DSBS Annual meeting
Bivariate pseudo-observations for recurrent event analysis with terminal events

Julie Kjærulff Furberg

29. April 2022

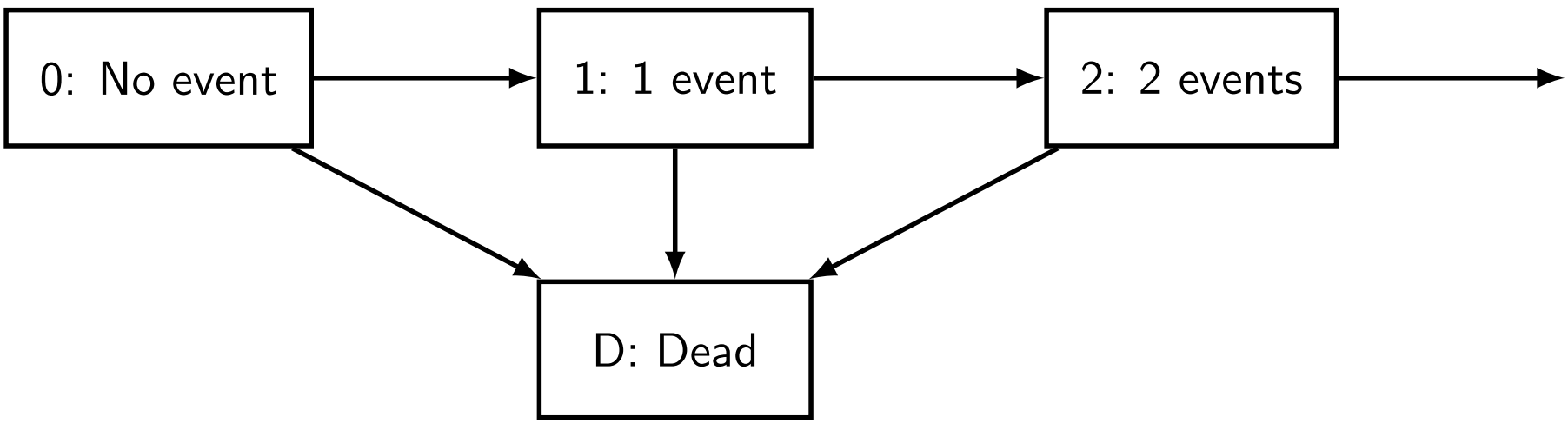
Outline

My PhD project :-)

- Industrial PhD at University of Copenhagen (UCPH) and Novo Nordisk (NN) (Feb 2020- (+ maternity leave x 1))
- Title of PhD project: "*Treatment effect measures for recurrent event endpoints with and without presence of terminal events*"
- Supervisors: Per Kragh Andersen (UCPH), Henrik Ravn (NN) and Trine Saugstrup (NN)

Recurrent events in the presence of terminal events

Desire: To model recurrent event data in a world where there are competing deaths. The underlying stochastic process could be described by this graph,



Topic for today: Bivariate pseudo-observations

- I will motivate our method (paper below) by an example based on simulated data

Lifetime Data Analysis
https://doi.org/10.1007/s10985-021-09533-5

Check for updates

Bivariate pseudo-observations for recurrent event analysis with terminal events

Julie K. Furberg[1] · Per K. Andersen[2] · Sofie Korn[3] · Morten Overgaard[4] · Henrik Ravn[1]

Received: 25 March 2021 / Accepted: 4 September 2021

Abstract
The analysis of recurrent events in the presence of terminal events requires special attention. Several approaches have been suggested for such analyses either using intensity models or marginal models. When analysing treatment effects on recurrent events in controlled trials, special attention should be paid to competing deaths and their impact on interpretation. This paper proposes a method that formulates a marginal model for recurrent events and terminal events simultaneously. Estimation is based

