MPR Survival Modelling in SPSS (with R)

G. MacKenzie

Centre of Biostatistics, University of Limerick, Ireland, and ENSAI, Statistics & Informatique, Rennes, France

ASSESS, York, October, 2016
Outline

- Introduction
- The GTDL Family
- The GTDL Model
- Crossing Hazards Data
- MPR Modelling
- Conclusions
Intro - MPR Models

These are models with two linear predictors. There is one linear predictor for modelling the *scale* parameter and another for modelling the *shape* parameter.

Thus, for example, we have

\[ \lambda = x'\beta \quad \text{scale} \]
\[ \gamma = z'\alpha \quad \text{shape} \]

Where $x$ and $z$ may contain the same covariates.

The story of the development of MPR follows.
We are all familiar with Cox’s PH semi-parametric regression survival model with hazard function

\[ \lambda(t; x) = \lambda_0(t) \exp(x' \beta) \]  

(1)

where, \( \lambda_0(t) \) is an unknown baseline hazard function and \( \beta \) is a vector or \( p \) unknown regression coefficients measuring the influence of the covariates, \( x \).

Although (1) has become a first choice model, many data sets do not follow the PH assumption.

So the question is then what to do?

Recall that Cox spent considerable effort on the multiple logistic function before turning to the PH model.
The GTDL family

One idea is to base survival on the logistic function - via the GTDL family.

(a) The GTDL PH model

\[ \lambda(t; x) = \pi(t\alpha + \gamma) \exp(x^'\beta) \]  

(b) The GTDL AL Model

\[ \lambda(t; x) = \lambda\phi\pi(\alpha\phi t) \]  

(c) The GTDL model (not PH, not AL)

\[ \lambda(t; x) = \lambda_0\pi(t\alpha + \gamma^*) \]  

where \( \pi(s) = \exp(s)/[1 + \exp(s)] \), \( \phi = \exp(x^'\beta) \) (the accelerator) and \( \gamma^* = x^'\beta \).
Some Properties & Extensions

The family is
- Wholly parametric
- Connected to Cure Rate models
- Generalised to Gamma Frailty family.

There are many papers tracing the development of the GTDL and extensions
The GTDL Model

Recalling that the two-parameter model has

\[ \lambda(t; x) = \lambda_0 \exp(t\alpha + x'\beta)/[1 + \exp(t\alpha + x'\beta)] \]

Then, looking at the shape parameter, \( \alpha \) we find:

\[ \alpha > 0 \Rightarrow \text{GTDL} \]

\[ \alpha = 0 \Rightarrow \text{Exponential} \]

\[ \alpha < 0 \Rightarrow \text{Cure Model*} \]

*Not much done on this yet - but care required.
The GTDL Gamma Frailty model

The conditional hazard is:

$$\lambda(t_i; x_i, u_i) = u_i \exp(t_i\alpha + x_i'\beta) / [1 + \exp(t_i\alpha + x_i'\beta)]$$

for $i = 1, \ldots, n$.

Whence integrating out the $u_i$ the marginal frailty distribution is

$$f(t; x) = \frac{\lambda p_i}{1 - \frac{\lambda\sigma^2}{\alpha}\log_e(q_ig_i)} \left[ 1 - \frac{\lambda\sigma^2}{\alpha} \log_e(q_ig_i) \right]^{-\frac{1}{\sigma^2}}$$

(5)

where:

$$q_i = \frac{1}{1 + \exp(t_i\alpha + x_i'\beta)}$$

$$g_i = 1 + \exp(x_i'\beta)$$

$$p_i = 1 - q_i$$

Other non-Gamma frailty assumptions work.
Survival of Gastric Cancer* patients

*Gastrointestinal Tumor Study Group (GTSG)(1982)
# Models Fitted - Gastric Cancer* patients

