Introduction to Continuous time to Event Data

The Timing of First Marriage:

Are There Religious Variations?

XIAOHE XU
Mississippi State University

CLARK D. HUDSPETH
Jacksonville State University

JOHN P. BARTKOWSKI
Mississippi State University

Using survey data from a nationally representative sample, this article explores how marriage timing varies across major religious denominations. Survival analysis indicates that net of statistical controls, Catholics, moderate Protestants, conservative Protestants, and Mormons marry significantly earlier than their unaffiliated counterparts. This holds true for women and men. However, no statistical differences emerge between Jews, liberal Protestants, and the unaffiliated. As surmised, auxiliary statistical tests reveal additional religious subcultural variations: (a) Jews tend to marry later than Catholics, conservative Protestants, and Mormons; (b) Catholics also marry later than conservative Protestants and Mormons; (c) no statistical difference surfaces between Mormons and conservative Protestants; and (d) differences between Catholics and liberal Protestants as well as between Jews and liberal Protestants are statistically negligible. These findings systematically support the denominational subcultural paradigm in the case of marriage timing.

Recall. . .

In the continuous time setting we have

$$F(t) = P(T \leq t) \qquad \text{The CDF}$$

Recall...

In the continuous time setting we have

$$
\begin{aligned}
F(t) &= P(T \leq t) && \text{The CDF} \\
f(t) &= F'(t) && \text{The PDF}
\end{aligned}
$$

Recall...

In the continuous time setting we have

$F(t) = P(T \leq t)$	The CDF
$f(t) = F'(t)$	The PDF
$S(t) = P(T > t) = 1 - F(t)$	The Survivor Function

Recall. . .

In the continuous time setting we have

$$
\begin{aligned}
F(t) &= P(T \leq t) && \text{The CDF}\\
f(t) &= F'(t) && \text{The PDF}\\
S(t) &= P(T > t) = 1 - F(t) && \text{The Survivor Function}\\
h(t) &= \lim_{\Delta t \to 0} \frac{P(t \leq T \leq t + \Delta t | T \geq t)}{\Delta t} && \\
&= \frac{f(t)}{S(t)} = -\frac{d}{dx} \log S(t) && \text{The Hazard Function}
\end{aligned}
$$

Recall. . .

In the continuous time setting we have

$$\begin{aligned}
F(t) &= P(T \leq t) && \text{The CDF} \\
f(t) &= F'(t) && \text{The PDF} \\
S(t) &= P(T > t) = 1 - F(t) && \text{The Survivor Function} \\
h(t) &= \lim_{\Delta t \to 0} \frac{P(t \leq T \leq t + \Delta t | T \geq t)}{\Delta t} && \\
&= \frac{f(t)}{S(t)} = -\frac{d}{dx} \log S(t) && \text{The Hazard Function} \\
H(t) &= \int_0^t h(s) ds && \text{The Cumulative Hazard Function.}
\end{aligned}$$

Likelihood Based Approaches

Modelling Censoring and Event Processes

- In order to account for **censoring**, we will typically jointly model (T_i, C_i).

Modelling Censoring and Event Processes

- In order to account for **censoring**, we will typically jointly model (T_i, C_i).
- We let $f(t_i; \theta)$ and $g(c_i; \phi)$ be the respective **densities**.

Modelling Censoring and Event Processes

- In order to account for **censoring**, we will typically jointly model (T_i, C_i).
- We let $f(t_i; \theta)$ and $g(c_i; \phi)$ be the respective **densities**.
- We let $\widetilde{t}_i$ be the **observed event time** and δ_i be the event indicator.

Modelling Censoring and Event Processes

- In order to account for **censoring**, we will typically jointly model (T_i, C_i).
- We let $f(t_i; \theta)$ and $g(c_i; \phi)$ be the respective **densities**.
- We let $\widetilde{t}_i$ be the **observed event time** and δ_i be the event indicator.
- Take $\mathcal{G}(c_i; \phi)$ to be the **survival curve** for censoring.

Modelling Censoring and Event Processes

- In order to account for **censoring**, we will typically jointly model (T_i, C_i).
- We let $f(t_i; \theta)$ and $g(c_i; \phi)$ be the respective **densities**.
- We let $\widetilde{t}_i$ be the **observed event time** and δ_i be the event indicator.
- Take $\mathcal{G}(c_i; \phi)$ to be the **survival curve** for censoring.
- Assume that censoring is **independent** of the event, $T_i \perp C_i$, then ..

$$L_i(\theta, \phi) = f(\widetilde{t}_i; \theta)^{\delta_i} \times S(\widetilde{t}_i; \theta)^{1-\delta_i} \times g(\widetilde{t}_i; \phi)^{1-\delta_i} \times \mathcal{G}(\widetilde{t}_i; \phi)^{\delta_i}.$$