Notation

- Let D^* denote the survival time and let $N^*(t)$ denote the number of recurrent events per subject by time t
- We assume that we have a single binary treatment covariate, Z
- Due to right-censoring, we observe $\{N(\cdot), D, \delta, Z\}$ for $N(t) = N^*(t \wedge C)$, $D = D^* \wedge C$, $\delta = I(D^* \leq C)$
- Let $\mu(t)$ denote the marginal mean function (in the presence of competing risks) and $S(t)$ denote the survival probability given by

$$\mu(t) = E(N^*(t)) = \int_0^t S(u^-) dR(u)$$

$$S(t) = P(D^* > t)$$

where $dR(t) = E(dN^*(t) \mid D^* \geq t)$

- For each individual, $i = 1, \ldots, n$, we observe $\{N_i(\cdot), D_i, \delta_i, Z_i\}$

A cautionary tale: Simulated data

- We simulate data using two frailty models such that

$$\lambda^D(t \mid Z, \nu) = \nu \exp(\gamma_D Z)\lambda_0^D$$
$$dR(t \mid Z, \nu) = \nu \exp\left(Z\left(\beta - \nu\lambda_0^D t(1 - \exp(\gamma_D))\right)\right)\, dt$$

  where $\nu > 0$ is a subject specific frailty term generated from a positive stable distribution with Laplace transform $\exp(-\nu^\rho)$, $\rho \in (0, 1]$
- ρ controls the association between recurrent events and death
- After some math, you can show that

$$\mu(t \mid Z) = \int_0^t S(u \mid Z)\, dR(u \mid Z) = \mu_0(t)\exp(\beta Z)$$
$$S(t \mid Z) = \exp\left(-\Lambda_0^D(t)\exp(\gamma Z)\right)$$

  where

$$\gamma = \gamma_D\rho, \quad \Lambda_0^D(t) = \left(\lambda_0^D t\right)^\rho, \quad \mu_0(t) = \frac{1}{\lambda_0^D}\left(-\exp\left(-\left(\lambda_0^D t\right)^\rho\right) + 1\right)$$

- Censoring is assumed to be uniform on $[0, 5]$

A cautionary tale: Simulated data

- Let $n = 1000$, $\beta = 0$, $\gamma_D = -2$, $\rho = 0.9$, $\lambda_0^D = 0.25$
- How does the data look?

		Status		
		0 (censoring)	1 (recurrent event)	2 (death)
Z	0	263	856	250
	1	433	827	54

Table: Overview of simulated data

- More deaths with $Z = 0$ as opposed to $Z = 1$
- Slightly more recurrent events with $Z = 0$ versus $Z = 1$

A cautionary tale: Simulated data

- Assume that we wish to model the recurrent events and estimate a treatment effect (a mean ratio)
- A naïve way to "handle" the competing deaths would be to treat these as censored when fitting a model for the marginal mean for the recurrent events
- We know that the difference between treatments in recurrent events is solely driven by the correlation between death and recurrent events (through ρ) since we have $\beta = 0$

