Software for the joint modelling of longitudinal and survival data: the JoineR package

Pete Philipson
Collaborative work with Ruwanthi Kolamunnage-Dona, Inês Sousa, Peter Diggle, Rob Henderson, Paula Williamson & Gerwyn Green

useR! conference 2010, NIST, Gaithersburg, MD
Outline

- Longitudinal and survival data
- Joint modelling
- The JoineR package
- Simulations and performance
- Application to real data: liver cirrhosis and CD4 cell counts
- Future work and plans
Longitudinal and survival data

- Longitudinal data
  - Focus on linear mixed-effects model

### Longitudinal sub-model

\[ Y_{ij} = X_1 i \beta_1 + R_1 i (t_{ij}) + \epsilon_{ij} \]

- \( R_1 = D_1 U_1 \) with \( U_1 \) multivariate Gaussian random effects and \( D_1 \) a random effects design matrix

- Survival data
  - Consider two alternatives for the event times \( F \)
    1. Cox proportional hazards
      \[ h_i(t) = h_0(t) \exp(X_2 i \beta_2 + R_2 i) \]
    2. Transformed Gaussian
      \[ F \sim LN(\mu_F, \sigma_F^2) \]

Philipson et al. Joint modelling software - JoineR
Joint modelling

Suitable for a range of objectives

1. Analysing repeated measures $Y$ in the presence of informative drop-out times $F$
2. Analysis of survival times $F$ acknowledging the association with $Y$, which may be a time-varying explanatory covariate subject to measurement error
3. Relationship between $Y$ and $F$ is of joint interest

Examples of two of these will be demonstrated later
Random effects (RE) joint model
- Sub-models linked through common random effects $U$
- Strength of association measured through parameter(s) $\gamma$, i.e. $R_2 = \gamma R_1$
- Model fitting achieved via EM algorithm

Transformation model
- Sub-models formulated as multivariate Gaussian
  \[(Y, \log F) \sim MVN(\mu, \Sigma)\]
- Linked through covariance structure
  \[
  \Sigma = \begin{pmatrix}
  \sigma_Y^2 & g(\theta) \\
  g'(\theta) & \sigma_F^2
  \end{pmatrix}
  \]
- Inverse probability methods - see Scharfstein et al
The JoineR package

- Longitudinal data formatting, visualising and simulation
- Joint model class and plotting function
- Simulating data from joint models
- Transformation model and random effects joint model fitting functions

Philipson et al. Joint modelling software - JoineR
Various simulation studies were carried out to test the software for each possible model. Functions to simulate data are part of the package.

```r
sim_intercept <- simjoint(n = 500, model = "int",
                          gamma = 3, ntms = 5)
```

- Options for continuous/categorical/factors
- Constant or parametric baseline hazard
- Balanced or unbalanced data
- User can choose level of drop-out/censoring and type of latent association
Plotting simulated data: random intercept model

Philipson et al.
Joint modelling software - JoineR
Plotting simulated data: random intercept and slope model

Philipson et al.

Joint modelling software - JoineR
Simulation study: results for RE model

**Intercept only model:** \( R_1 = U_0, \ R_2 = \gamma R_1 \)

<table>
<thead>
<tr>
<th></th>
<th>( \beta_{11} )</th>
<th>( \beta_{12} )</th>
<th>( \beta_{21} )</th>
<th>( \beta_{22} )</th>
<th>( \gamma )</th>
<th>( \sigma^2_0 )</th>
<th>( \sigma^2_{\epsilon} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>250</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.01</td>
<td>0.98</td>
<td>0.49</td>
</tr>
<tr>
<td>500</td>
<td>1.00</td>
<td>1.00</td>
<td>0.99</td>
<td>0.99</td>
<td>0.98</td>
<td>1.00</td>
<td>0.50</td>
</tr>
<tr>
<td>1000</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.99</td>
<td>1.00</td>
<td>0.50</td>
</tr>
<tr>
<td>True</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
</tr>
</tbody>
</table>

*Table:* Simulation results from intercept only model

**Intercept and slope models:** \( R_1 = U_0 + U_1 t, \ R_2 = \gamma R_1 \)

<table>
<thead>
<tr>
<th></th>
<th>( \beta_{11} )</th>
<th>( \beta_{12} )</th>
<th>( \beta_{21} )</th>
<th>( \beta_{22} )</th>
<th>( \gamma )</th>
<th>( \sigma^2_0 )</th>
<th>( \sigma^2_{\epsilon} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>250</td>
<td>1.00</td>
<td>0.99</td>
<td>0.99</td>
<td>1.00</td>
<td>0.25</td>
<td>0.99</td>
<td>1.99</td>
</tr>
<tr>
<td>500</td>
<td>1.00</td>
<td>0.99</td>
<td>1.01</td>
<td>1.00</td>
<td>0.25</td>
<td>0.99</td>
<td>1.99</td>
</tr>
<tr>
<td>1000</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.25</td>
<td>1.00</td>
<td>2.00</td>
</tr>
<tr>
<td>True</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.25</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

*Table:* Simulation results from intercept and slope model
Data on almost 500 patients from a randomised clinical trial of prednisone for liver cirrhosis patients. Further details can be found in Andersen et al.

