Dynamic frailty models for recurrent events data

Theodor Adrian Balan   Hein Putter

Department of Medical Statistics and Bioinformatics

Leiden University Medical Center

PTA International Conference, 2016

The data for one subject...

- Events at $t_1 < t_2 < ... < t_n$ over $[0, \tau]$
- Possibly time-dependent covariates $\mathbf{x}_i(t)$
- Unobserved random effect Z_i (*frailty*)

The data for one subject...

- Events at $t_1 < t_2 < ... < t_n$ over $[0, \tau]$
- Possibly time-dependent covariates $\mathbf{x}_i(t)$
- Unobserved random effect Z_i (*frailty*)

The canonical framework is to assume that the process is conditionally Poisson, i.e.

$$\lambda_i(t|Z_i) = Z_i \lambda_0(t) e^{\gamma' \mathbf{x}_i(t)}$$

Interpretation

- Assume that the process is Poisson **conditional** on Z_i (relaxation of the Poisson assumption)
- The regular frailty stands for "unobserved heterogeneity" due to missing covariates that are **time-constant** and have a **time-constant effect**
- Assumptions of the frailty model are rarely checked

Interpretation

- Assume that the process is Poisson **conditional** on Z_i (relaxation of the Poisson assumption)
- The regular frailty stands for "unobserved heterogeneity" due to missing covariates that are **time-constant** and have a **time-constant effect**
- Assumptions of the frailty model are rarely checked

A motivational quote

A criticism of commonly used random effects models, though, is that the random effects are time- invariant and are too simplistic for complex processes.
– Cook & Lawless, *The Statistical Analysis of Recurrent Events* (2007)

Dynamic frailty extension

Idea

An unobserved **time-dependent** random effect $Z_i(t)$ and that the process is conditionally Poisson, i.e.

$$\lambda_i(t|Z_i(\cdot)) = Z_i(t)\lambda_0(t)e^{\gamma' \mathbf{x}_i(t)}$$

Interpretation

- The dynamic frailty stands for "unobserved heterogeneity" due to missing covariates, possibly **time-dependent** and with a possibly **time-dependent effect**
- Previous work done by Yashin & Manton (1997), Gjessing et al (2003), Aalen & Gjessing (2004), Putter & van Houwelingen (2015)

The dynamic frailty construction

General idea is to have two parts:

- A part that models the distribution of $Z_i(t)$ for a fixed t for some parameter θ
- A part that models the correlation between $Z_i(t_1)$ and $Z_i(t_2)$ for some parameter λ

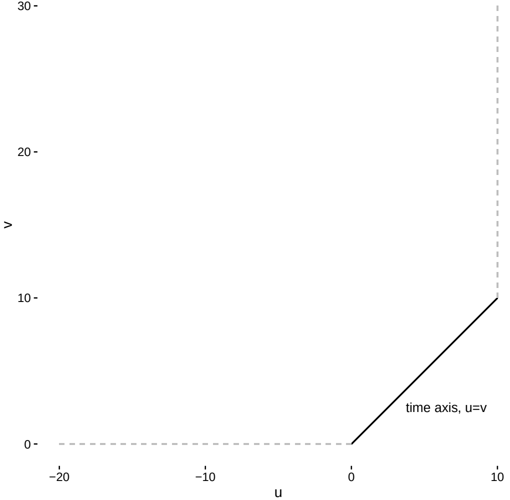
The dynamic frailty construction

General idea is to have two parts:

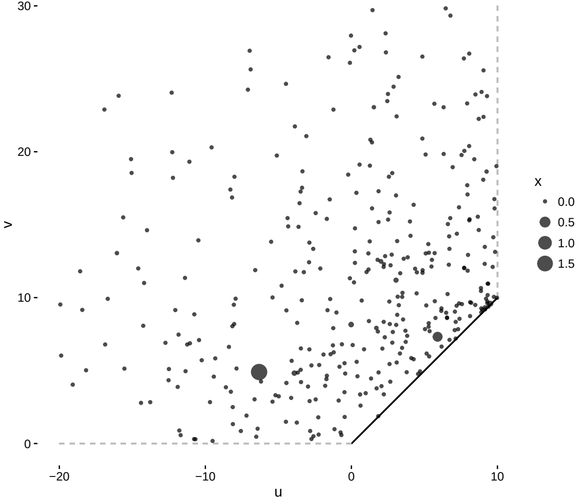
- A part that models the distribution of $Z_i(t)$ for a fixed t for some parameter θ
- A part that models the correlation between $Z_i(t_1)$ and $Z_i(t_2)$ for some parameter λ

The construction used is that of Putter & van Houwelingen, *Dynamic frailty models based on compound birth-death processes*, Biostatistics (2015).

