Software Implementation for Wulfsohn-Tsiatis

Pete Philipson

University of Newcastle Upon Tyne

Outline

- Longitudinal data

Outline

- Longitudinal data
- Survival data

Outline

- Longitudinal data
- Survival data
- Joint model

Outline

- Longitudinal data
- Survival data
- Joint model
- Computational implementation

Longitudinal data

- Random effects model

$$Y_{ij} = \boldsymbol{x}_{1i}(t_{ij})^T \boldsymbol{\beta}_1 + U_{0i} + U_{1i} t_{ij} + \epsilon_{ij}$$

Longitudinal data

- Random effects model

$$Y_{ij} = \boldsymbol{x}_{1i}(t_{ij})^T \boldsymbol{\beta}_1 + U_{0i} + U_{1i} t_{ij} + \epsilon_{ij}$$

  - where

$$\begin{pmatrix} U_{0i} \\ U_{1i} \end{pmatrix} \sim \left(N \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{00} & \sigma_{01} \\ \sigma_{10} & \sigma_{11} \end{pmatrix} \right),$$

  or, alternatively,

$$\boldsymbol{U_i} \sim N(\boldsymbol{0}, \boldsymbol{V}).$$

Longitudinal data

- Random effects model

$$Y_{ij} = \boldsymbol{x}_{1i}(t_{ij})^T \boldsymbol{\beta}_1 + U_{0i} + U_{1i}t_{ij} + \epsilon_{ij}$$

  - where

$$\begin{pmatrix} U_{0i} \\ U_{1i} \end{pmatrix} \sim \left(N \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{00} & \sigma_{01} \\ \sigma_{10} & \sigma_{11} \end{pmatrix} \right),$$

    or, alternatively,

$$\boldsymbol{U_i} \sim N(\boldsymbol{0}, \boldsymbol{V}).$$

  - Also,

$$\epsilon_{ij} \sim N(0, \sigma_\epsilon^2).$$

Survival data

- Survival data with hazard

$$\alpha(t; \boldsymbol{x}_{2i}, \boldsymbol{U_i}) = \alpha_0(t)\text{exp}\{\boldsymbol{x}_{2i}(t)^T\boldsymbol{\beta}_2 + \gamma(U_{0i} + U_{1i}t) + U_{2i}\}$$

Survival data

- Survival data with hazard

$$\alpha(t; \boldsymbol{x}_{2i}, \boldsymbol{U_i}) = \alpha_0(t)\text{exp}\{\boldsymbol{x}_{2i}(t)^T\boldsymbol{\beta}_2 + \gamma(U_{0i} + U_{1i}t) + U_{2i}\}$$

- Observe time s_i with associated failure indicator

$$\Delta_i = \begin{cases} 0 & \text{censored,} \\ 1 & \text{failure.} \end{cases}$$

Survival data

- Survival data with hazard

$$\alpha(t; \boldsymbol{x}_{2i}, \boldsymbol{U_i}) = \alpha_0(t)\text{exp}\{\boldsymbol{x}_{2i}(t)^T\boldsymbol{\beta}_2 + \gamma(U_{0i} + U_{1i}t) + U_{2i}\}$$

- Observe time s_i with associated failure indicator

$$\Delta_i = \begin{cases} 0 & \text{censored,} \\ 1 & \text{failure.} \end{cases}$$

- Joint model - idea is to combine longitudinal and survival elements in a larger meta-model, see Wulfsohn and Tsiatis (1997)

Joint model I

- Observed data likelihood

$$\prod_{i=1}^{m}\left[\int_{-\infty}^{\infty}\left\{\prod_{j=1}^{n_i} f(y_{ij} \mid \cdot)\right\} f(s_i, \Delta_i \mid \cdot) f(\boldsymbol{U}_i \mid \boldsymbol{V}) \; d\boldsymbol{U}_i\right]$$

Joint model I

- Observed data likelihood

$$\prod_{i=1}^{m}\left[\int_{-\infty}^{\infty}\left\{\prod_{j=1}^{n_i} f(y_{ij} \mid \cdot)\right\} f(s_i, \Delta_i \mid \cdot) f(\boldsymbol{U}_i \mid \boldsymbol{V}) \; d\boldsymbol{U}_i\right]$$

- Complete data likelihood

Joint model I

- Observed data likelihood

$$\prod_{i=1}^{m} \left[\int_{-\infty}^{\infty} \left\{ \prod_{j=1}^{n_i} f(y_{ij} \mid \cdot) \right\} f(s_i, \Delta_i \mid \cdot) f(\boldsymbol{U}_i \mid \boldsymbol{V}) \; d\boldsymbol{U}_i \right]$$

- Complete data likelihood
- EM algorithm

Joint model I

- Observed data likelihood

$$\prod_{i=1}^{m}\left[\int_{-\infty}^{\infty}\left\{\prod_{j=1}^{n_i} f(y_{ij}\mid\cdot)\right\} f(s_i,\Delta_i\mid\cdot) f(\boldsymbol{U}_i\mid\boldsymbol{V})\; d\boldsymbol{U}_i\right]$$

