The *conditional* hazard also comes in two variants

$$\lambda_{1|2}(t_1|T_2{=}t_2) = \lim \Pr\{T_1 \in [t_1, t_1 + dt) | T_1 \geq t_1, T_2 = t_2\}/dt$$

given that unit 2 failed at t_2, and

$$\lambda_{1|2}(t_1|T_2{\geq}t_2) = \lim \Pr\{T_1 \in [t_1, t_1 + dt) | T_1 \geq t_1, T_2 \geq t_2\}/dt$$

given that unit 2 survived to just before t_2.

The two types of conditional hazard together completely determine the joint distribution, see Cox and Oakes (1975).

Finally we have the *marginal* hazards we already know. If T_1 and T_2 are independent the joint hazard is the sum of the marginals.

Frailty Models

A popular approach to modeling multivariate survival is to assume the existence of a shared random effect θ such that T_1 and T_2 are independent given θ:

$$S_{12}(t_1, t_2|\theta) = S_1(t_1|\theta)S_2(t_2|\theta)$$

Typically we assume that frailty acts multiplicatively on the conditional hazard, so that

$$\lambda_j(t|\theta) = \lambda_{0j}(t)\theta \quad \text{and} \quad S_j(t|\theta) = S_{0j}(t)^\theta$$

for some baseline hazard and survival functions with $j = 1, 2$.

Usually the baseline hazard is the same for all failure times. This makes most sense when the events are exchangeable, for example spells of unemployment. Otherwise covariates may be used, for example to distinguish risks for males and females.

Frailty Distributions

A common assumption about shared frailty is that it follows a gamma distribution. If frailty is gamma with mean one and variance σ^2 the joint survival function is

$$S_{12}(t_1, t_2) = \left(\frac{1}{1 + \sigma^2 \Lambda_{01}(t_1) + \sigma^2 \Lambda_{02}(t_2)} \right)^{1/\sigma^2}$$

An alternative assumption that also yields an explicit solution for the survival function is inverse Gaussian frailty.

A third option is to use a non-parametric estimator of the frailty distribution, which leads to a discrete mixture where θ takes values $\theta_1, \ldots \theta_k$ with probabilities $\pi_1, \ldots, \pi_k$ adding to one. In this case

$$S_{12}(t_1, t_2) = \sum_{j=1}^{k} e^{-\theta_j [\Lambda_{01}(t_1) + \Lambda_{02}(t_2)]} \pi_j$$

see Laird (1978) and Heckman and Singer (1984).

Clayton's Model

Clayton (1978) proposed a bivariate survival model where the two conditional hazards for T_1 given $T_2 = t_2$ and given $T_2 \geq t_2$ are proportional:

$$\frac{\lambda_{1|2}(T_1|T_2 = t_2)}{\lambda_{1|2}(T_1|T_2 \geq t_2)} = 1 + \phi$$

In words, the risk for unit 1 at time t_1 given that the other unit failed at t_2 is $1 + \phi$ times the risk at t_1 given that the other unit survived to t_2.

A remarkable result is that this model is exactly equivalent to a multiplicative frailty model with gamma-distributed shared frailty and $\sigma^2 = \phi$.

An important implication of this result is that shared frailty models are clearly identified, as the choice of frailty distribution has observable consequences.

It also gives a new interpretation to σ^2.

Oakes's Interpretation

Oakes (1982) shows that ϕ (and thus σ^2) is closely related to a measure of ordinal association known as Kendall's τ (tau).

Given a bivariate sample of data on (T_1, T_2), Kendall considers all pairs of observations, calls the pair concordant if the rank order is the same and discordant otherwise, and computes

$$\tau = \frac{\text{concordant pairs} - \text{discordant pairs}}{\text{number of pairs}}$$

Oakes extends this to censored data by focusing on pairs where the order can be established, and shows that under gamma frailty

$$E(\hat{\tau}) = \frac{\phi}{\phi + 2}$$

which provides a nice justification for interpreting ϕ (and σ^2) as a measure of ordinal association between kindred lifetimes.

Multivariate Extensions

These ideas extend directly to the multivariate case, please refer to the notes for details. A few important facts:

Clayton shows that with multivariate failures and gamma frailty the ratio of the risk for one unit when m have failed at given durations, to the risk if all had survived to the same durations is

$$1 + m\phi$$

which reduces to $1 + \phi$ in the bivariate case, still with $\phi = \sigma^2$.

Oakes shows that we can interpret the ratio

$$\frac{\phi}{2 + \phi}$$

as a measure of association between any two of the m failure times.

Shared frailty models allow only for positive association between kindred lifetimes, but cover the entire range from independence to maximum possible positive association.

Software Notes

Stata's `streg` fits parametric proportional hazard models with gamma or inverse Gaussian shared frailty. PWE models with log-normal or gamma frailty can also be fit using `xtpoisson`. Cox models with gamma or inverse Gaussian frailty can be fitted with `stcox`, but in my experience this command is very slow.