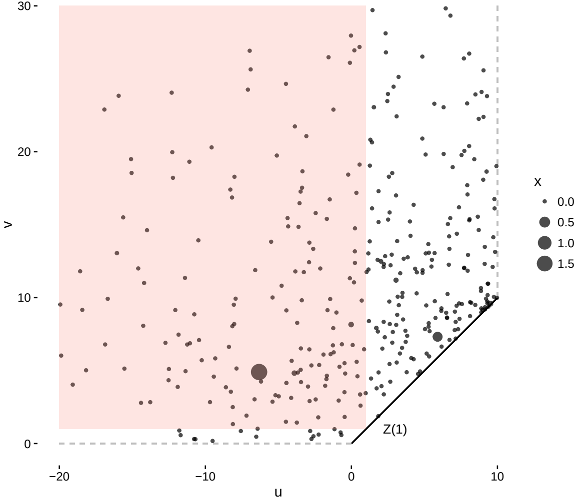
The dynamic frailty construction



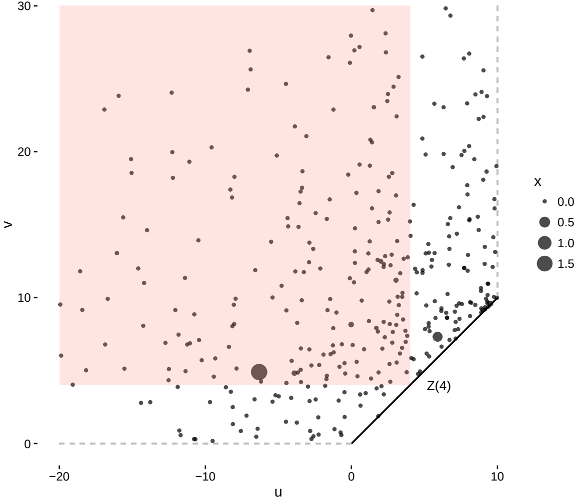
The dynamic frailty construction



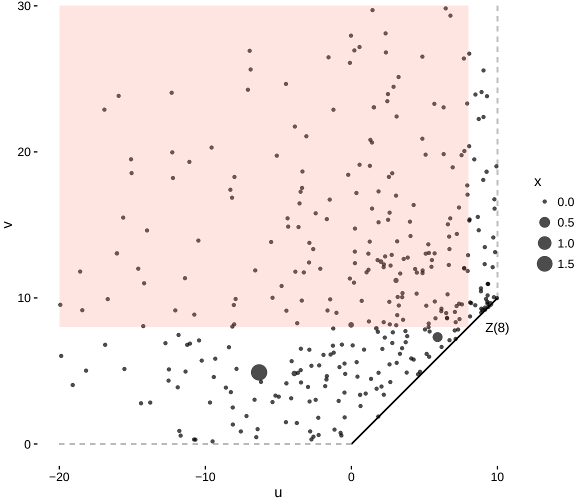
The dynamic frailty construction



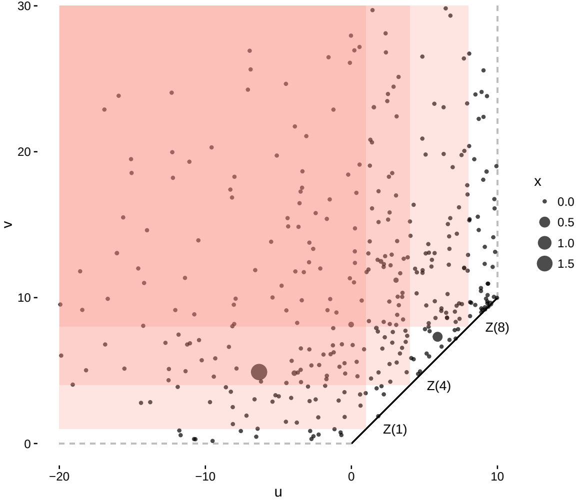
The dynamic frailty construction



The dynamic frailty construction

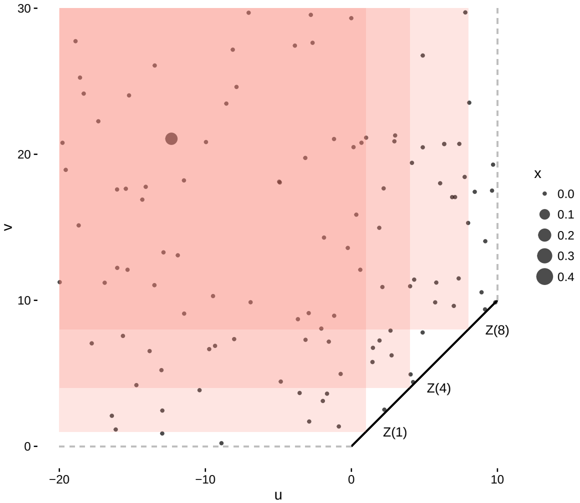


The dynamic frailty construction



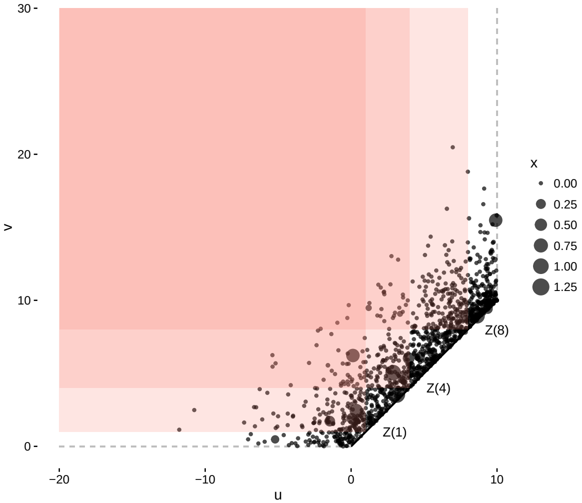
The dynamic frailty construction

High correlation



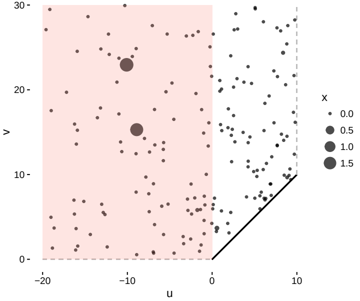
The dynamic frailty construction

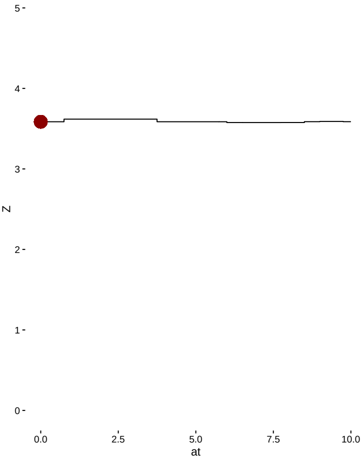
Low correlation



The dynamic frailty construction

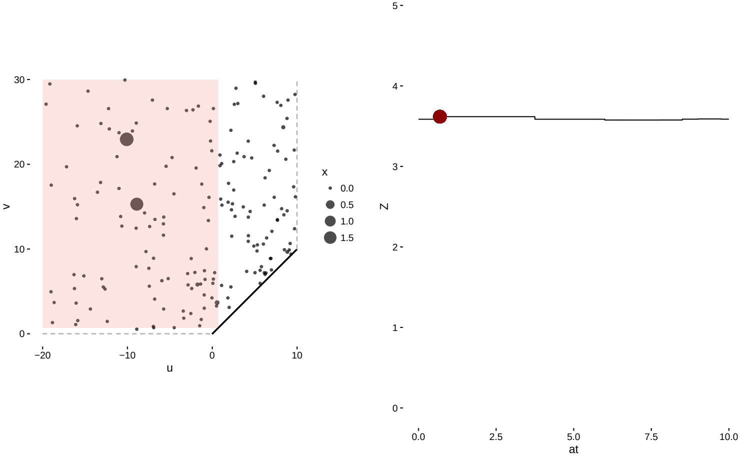
High correlation





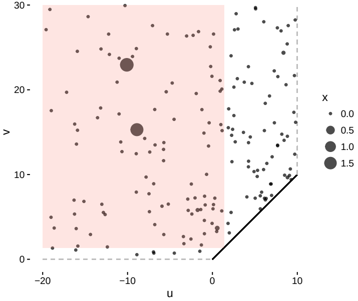