A cautionary tale: Simulated data

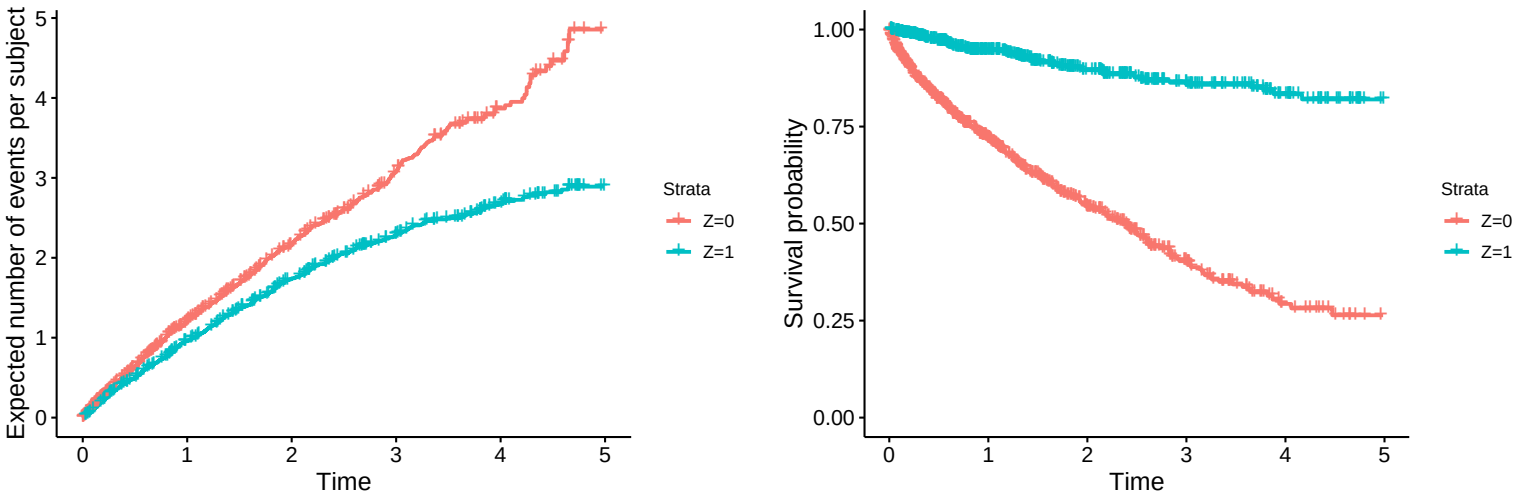


Figure:
LHS: Nelson-Aalen estimates of the marginal mean function per treatment - *incorrectly* censoring for death
RHS: Kaplan-Meier estimates of survival probability per treatment

A cautionary tale: Simulated data

- Similarly, it is possible to fit a marginal (LWYY[1]) model of the form

$$\mu(t \mid Z) = \mu_0(t) \exp(\beta Z)$$

  where one treats the deaths as censoring
- This leads to $\hat{\beta} = -0.3056$ ($\text{se}(\hat{\beta}) = 0.0614$) with a corresponding Wald test statistic of $z = -4.976$ ($P(\tilde{Z} > |z|) < 0.001$)
- This implies that we expect $\exp(-0.31) = 0.73$ fewer events on average with $Z = 1$ versus $Z = 0$
- But we know that $\beta = 0$...
- We are misjudging the situation (type I error). This is due to the relationship between the recurrent events and death. More recurrent events imply more deaths (and vice versa)

[1] Lin, Wei, Ying and Yang - *Semiparametric regression for the mean and rate functions of recurrent events* (2000)

A two-dimensional approach to a complex reality: The bivariate pseudo-observation model

- We suggest a two-dimensional modelling procedure that enables inference on recurrent events and death simultaneously
- $(\mu(t), S(t))$ is modelled conditional on the covariates Z using a regression model based on bivariate pseudo-observations
- Here, competing deaths are addressed in an appropriate way in the analysis of recurrent events
- Moreover, we gain knowledge on both $\mu(t)$ and $S(t)$ as well as their mutual relationship

Quick intro to pseudo-observations[2]

- Let X_i denote a survival time for subject i
- We are interested in some function, f, of $X = (X_1, \ldots, X_n)$
- Let the parameter of interest, $\theta = E(f(X))$
- Assume that a sufficiently nice estimator $\hat{\theta}$ of θ exists
- Then, the pseudo-observation for individual i is given by

$$\hat{\theta}_i = n\hat{\theta} - (n-1)\hat{\theta}^{-i},$$

  where $\hat{\theta}^{-i}$ denotes the the estimate when leaving out individual i
- Now, $f(X_i)$ can be replaced by the pseudo-observations $\hat{\theta}_i$
- Finally, $\hat{\theta}_i$ can be used as the outcome variable (instead of $f(X_i)$) in a generalized linear regression model with some link function g

$$g(E(f(X) \mid Z)) = \xi^T Z$$

[2] You can read more here: Andersen and Perme - *Pseudo-observations in survival analysis* (2010)