In R the packages `frailtypack` and the newer `parfm` have functions to fit parametric models with shared frailty. PWE models with log-normal frailty can also be fit via the Poisson trick with `lme4`. The `coxph` function lets you add a `frailty` term to a model formula, but a better approach is Therneau's `coxme`, which includes the `coxme` function to fit mixed Cox survival models with Gaussian random effects.

The computing logs illustrate shared frailty models using a PWE model in Stata and a Cox model in R.

Child Mortality in Guatemala

Our illustrative example uses data on child mortality in Guatemala, first analyzed by Pebley and Stupp (1987) using a PWE model, and then by Guo and Rodríguez (1992) adding gamma frailty at the family level.

The table on the right summarizes parameter estimates. See the computing logs for variable definitions and other details.

The exponentiated coefficients represent subject-specific hazard ratios. The only change of note is the coefficient for previous child death, which goes from 10.3% excess risk to 7.3% lower risk.

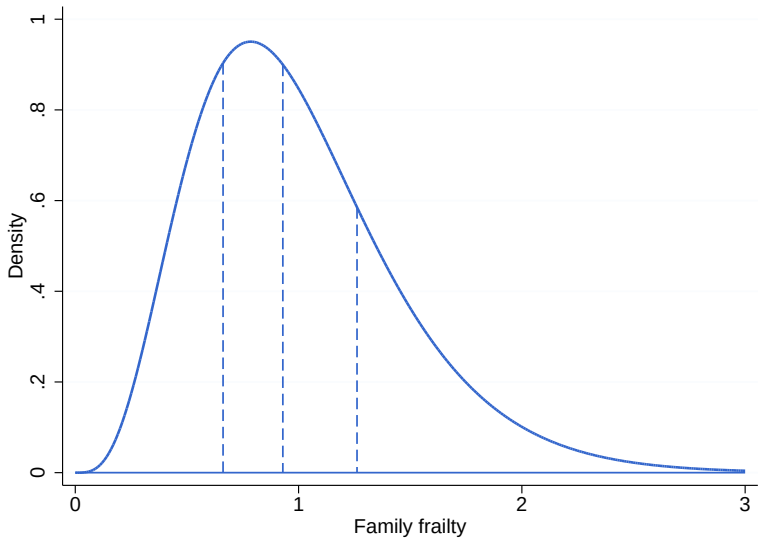
Clearly this variable was acting as a proxy for unobserved family effects, now captured by the random effect.

Variable	pwe	gamma
_t		
a0	0.338	0.371
a1to5	0.025	0.027
a6to11	0.030	0.034
a12to23	0.018	0.020
a24up	0.003	0.004
mage	0.861	0.856
mage2	1.003	1.003
borde	1.064	1.059
pdead	1.103	0.927
p0014	1.714	1.774
p1523	0.885	0.908
p2435	0.772	0.796
p36up	0.676	0.690
i011a1223	2.247	2.210
i011a24p	4.934	4.960
i1223a24p	1.076	1.077
ln_the		
_cons		0.214

The Variance of Frailty

The estimated variance of frailty is 0.214. This implies modest association between the lifetimes of siblings, a rank correlation of 0.097, but translates into substantial Clayton hazard ratios.

The quartiles of the estimated frailty distribution are 0.662, 0.930 and 1.262. Thus, families with frailty at Q1 have 29% lower risk, and those in Q3 have 36% higher risk, than families at median frailty.

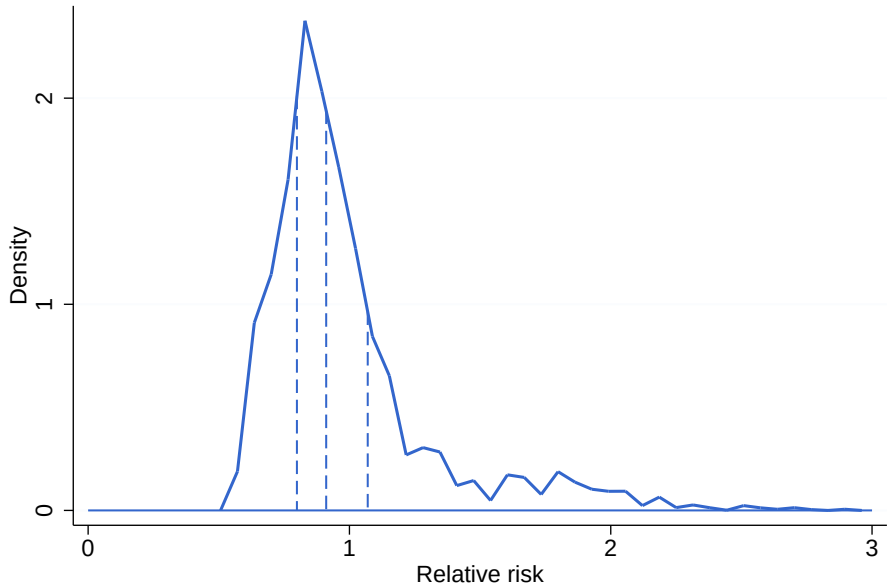


Testing significance of the variance requires care because the null hypothesis is on a boundary of the parameter space. The statistic can be treated as a 50:50 mix of χ^2_0 and χ^2_1, or conservatively as χ^2_1. Here we get 3.3, which is clearly significant.