The dynamic frailty construction

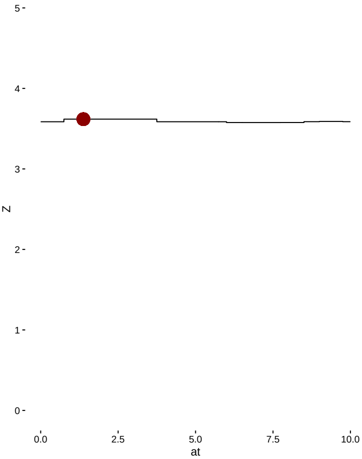
High correlation



The dynamic frailty construction

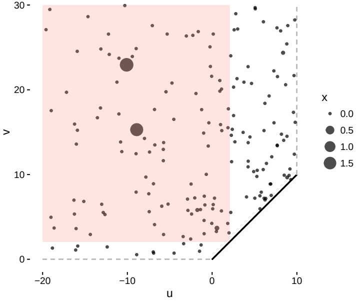
High correlation

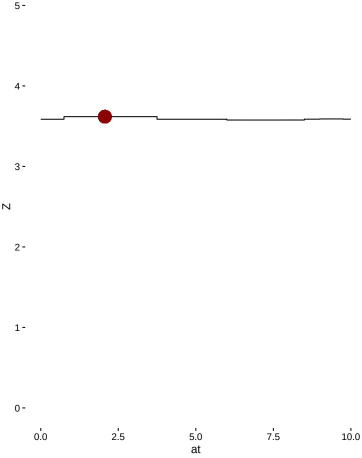




The dynamic frailty construction

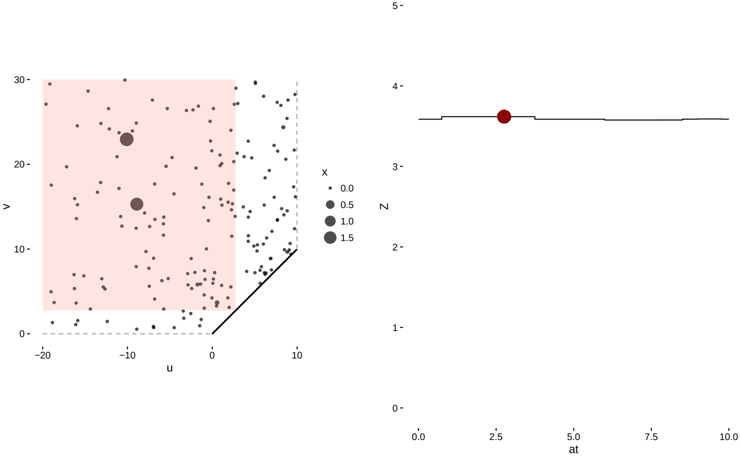
High correlation





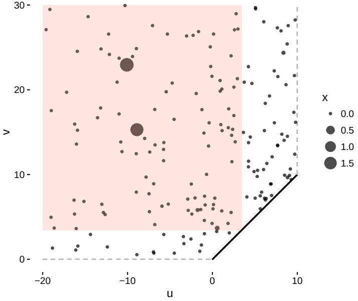
The dynamic frailty construction

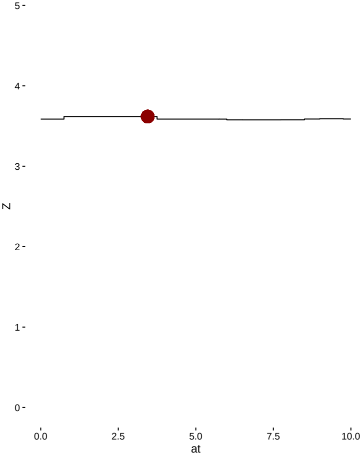
High correlation



The dynamic frailty construction

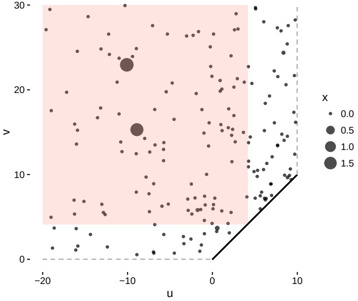
High correlation

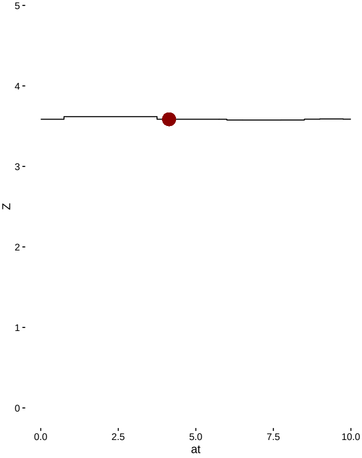




The dynamic frailty construction

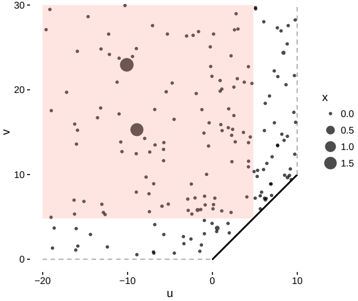
High correlation

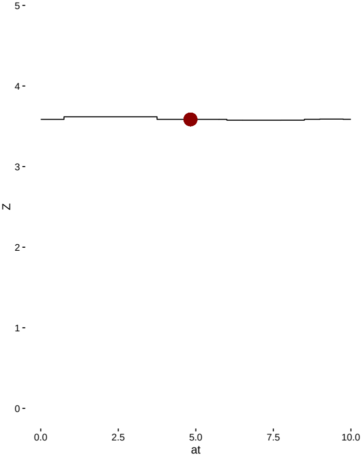




The dynamic frailty construction

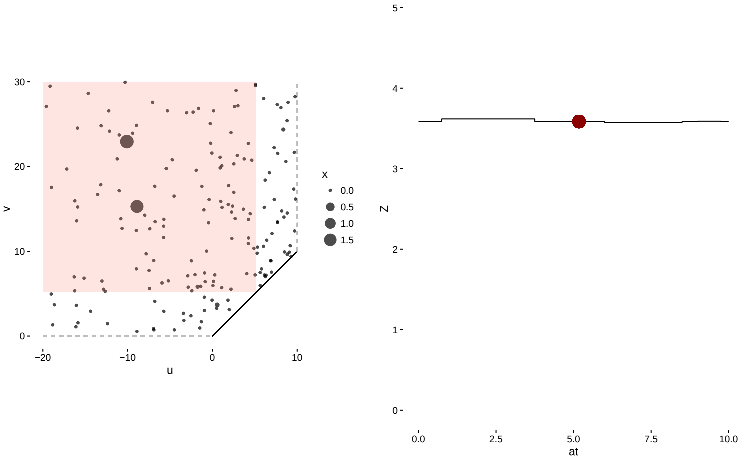