Bivariate pseudo-observation model I

- The bivariate pseudo-observation model considers the marginal target parameter

$$\theta = \begin{pmatrix} \mu(t) \\ S(t) \end{pmatrix} = \begin{pmatrix} E(N^*(t)) \\ E(I(D^* > t)) \end{pmatrix}$$

- As an estimator for θ, we will consider

$$\hat{\theta} = \begin{pmatrix} \hat{\mu}(t) \\ \hat{S}(t) \end{pmatrix}$$

where

$$\hat{\mu}(t) = \int_0^t \hat{S}(u^-)\, d\hat{R}(u)$$

Here $\hat{R}(t)$ denotes the Nelson-Aalen estimator of $R(t)$ and $\hat{S}(t)$ denotes the Kaplan-Meier estimator of $S(t)$

Bivariate pseudo-observation model II

- For a given time $t \in [0, \tau]$, the pseudo-observation for subject i is given by

$$\hat{\theta}_i = \begin{pmatrix} \hat{\mu}_i(t) \\ \hat{S}_i(t) \end{pmatrix} = \begin{pmatrix} n\hat{\mu}(t) - (n-1)\hat{\mu}^{-i}(t) \\ n\hat{S}(t) - (n-1)\hat{S}^{-i}(t) \end{pmatrix}$$

where $\hat{\mu}^{-i}(\cdot)$ and $\hat{S}^{-i}(\cdot)$ are the estimates based on leaving individual i out from the computation

- We formulate the following generalised linear model,

$$g\begin{pmatrix} \mu(t \mid Z) \\ S(t \mid Z) \end{pmatrix} = \xi^T Z$$

with $g(x, y) = (\log(x), \text{cloglog}(y)) = (\log(x), \log(-\log(y)))$ and for a covariate vector Z

Bivariate pseudo-observation model III

- We assume that

$$\mu(t \mid Z) = \mu_0(t) \exp(\beta^T Z)$$
$$S(t \mid Z) = \exp(-\Lambda_0^D(t) \exp(\gamma^T Z))$$

- Then

$$g\begin{pmatrix}\mu(t \mid Z) \\ S(t \mid Z)\end{pmatrix} = \begin{pmatrix}\log\left(\mu_0(t)\right) + \beta^T Z \\ \log\left(\Lambda_0^D(t)\right) + \gamma^T Z\end{pmatrix}$$

- The model parameters ξ is estimated using generalised estimating equations (GEE) by doing a regression of $\hat{\theta}_i$ on Z_i

Bivariate pseudo-observation model IV

- In the paper, we argue that

$$\begin{pmatrix} \hat{\beta} \\ \hat{\gamma} \end{pmatrix} \overset{as}{\sim} \mathcal{N}\left(\begin{pmatrix} \beta \\ \gamma \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} \right), \quad \Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix},$$

  where $\hat{\beta}$ and $\hat{\gamma}$ denote the estimates of β and γ from the GEE procedure
- Σ can be estimated using a sandwich covariance estimator implemented in standard software (however these are expected to be slightly conservative!)
- The bivariate normality makes everything nice in terms of hypothesis testing

Hypothesis tests I

- **Local tests**
  1. $H_0 : \beta = 0$ versus $H_a : \beta \neq 0$. Reject H_0 if $P(|U| > |\tilde{\beta}|) \leq \alpha$, where $\tilde{\beta} = \frac{\hat{\beta}}{\sqrt{\text{var}(\hat{\beta})}}$ and $U \sim \mathcal{N}(0,1)$
  2. $H_0 : \gamma = 0$ versus $H_a : \gamma \neq 0$. Reject H_0 if $P(|U| > |\tilde{\gamma}|) \leq \alpha$, where $\tilde{\gamma} = \frac{\hat{\gamma}}{\sqrt{\text{var}(\hat{\gamma})}}$ and $U \sim \mathcal{N}(0,1)$
- **Global test**
  1. $H_0 : \beta = \gamma = 0$ versus $H_a : \beta \neq 0$ or $\gamma \neq 0$. Let

$$T_{global} = \begin{pmatrix} \hat{\beta} & \hat{\gamma} \end{pmatrix} \hat{\Sigma}^{-1} \begin{pmatrix} \hat{\beta} \\ \hat{\gamma} \end{pmatrix}$$