Observed and Unobserved Effects

It is interesting to compare the magnitude of the estimated unobserved family effects with the relative risks corresponding to observed characteristics of the child and mother.

The figure on the right shows the estimated density of the risks at birth. The quartiles are 0.799, 0.911 and 1.070. Thus, children in Q1 have 12.3% lower, and those in Q3 have 17.5% higher risk than those at the median.



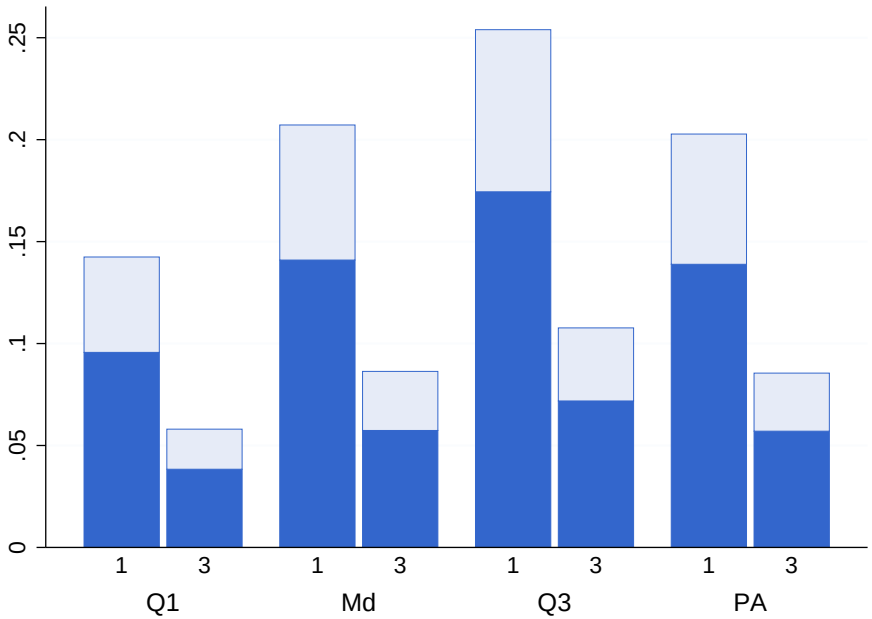
Clearly the unobserved family effects are larger than the observed child and family effects.

See the computing logs for details of the calculations. For the plot I scaled the hazards to have mean one.

Subject-Specific and Population Average Probabilities

We can translate the results into a more convenient scale by calculating subject-specific and population average probabilities. I use preceding birth interval as an example.

These are subject-specific probabilities of infant and child death for a 26-year old mother having a 2nd child, who has not experienced a child death before, has a preceding birth interval of one or three years, and her frailty is in each quartile.



I also show the corresponding population average probabilities. Differences between the average mother and the population average are modest because selection hasn't had much time to operate by ages one and five.

Marginal and Joint Probabilities

The final calculation concerns the marginal and joint probabilities of infant and child death for two children in the same family.

It doesn't make sense to fix the mother's age at 26 unless she has twins, so I did the calculations for a second birth at age 26 and a third birth at age 29. Here are the probabilities for age five

2nd	3rd Child		
Child	died	survived	All
died	.0090	.0765	.0855
survived	.0793	.8351	.9144
All	.0883	.9116	1.000

The odds-ratio for this 2 by 2 table is 1.239, so the odds of one child dying by age five are 23.9% higher if the other child died by age five. (Also, the joint survival is slightly higher than the product of the marginal probabilities.)

Log-Normal Frailty

In the computing logs I also fit this model using log-normal frailty via the equivalence with Poisson regression. The estimates of the parameters are quite robust to the choice of frailty distribution.

A nice feature of log-normal frailty is that we can write the model as

$$\log \lambda(t|x,\theta) = \log \lambda_0(t) + x'\beta + \sigma z$$

where z is standard normal and $\theta = e^{\sigma z}$. This leads to interpreting σ as just another coefficient. In our example $\hat{\sigma} = 0.442$, so a one st.dev. increase in log-frailty is associated with 55.6% higher risk.

The estimated quartiles are Q1=0.742 and Q3=1.348, so these families have 26% lower and 35% higher risk than families at the median. The results are very similar to those under gamma frailty.

A disadvantage of log-normal frailty is the need for Gaussian quadrature to calculate unconditional probabilities.