Modelling Censoring and Event Processes

- In order to account for **censoring**, we will typically jointly model (T_i, C_i).
- We let $f(t_i; \theta)$ and $g(c_i; \phi)$ be the respective **densities**.
- We let $\widetilde{t}_i$ be the **observed event time** and δ_i be the event indicator.
- Take $\mathcal{G}(c_i; \phi)$ to be the **survival curve** for censoring.
- Assume that censoring is **independent** of the event, $T_i \perp C_i$, then ..

$$L_i(\theta, \phi) = f(\widetilde{t}_i; \theta)^{\delta_i} \times S(\widetilde{t}_i; \theta)^{1-\delta_i} \times g(\widetilde{t}_i; \phi)^{1-\delta_i} \times \mathcal{G}(\widetilde{t}_i; \phi)^{\delta_i}.$$

- Assume that censoring is **uninformative** so that $\theta \cap \phi = \emptyset$, then ..

$$L_i(\theta, \phi) = f(\widetilde{t}_i; \theta)^{\delta_i} \times S(\widetilde{t}_i; \theta)^{1-\delta_i}.$$

Re-writing based on Hazard

Since $h(t) = \frac{f(t)}{S(t)}$ then

$$L_i(\theta, \phi) = f(\tilde{t}_i; \theta)^{\delta_i} \times S(\tilde{t}_i; \theta)^{1-\delta_i} = h(\tilde{t}_i; \theta)^{\delta_i} S(\tilde{t}_i; \theta).$$

If we specify a **specific distribution** this can be worked out.

Location-Scale Families

Definition

Suppose that a random variable Y can be written as

$$Y = a + bX,$$

for X in the **same family of distributions** as Y.

Then the distribution of X and Y is called a **location-scale family**.

Example

If $X \sim N(0,1)$ and $Y \sim N(\mu, \sigma^2)$, then

$$Y \overset{d}{=} \mu + \sigma X.$$

If $Y \sim \mathrm{Exp}(\rho)$ then

$$\log(Y) = \log \rho + W,$$

where W has an **extreme value distribution**.

Location-Scale “Regression”

Breakdown into Non-Normal Errors

Note that if, for **some transformation**, we have

$$Y = g(T) = \mu + W,$$

where W is considered **an error term**, this looks like a regression model.

Breakdown into Non-Normal Errors

Note that if, for **some transformation**, we have

$$Y = g(T) = \mu + W,$$

where W is considered **an error term**, this looks like a regression model.

Estimation of the distribution becomes estimation of μ.

Parametric Estimation

- We assume **independent censoring**.

Parametric Estimation

- We assume **independent censoring**.
- We assume **non-informative censoring**.

Parametric Estimation

- We assume **independent censoring**.
- We assume **non-informative censoring**.
- We specify a **parametric form** for the distribution, typically a **location-scale family**.

Parametric Estimation

- We assume **independent censoring**.
- We assume **non-informative censoring**.
- We specify a **parametric form** for the distribution, typically a **location-scale family**.
- We find the **ML estimator**.

Parametric Estimation

- We assume **independent censoring**.
- We assume **non-informative censoring**.
- We specify a **parametric form** for the distribution, typically a **location-scale family**.
- We find the **ML estimator**.
- This process will be expanded to allow for **covariates**, later.

Truncation

Sample Bias

Truncation occurs when our sample contains only individuals with $T_i > u$ or $T_i < u$ for some threshold u.

When we have truncation we need to run a **conditional analysis**. That is, we have to condition on $T_i > u$.

This way, we will consider

$$L_i(\theta) = f(t_i | T > u; \theta) = \frac{f(t_i; \theta)}{S(u; \theta)}.$$

Summary

- Continuous time survival data can be analyzed with **parametric likelihood analysis**.

Summary

- Continuous time survival data can be analyzed with **parametric likelihood analysis**.
- **Location-scale families** provide a convenient, regression-type formualtion.

Summary

- Continuous time survival data can be analyzed with **parametric likelihood analysis**.
- **Location-scale families** provide a convenient, regression-type formualtion.
- **Truncation** is a problem which requires a **conditional analysis**.

Summary

- Continuous time survival data can be analyzed with **parametric likelihood analysis**.
- **Location-scale families** provide a convenient, regression-type formualtion.
- **Truncation** is a problem which requires a **conditional analysis**.
- **Standard likelihood** techniques are used.

What is Next?

Coming up. . .

- We will explore **likelihood derivation** in full.

Coming up. . .

- We will explore **likelihood derivation** in full.
- We will work through **location-scale families** and show the implied **log-linear regression** models.

Coming up. . .

- We will explore **likelihood derivation** in full.
- We will work through **location-scale families** and show the implied **log-linear regression** models.
- We will fit **parametric likelihood models** in R.

Coming up...

- We will explore **likelihood derivation** in full.
- We will work through **location-scale families** and show the implied **log-linear regression** models.
- We will fit **parametric likelihood models** in R.
- Then... two final types of models for **continuous time survival analysis**.