High correlation





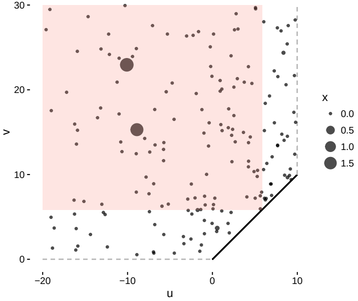
The dynamic frailty construction

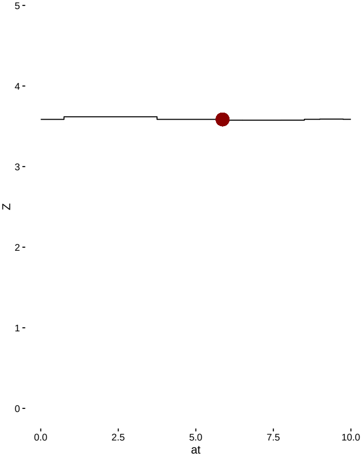
High correlation



The dynamic frailty construction

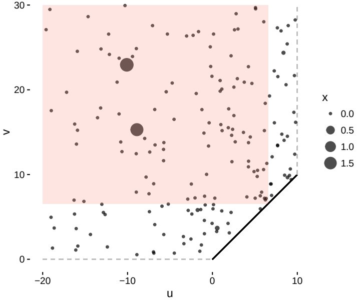
High correlation

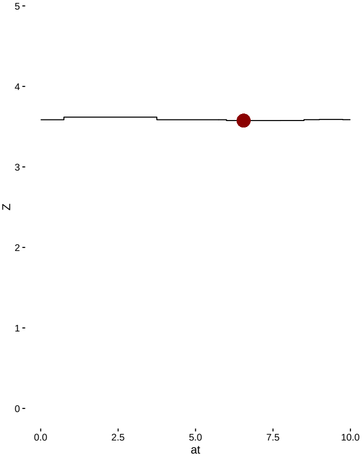




The dynamic frailty construction

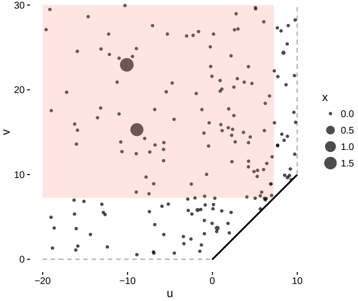
High correlation

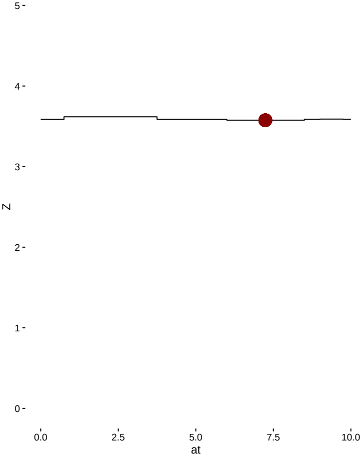




The dynamic frailty construction

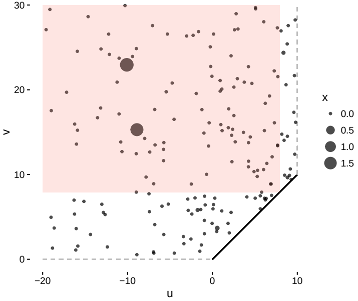
High correlation

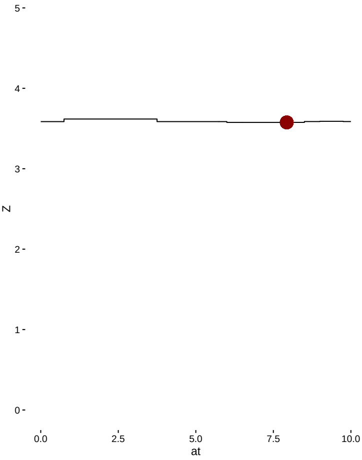




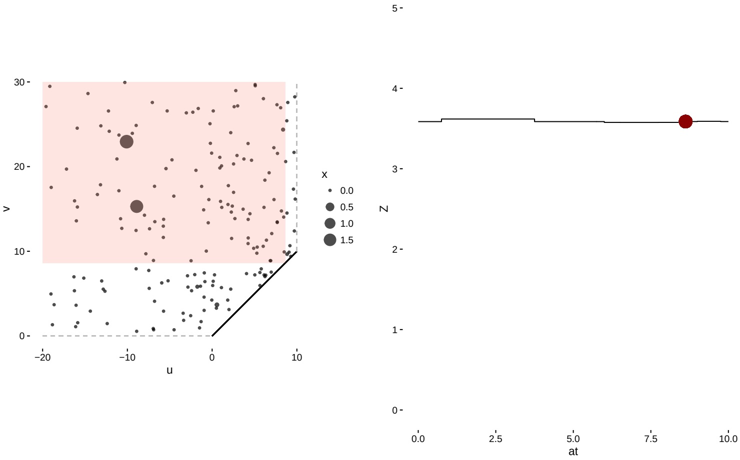
The dynamic frailty construction

High correlation



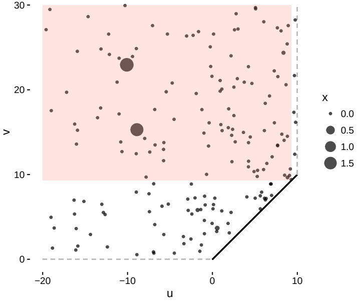


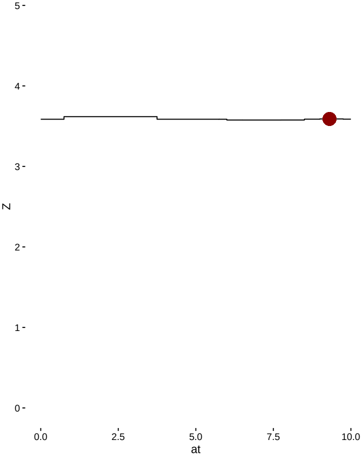
The dynamic frailty construction
High correlation