  H_0 is rejected if $P(Y > T_{global}) \leq \alpha$, where Y is χ^2-distributed with 2 degrees of freedom

Hypothesis tests II

- **Sequential test**
  1. Order the values of $\tilde{\beta}$ and $\tilde{\gamma}$ according to absolute size. Assume that $|\tilde{\beta}| < |\tilde{\gamma}|$ without loss of generality
  2. Let $H_l : \eta = 0$ be the ordered hypotheses for $\eta = \{\beta, \gamma\}$ with $l = \{1, 2\}$ representing the order
  3. H_1 is rejected if $P(|\max(W_1, W_2)| > |\tilde{\beta}|) \leq \alpha$ where (W_1, W_2) is a bivariate standard normal variable with unit variance and correlation given by the correlation between $\tilde{\beta}$ and $\tilde{\gamma}$
  4. If H_1 is rejected, H_2 is tested. H_2 is rejected if $P(|W_2| > |\tilde{\gamma}|) \leq \alpha$

  This procedure maintains type I error control

The cautionary tale revisited

- If we fit the bivariate pseudo-observation model to the simulated data at time points $t = (2, 3, 4)$, we obtain the following estimates

$$\hat{\xi} = \begin{pmatrix} \hat{\beta} \\ \hat{\gamma} \end{pmatrix} = \begin{pmatrix} 0.0426 \\ -1.7848 \end{pmatrix}, \qquad \hat{\Sigma} = \begin{pmatrix} 0.0041 & -0.0014 \\ -0.0014 & 0.0258 \end{pmatrix}$$

- **Local tests**
  1. $\tilde{\beta} = \frac{\hat{\beta}}{\hat{\sigma}_{11}} = \frac{0.0426}{0.0645} = 0.6619$ (p-value of 0.51)
  2. $\tilde{\gamma} = \frac{\hat{\gamma}}{\hat{\sigma}_{22}} = \frac{-1.7848}{0.1607} = -11.1076$ (p-value < 0.0001)
- **Global test**

$$T_{global} = \begin{pmatrix} \hat{\beta} & \hat{\gamma} \end{pmatrix} \hat{\Sigma}^{-1} \begin{pmatrix} \hat{\beta} \\ \hat{\gamma} \end{pmatrix} = 124.1278 \text{ (pvalue} < 0.0001)$$

- **Sequential test**
  1. Start with $\tilde{\gamma} = -11.1076$
  2. We reject H_1 since $P(|\max(W_1, W_2)| > |\tilde{\gamma}|) < 0.0001$
  3. We cannot reject H_2 since $P(|W_2| > |\tilde{\beta}|) = 0.5080$

The cautionary tale revisited

- By using the bivariate pseudo-observation model we can disentangle the effects of treatment on recurrent events and death
- We do not make a type I error like we do if blindly fit a LWYY model to the data
- Moreover, we gain knowledge on how these effects are correlated. The correlation matrix, $\hat{\Omega}$, corresponding to $\hat{\Sigma}$ is given by

$$\hat{\Omega} = \begin{pmatrix} 1 & -0.1367 \\ -0.1367 & 1 \end{pmatrix}$$

This matches the pattern we see, the negative effect on death for $Z = 0$ is carried over to the effect on recurrent events (although $\beta = 0$) through ρ

Summary

- The bivariate pseudo-observation model is a powerful tool for understanding treatment effects (or other covariate effects) on recurrent events and death simultaneously
- Various link functions may be chosen in the linear model (corresponding to different estimands)
- This approach actually analyses the two-dimensional (or multi-dimensional) nature of data instead of trying to squeeze it into an one-dimensional issue
- `recurrentpseudo` R-package on the way...
- Think twice before you close your eyes to mortality when conducting a recurrent event analysis!