- Complete data likelihood
- EM algorithm
- Maximisation (M) step

Joint model I

- Observed data likelihood

$$\prod_{i=1}^{m}\left[\int_{-\infty}^{\infty}\left\{\prod_{j=1}^{n_i} f(y_{ij} \mid \cdot)\right\} f(s_i, \Delta_i \mid \cdot) f(\boldsymbol{U}_i \mid \boldsymbol{V}) \; d\boldsymbol{U}_i\right]$$

- Complete data likelihood
- EM algorithm
- Maximisation (M) step
  - score equations

Joint model I

- Observed data likelihood

$$\prod_{i=1}^{m}\left[\int_{-\infty}^{\infty}\left\{\prod_{j=1}^{n_i} f(y_{ij} \mid \cdot)\right\} f(s_i, \Delta_i \mid \cdot) f(\boldsymbol{U}_i \mid \boldsymbol{V}) \; d\boldsymbol{U}_i\right]$$

- Complete data likelihood
- EM algorithm
- Maximisation (M) step
  - score equations
  - maximum likelihood estimates

Joint model I

- Observed data likelihood

$$\prod_{i=1}^{m} \left[\int_{-\infty}^{\infty} \left\{ \prod_{j=1}^{n_i} f(y_{ij} \mid \cdot) \right\} f(s_i, \Delta_i \mid \cdot) f(\boldsymbol{U}_i \mid \boldsymbol{V}) \; d\boldsymbol{U}_i \right]$$

- Complete data likelihood
- EM algorithm
- Maximisation (M) step
  - score equations
  - maximum likelihood estimates
  - Newton-Raphson iterative algorithm

Joint model II

- Expectation (E) step

Joint model II

- Expectation (E) step
  - conditional expectations of the form

$$E\{h(\boldsymbol{U}_i) \mid s_i, \Delta_i, \boldsymbol{Y}_i, \hat{\boldsymbol{\theta}}\}$$

where

$$\boldsymbol{\theta} = (\boldsymbol{\beta_1}, \boldsymbol{\beta_2}, \boldsymbol{V}, \sigma_\epsilon^2, \gamma, \boldsymbol{\alpha}_0)$$

Joint model II

- Expectation (E) step
  - conditional expectations of the form

$$E\{h(\boldsymbol{U}_i) \mid s_i, \Delta_i, \boldsymbol{Y}_i, \hat{\boldsymbol{\theta}}\}$$

where

$$\boldsymbol{\theta} = (\boldsymbol{\beta_1}, \boldsymbol{\beta_2}, \boldsymbol{V}, \sigma_\epsilon^2, \gamma, \boldsymbol{\alpha}_0)$$

  - require appropriate density

$$f(\boldsymbol{U}_i \mid s_i, \Delta_i, \boldsymbol{Y}_i, \hat{\boldsymbol{\theta}})$$

Joint model II

- Expectation (E) step
  - conditional expectations of the form

$$E\{h(\boldsymbol{U}_i) \mid s_i, \Delta_i, \boldsymbol{Y}_i, \hat{\boldsymbol{\theta}}\}$$

where

$$\boldsymbol{\theta} = (\boldsymbol{\beta_1}, \boldsymbol{\beta_2}, \boldsymbol{V}, \sigma_\epsilon^2, \gamma, \boldsymbol{\alpha}_0)$$

  - require appropriate density

$$f(\boldsymbol{U}_i \mid s_i, \Delta_i, \boldsymbol{Y}_i, \hat{\boldsymbol{\theta}})$$

  - transformation of variables

Joint model II

- Expectation (E) step
  - conditional expectations of the form

$$E\{h(\boldsymbol{U}_i) \mid s_i, \Delta_i, \boldsymbol{Y}_i, \hat{\boldsymbol{\theta}}\}$$

    where

$$\boldsymbol{\theta} = (\boldsymbol{\beta_1}, \boldsymbol{\beta_2}, \boldsymbol{V}, \sigma_\epsilon^2, \gamma, \boldsymbol{\alpha}_0)$$

  - require appropriate density

$$f(\boldsymbol{U}_i \mid s_i, \Delta_i, \boldsymbol{Y}_i, \hat{\boldsymbol{\theta}})$$

  - transformation of variables
  - Gauss-Hermite quadrature

Joint model II

- Expectation (E) step
  - conditional expectations of the form

$$E\{h(\boldsymbol{U}_i) \mid s_i, \Delta_i, \boldsymbol{Y}_i, \hat{\boldsymbol{\theta}}\}$$

where

$$\boldsymbol{\theta} = (\boldsymbol{\beta_1}, \boldsymbol{\beta_2}, \boldsymbol{V}, \sigma_\epsilon^2, \gamma, \boldsymbol{\alpha}_0)$$

  - require appropriate density

$$f(\boldsymbol{U}_i \mid s_i, \Delta_i, \boldsymbol{Y}_i, \hat{\boldsymbol{\theta}})$$