The dynamic frailty construction

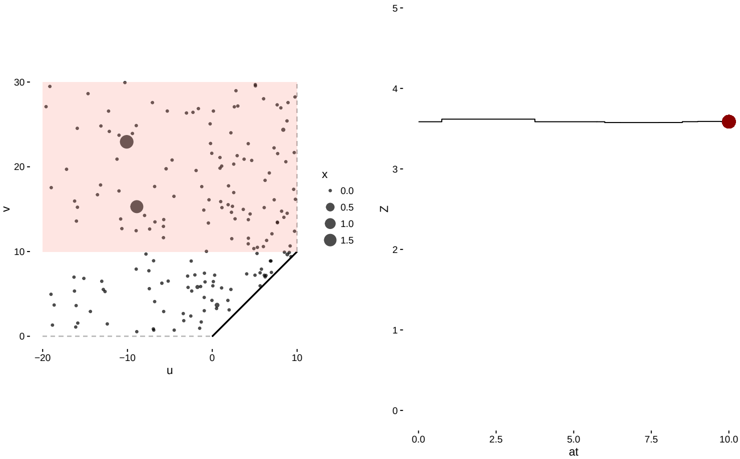
High correlation





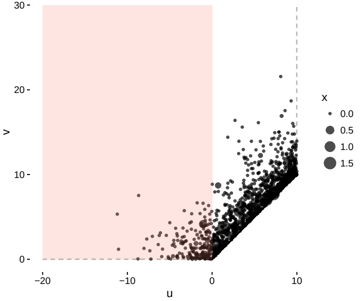
The dynamic frailty construction

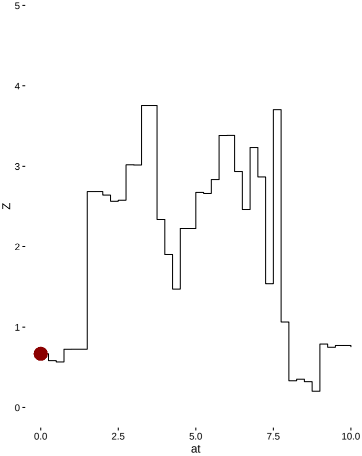
High correlation



The dynamic frailty construction

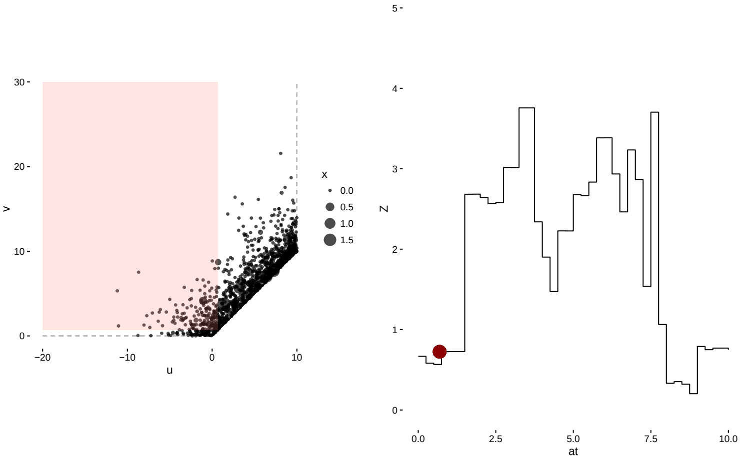
Low correlation





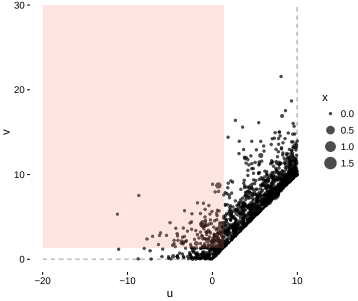
The dynamic frailty construction

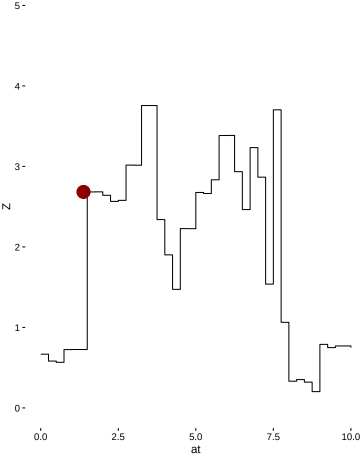
Low correlation



The dynamic frailty construction

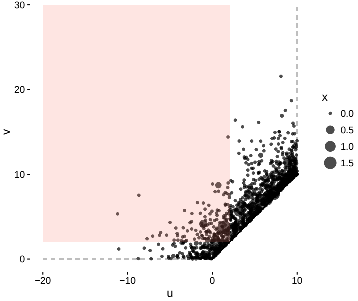
Low correlation

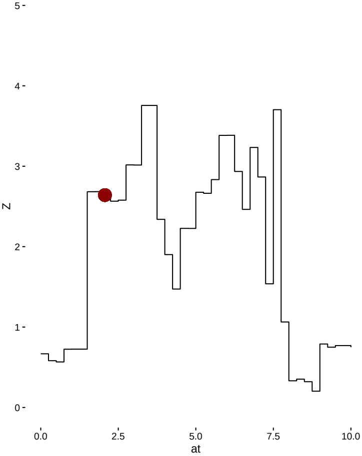




The dynamic frailty construction

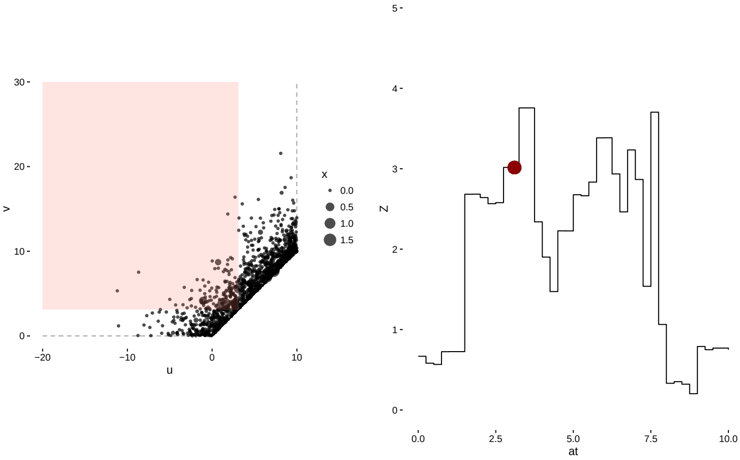