- We can fit a joint model using JoineR

```r
fit_int_slope <- joint(Y ~ int + P + tt + P_tt + tt0 + P_tt0, 
                       "id", "tt", Surv(s,cen)~sP, 
                       data = liverJointData, 
                       longsep = T, survsep = T)

fit_int_slope <- joint(Y ~ int + P + tt + P_tt + tt0 + P_tt0, 
                       "id", "tt", Surv(s,cen)~sP, 
                       data = liverJointData, 
                       longsep = T, survsep = T, gpt = 15)

fit_quadratic <- joint(Y ~ int + P + tt + P_tt + tt0 + P_tt0, 
                       "id", "tt", Surv(s,cen)~sP, 
                       data = liverJointData, 
                       model = "quad", longsep = T, survsep = T)
```
Liver cirrhosis data

Time

Censored

Failed

Philipson et al.
Joint modelling software - JoineR
### Application: liver cirrhosis data (ctd.)

#### Parameter Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Separate analysis</th>
<th>Joint analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Longitudinal</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>69.99</td>
<td>70.31</td>
</tr>
<tr>
<td>Treatment, $P$</td>
<td>11.63</td>
<td>11.28</td>
</tr>
<tr>
<td>Time, $t$</td>
<td>1.33</td>
<td>0.25</td>
</tr>
<tr>
<td>$P \times t$</td>
<td>-1.59</td>
<td>-1.24</td>
</tr>
<tr>
<td>$t = 0, B$</td>
<td>-1.15</td>
<td>-1.48</td>
</tr>
<tr>
<td>$P \times B$</td>
<td>-11.80</td>
<td>-11.45</td>
</tr>
<tr>
<td><strong>Survival</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment</td>
<td>-0.10</td>
<td>-0.08</td>
</tr>
<tr>
<td><strong>Association</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma$</td>
<td>-</td>
<td>-0.04</td>
</tr>
</tbody>
</table>

Philipson et al.

Joint modelling software - JoineR
Data collected on 467 HIV-infected patients to compare efficacy and safety of two antiretroviral drugs. Further details in Guo & Carlin and data available from Brad Carlin’s software page.

- We can fit a joint model using **JoineR**

```r
fit_int <- joint(Y~ tt + tt_drug + gen + prev + strat,``id'',``tt'',
                 Surv(s,cen)~sgrp + sgen + sprev + sstrat, model = ``int'',
                 data = CarlinJointData, longsep = T, survsep = T)
```
## Application II: CD4 cell count data (ctd.)

### Parameter Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Separate analysis</th>
<th>Joint analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Longitudinal</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>8.00</td>
<td>7.96</td>
</tr>
<tr>
<td>Time</td>
<td>-0.16</td>
<td>-0.17</td>
</tr>
<tr>
<td>Time × Drug</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Gender</td>
<td>-0.15</td>
<td>-0.12</td>
</tr>
<tr>
<td>Prev OI</td>
<td>-2.31</td>
<td>-2.34</td>
</tr>
<tr>
<td>Stratum</td>
<td>-0.11</td>
<td>-0.14</td>
</tr>
<tr>
<td><strong>Survival</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drug</td>
<td>0.22</td>
<td>0.30</td>
</tr>
<tr>
<td>Gender</td>
<td>-0.17</td>
<td>-0.17</td>
</tr>
<tr>
<td>Prev OI</td>
<td>0.65</td>
<td>0.65</td>
</tr>
<tr>
<td>Stratum</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td><strong>Association</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma$</td>
<td>-</td>
<td>-0.23</td>
</tr>
</tbody>
</table>

Philipson et al.  
Joint modelling software - JoineR
Future work

- Deposit on CRAN
- Added flexibility for latent structure in model fitting - user can choose \( D_1, D_2 \)
- More flexibility in simulation routines
- See the project website at http://www.liv.ac.uk/joine-r/index.html