  - transformation of variables
  - Gauss-Hermite quadrature
- Implementation…

Software implementation

- Begin with separate analyses of longitudinal and survival components using C

Software implementation

- Begin with separate analyses of longitudinal and survival components using C
  - Use standard likelihood methods and EM algorithm to obtain starting values for joint analysis

Software implementation

- Begin with separate analyses of longitudinal and survival components using C
  - Use standard likelihood methods and EM algorithm to obtain starting values for joint analysis
  - Run EM until convergence achieved

Software implementation

- Begin with separate analyses of longitudinal and survival components using C
  - Use standard likelihood methods and EM algorithm to obtain starting values for joint analysis
  - Run EM until convergence achieved
- Generalisation of code to fit alternative data-sets

Software implementation

- Begin with separate analyses of longitudinal and survival components using C
  - Use standard likelihood methods and EM algorithm to obtain starting values for joint analysis
  - Run EM until convergence achieved
- Generalisation of code to fit alternative data-sets
- Flexibility in choice of latent association, frailty etc.

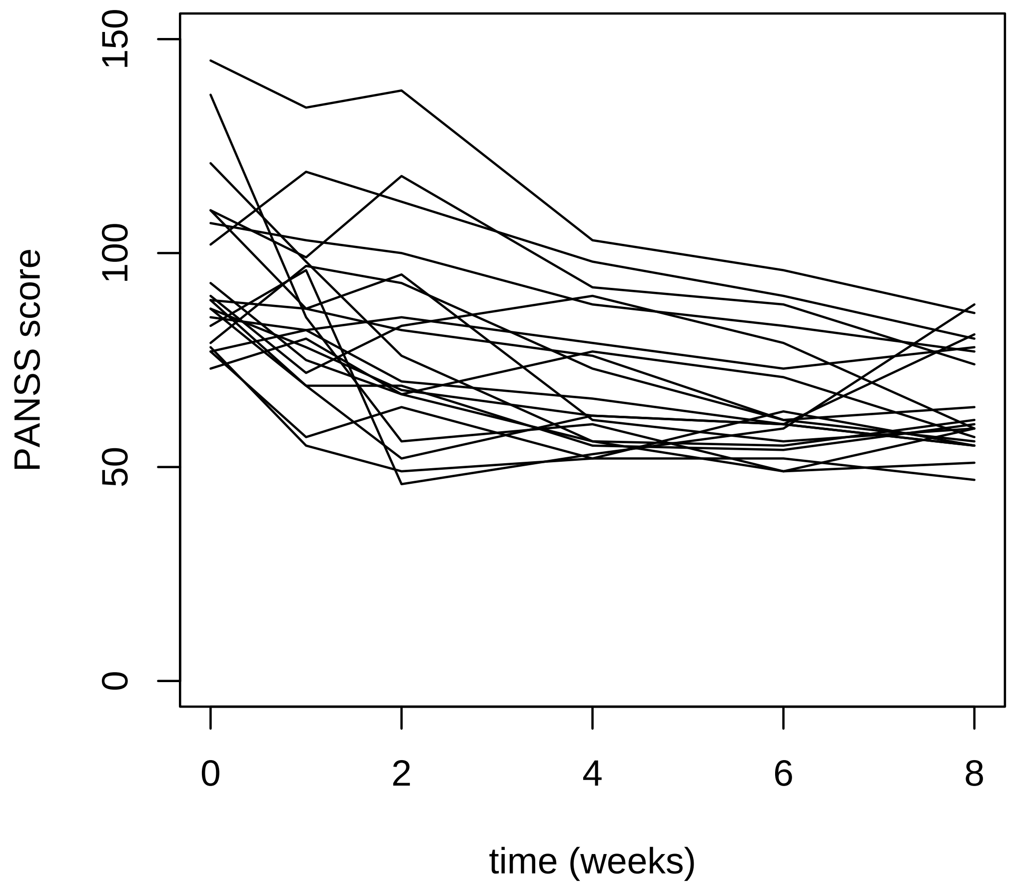
Software implementation

- Begin with separate analyses of longitudinal and survival components using C
  - Use standard likelihood methods and EM algorithm to obtain starting values for joint analysis
  - Run EM until convergence achieved
- Generalisation of code to fit alternative data-sets
- Flexibility in choice of latent association, frailty etc.
- Running C from R

Software implementation

- Begin with separate analyses of longitudinal and survival components using C
  - Use standard likelihood methods and EM algorithm to obtain starting values for joint analysis
  - Run EM until convergence achieved
- Generalisation of code to fit alternative data-sets
- Flexibility in choice of latent association, frailty etc.
- Running C from R
- Comprehensive R Archive Network (CRAN)

PANSS data



Bibliography

Wulfsohn, M. S. and Tsiatis, A. A. (1997). A Joint Model for Survival and Longitudinal Data Measured with Error. *Biometrics*, **53**, 330-339.