Low correlation





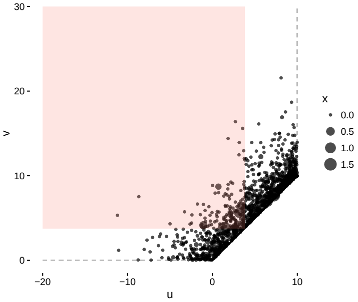
The dynamic frailty construction

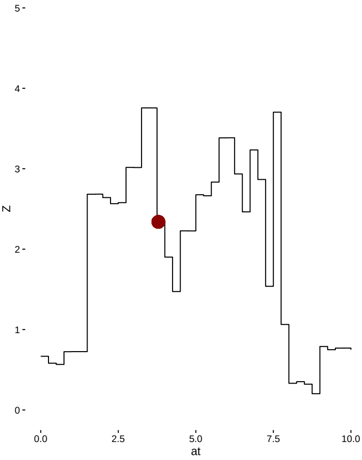
Low correlation



The dynamic frailty construction

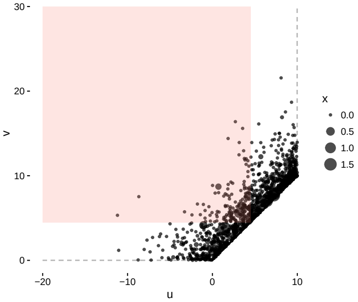
Low correlation

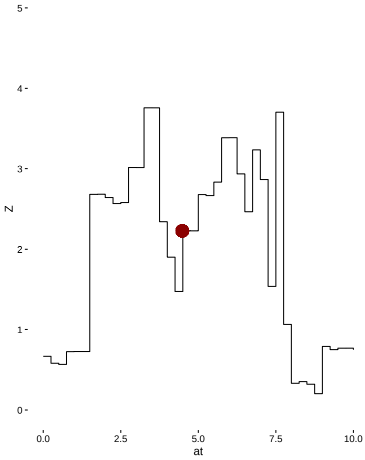




The dynamic frailty construction

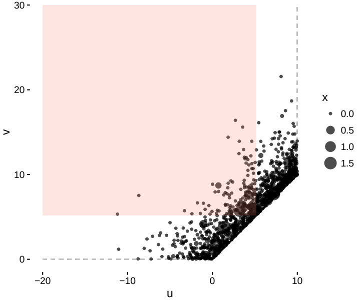
Low correlation

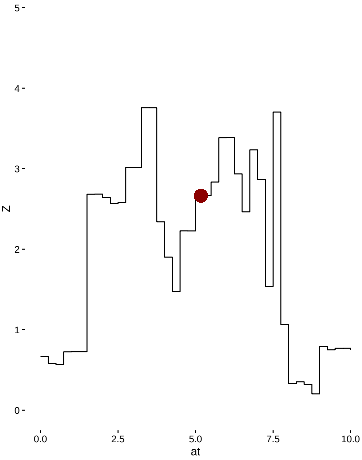




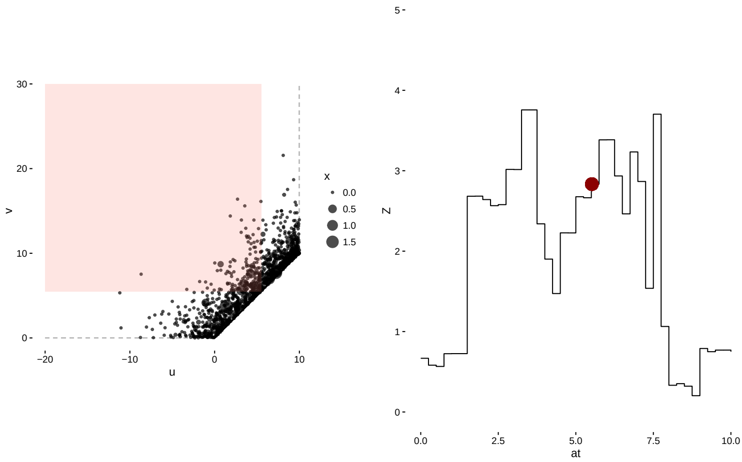
The dynamic frailty construction

Low correlation



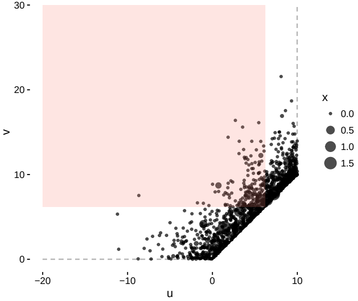


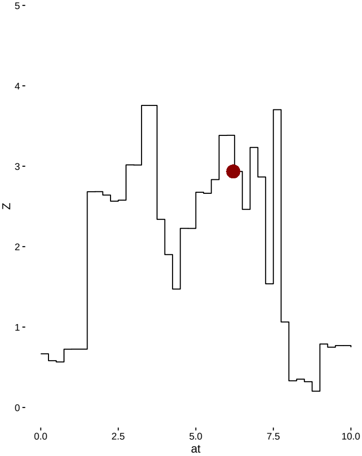
The dynamic frailty construction
Low correlation