**Table**: Models fitted and their marginal mles & (s.e.)

<table>
<thead>
<tr>
<th>Model</th>
<th>$\hat{\alpha}_0$</th>
<th>$\hat{\alpha}_1$</th>
<th>$\hat{\beta}_0$</th>
<th>$\hat{\beta}_1$</th>
<th>$\hat{\sigma}^2$</th>
<th>$\hat{\ell}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cox</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.106</td>
<td>-</td>
<td>-307.47</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>(0.223)</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Cox GF</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-1.146</td>
<td>1.717</td>
<td>-306.50</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>(0.675)</td>
<td>(1.024)</td>
<td></td>
</tr>
<tr>
<td>GTDL</td>
<td>-0.832</td>
<td>-0.094</td>
<td>1.494</td>
<td>-1.380</td>
<td>-</td>
<td>-132.55</td>
</tr>
<tr>
<td></td>
<td>(0.242)</td>
<td>(0.192)</td>
<td>(0.666)</td>
<td>(0.822)</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>GTDL GF</td>
<td>-0.789</td>
<td>3.499</td>
<td>2.380</td>
<td>-4.612</td>
<td>0.400</td>
<td>-127.89</td>
</tr>
<tr>
<td></td>
<td>(0.326)</td>
<td>(1.408)</td>
<td>(1.413)</td>
<td>(1.676)</td>
<td>(0.176)</td>
<td></td>
</tr>
</tbody>
</table>

*Gastrointestinal Tumor Study Group (GTSG)(1982)
Goodness of fit - Gastric Cancer* patients

*Gastrointestinal Tumor Study Group (GTSG)(1982)
MPR Survival Models

These are characterized by two regression models - one for the scale parameter and one for the shape parameter:

**Scale** - $\lambda$
- controls the magnitude of the hazard.
- larger the hazard $\Rightarrow$ the greater risk $\Rightarrow$ the shorter the lifetime.
- usually modelled as a function of covariates.

**Shape** - $\gamma$
- controls the evolution of the hazard.
- constant, increasing, decreasing, bathtub, unimodel.
- usually left constant, but now modelled as a function of covariates.
MPR Weibull set up

The hazard function is given by

\[ \lambda(t; x, z) = \exp(x' \beta) \exp(z' \alpha) t^{\exp(z' \alpha) - 1} \]  

(6)

where the Weibull hazard is

\[ \lambda(t) = \lambda \gamma t^{\gamma - 1} \quad (\lambda > 0, \gamma > 0) \]

and the MPR specification is

\[ \log \lambda = x' \beta \]
\[ \log \gamma = z' \alpha \]

note that \( x \) and \( z \) may contain the same covariates.
MPR Weibull Hazard Ratio

\[ \lambda(t; c, \tilde{x}, \tilde{z}) = \exp(c\beta_1 + \tilde{x}'\beta) \exp(c\alpha_1 + \tilde{z}'\alpha)t^{\exp(c\alpha_1 + \tilde{z}'\alpha) - 1} \]  \tag{7} 

whence the hazard ratio is

\[ \psi(t; \tilde{z}) = \frac{\lambda(t; c = 1, \tilde{x}, \tilde{z})}{\lambda(t; c = 0, \tilde{x}, \tilde{z})} = \exp(\beta_1) \exp(\alpha_1) t^{\exp(\tilde{z}'\alpha)[\exp(\alpha_1) - 1]} \] \tag{8} 

Model is PH for \( \alpha_1 = 0 \)
Hazard Ratio increases for \( \alpha_1 > 0 \)
Hazard Ratio decreases for \( \alpha_1 < 0 \)
Hazard Ratio depends on \( \tilde{z} \)
NI Lung Cancer Study

This study was carried out by my MD student Dr. Paula Wilkinson (PW).