The dynamic frailty construction

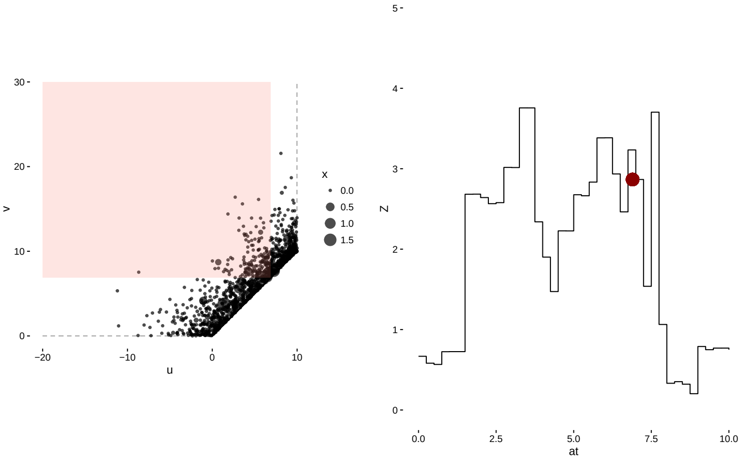
Low correlation





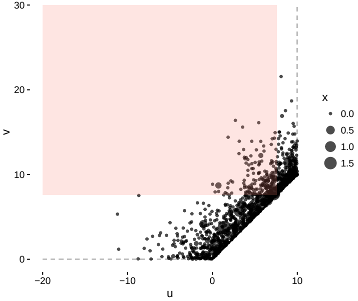
The dynamic frailty construction

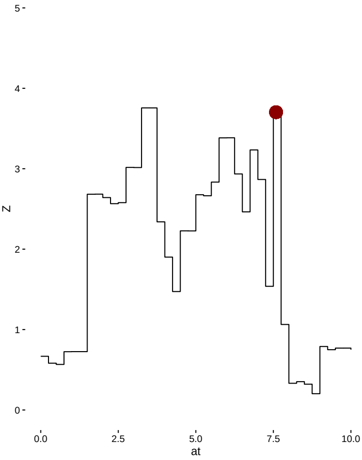
Low correlation



The dynamic frailty construction

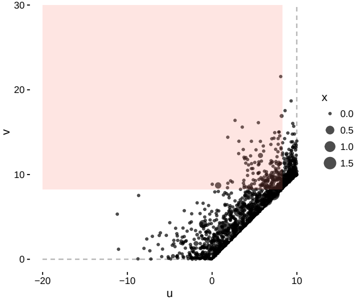
Low correlation

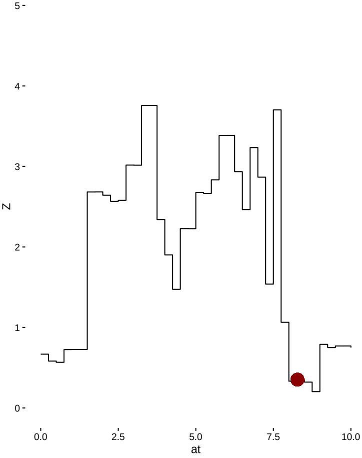




The dynamic frailty construction

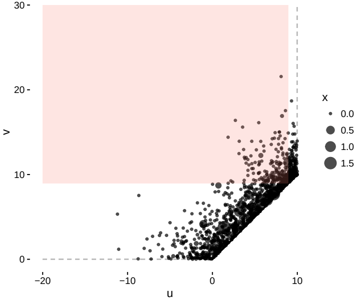
Low correlation

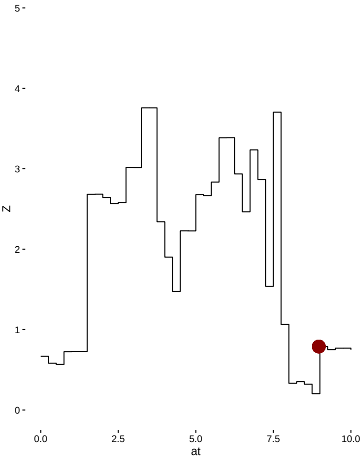




The dynamic frailty construction

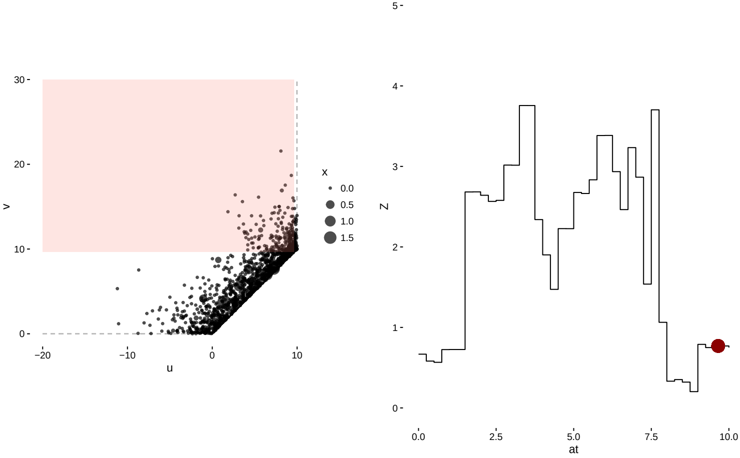
Low correlation



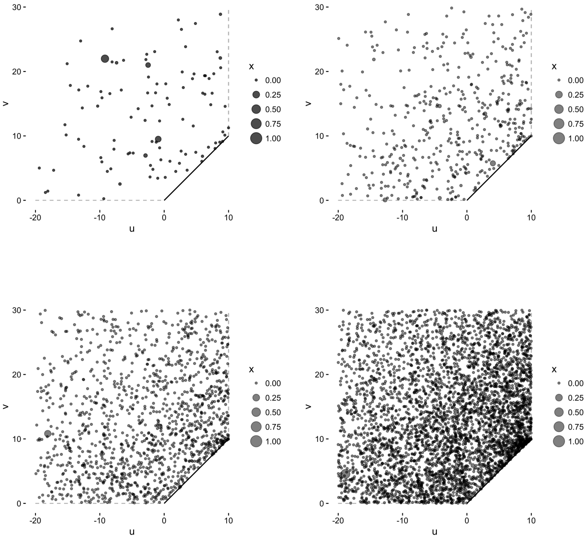


The dynamic frailty construction

Low correlation



Limiting process



Choices for the frailty

The distribution

The key property of the distribution is “infinite divisbility”, positive and continuous density $f_\theta(x)$

- the gamma distribution
- the positve stable distribution
- the compound Poisson distribution
- all members of the Power Variance Functions (PVF) family (Hougaard, 1986)

The correlation

The correlation is controlled by the intensity of the Poisson process. The key is to have an intensity g for the locations (u, v), i.e.

$$g(u, v; \lambda) \propto e^{-\lambda(v-u)}$$

General idea of estimation

From Nielsen et al (1992): for a grid of θ, maximize the log-likelihood w.r.t γ and h_0 with an EM algorithm.

- E step: calculate $\widehat{Z}_i = E[Z_i | data]$
- M step: Cox model with $\log \widehat{Z}_i$ as time-constant offset

From Putter & van Houwelingen (2015): for a grid of (λ, θ), maximize the log-likelihood w.r.t γ and h_0 with an EM algorithm.

- E step: calculate $\widehat{Z}_i(t) = E[Z_i(t) | data]$ at all the event time points in the data
- M step: Cox model with $\log \widehat{Z}_i(t)$ as time-dependent offset

General idea of estimation

From Nielsen et al (1992): for a grid of θ, maximize the log-likelihood w.r.t γ and h_0 with an EM algorithm.

- E step: calculate $\widehat{Z}_i = E[Z_i | data]$
- M step: Cox model with $\log \widehat{Z}_i$ as time-constant offset

From Putter & van Houwelingen (2015): for a grid of (λ, θ), maximize the log-likelihood w.r.t γ and h_0 with an EM algorithm.

- E step: calculate $\widehat{Z}_i(t) = E[Z_i(t) | data]$ at all the event time points in the data
- M step: Cox model with $\log \widehat{Z}_i(t)$ as time-dependent offset

General idea of estimation

From Nielsen et al (1992): for a grid of θ, maximize the log-likelihood w.r.t γ and h_0 with an EM algorithm.