- Multi-source population study of incident cases in NI.
- Followed for c18 months (PW).
- Survival Time = Time from Diagnosis to Death or Censoring.
- Some 693 (77%) had died by the censoring date (30th May, 1993).
- The influence of 9 covariates were analysed: Age, Sex, Treatment, WHO Status, Cell type, Sodium level, Albumen level, Metastases and Smoking category.
Lung Cancer (Wilkinson, 1995): Treatment Model

\[ \Delta_{AIC} = AIC_{PH} - AIC_{MPR} = 3933.5 - 3896.2 = 37.3. \]
Lung Cancer: MPR Variable Selection

MPR Variable Selection: Lung Cancer

<table>
<thead>
<tr>
<th></th>
<th>PH Weibull</th>
<th>MPR Weibull</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>$\beta$</td>
<td>$\beta, \alpha$</td>
</tr>
<tr>
<td>Age Group</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>WHO Status</td>
<td>$\beta$</td>
<td>$\beta$</td>
</tr>
<tr>
<td>Sex</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Smoker</td>
<td>$\beta$</td>
<td>$\alpha$</td>
</tr>
<tr>
<td>Cell Type</td>
<td>$\beta$</td>
<td>$\beta$</td>
</tr>
<tr>
<td>Metastases</td>
<td>$\beta$</td>
<td>$\beta, \alpha$</td>
</tr>
<tr>
<td>Sodium</td>
<td>$\beta$</td>
<td>$\beta$</td>
</tr>
<tr>
<td>Albumen</td>
<td>$\beta$</td>
<td>$\beta, \alpha$</td>
</tr>
</tbody>
</table>

| $AIC$                  | 3723.1     | 3679.7      |
| $\Delta_{AIC}$         | 43.4       | 0.0         |
| $\text{dim}(\theta)$  | 22         | 30          |
Conclusions

- We have traced the development of some NON-PH parametric models.
- From the GTDL family to the MPR class
- These models provide deep insights into the survival process.
- They can handle crossing hazards and wide variety of non-PH behaviour
- So now we consider how to fit them in SPSS via R
Collaborators

Former PhD Students and collaborators...

- Brigadier-General, Dr. Yasin Al Tawarah (Keele) - LPH and LAL.
- Dr. Milica Blagojevic (Keele) - Early Gamma frailty versions of the GTDL.
- Dr. Joseph Lynch (UL) - Later frailty models and structural dispersion.
- Professor IL Do Ha (UL/SK) - h-likelihood GTDL non-Gamma frailty and robustness.
- Dr. Kevin Burke (UL) - MPR models.
Key References -1


Key References - 2


Some Properties & Extensions

The family is

- Wholly parametric
- Connected to Cure Rate models
- Generalised to Gamma Frailty family.

Some history:
- MacKenzie (1996), JRSS D, Introduced basic model
- MacKenzie (1997), SIM, Extended to MV survival
- Blagojevic et al (2003), IWSM, Extended to Gamma Frailty
- Ha & MacKenzie (2010), SM, Robustness of frailty versions
- Lynch & MacKenzie (2014), Springer Book Chapter, analysed Breast cancer survival in the West Midlands, UK.
Some Notation

Let $c$ be a binary covariate common to both scale and shape. Let it be the first covariate, ie, $x_1 = z_1 = c$.

Then let

$$x' \beta = \beta_0 + c \beta_1 + x_2 \beta_2 + \cdots + x_p \beta_p$$

$$= x_1 \beta_1 + \tilde{x}' \beta \quad (9)$$

and

$$z' \alpha = z_1 \alpha_1 + \tilde{z}' \alpha \quad (10)$$

where $\tilde{x}' = (1, 0, x_2, \ldots x_p)$ and $\tilde{z}' = (1, 0, z_2, \ldots z_p)$. 
MPR Lung Cancer CIs

MLE Simulation: Hazard Ratios with CIs

Time (months)

Radiotherapy

Chemo+Radio

Surgery

Chemotherapy

Hazard Ratio

0.0 0.4 0.8 1.2 1.6

0 5 10 15 20

ASSESS, York, October, 2016