- E step: calculate $\widehat{Z}_i = E[Z_i | data]$
- M step: Cox model with $\log \widehat{Z}_i$ as time-constant offset

From Putter & van Houwelingen (2015): for a grid of (λ, θ), maximize the log-likelihood w.r.t γ and h_0 with an EM algorithm.

- E step: calculate $\widehat{Z}_i(t) = E[Z_i(t) | data]$ at all the event time points in the data
- M step: Cox model with $\log \widehat{Z}_i(t)$ as time-dependent offset

General idea of estimation

From Nielsen et al (1992): for a grid of θ, maximize the log-likelihood w.r.t γ and h_0 with **EM algorithm**

- E step: calculate $\widehat{Z}_i = E[Z_i | data]$
- M step: Cox model with $\log \widehat{Z}_i$ as time-constant offset

From Putter & van Houwelingen (2015): for a grid of (λ, θ), maximize the log-likelihood w.r.t γ and h_0 with an **EM algorithm**.

- E step: calculate $\widehat{Z}_i(t) = E[Z_i(t) | data]$ at all the event time points in the data
- M step: Cox model with $\log \widehat{Z}_i(t)$ as time-dependent offset

General idea of estimation

From Nielsen et al (1992): for a grid of θ, maximize the log-likelihood w.r.t γ and h_0 with **EM algorithm**

- E step: calculate $\widehat{Z}_i = E[Z_i | data]$
- M step: Cox model with $\log \widehat{Z}_i$ as time-constant offset

From Putter & van Houwelingen (2015): for a grid of (λ, θ), maximize the log-likelihood w.r.t γ and h_0 with an **EM algorithm**.

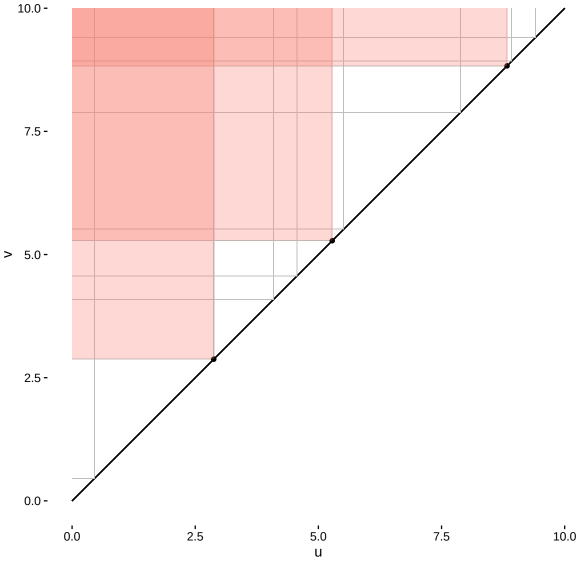
- E step: calculate $\widehat{Z}_i(t) = E[Z_i(t) | data]$ at all the event time points in the data
- M step: Cox model with $\log \widehat{Z}_i(t)$ as time-dependent offset

Costly E step

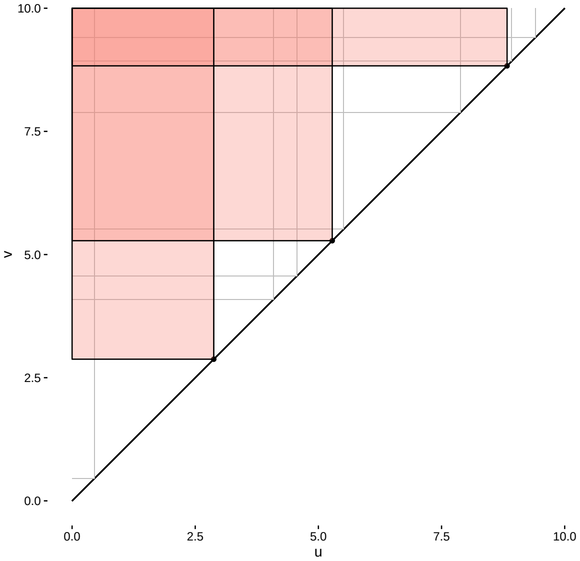
The E step is slow because of two factors:

- Large number of events / individual
- Large number of event time points in the data set

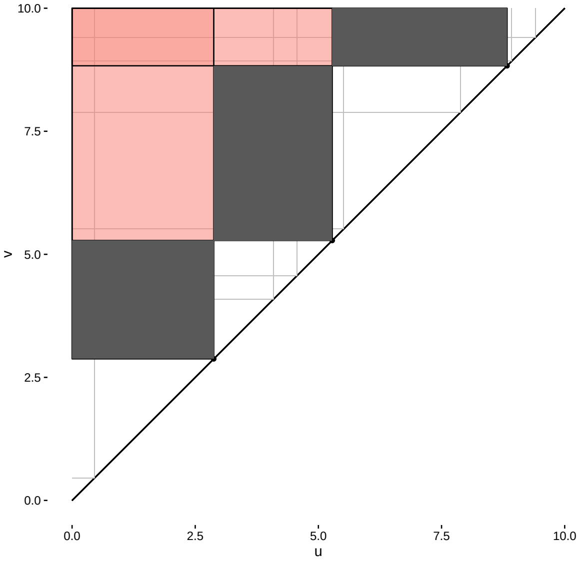
E step



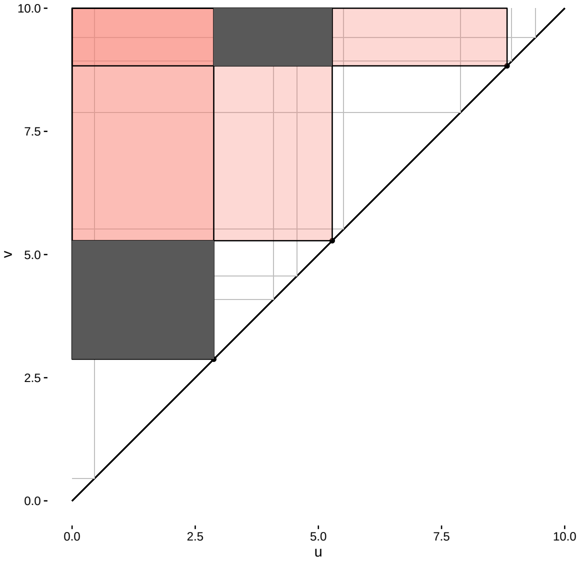
E step



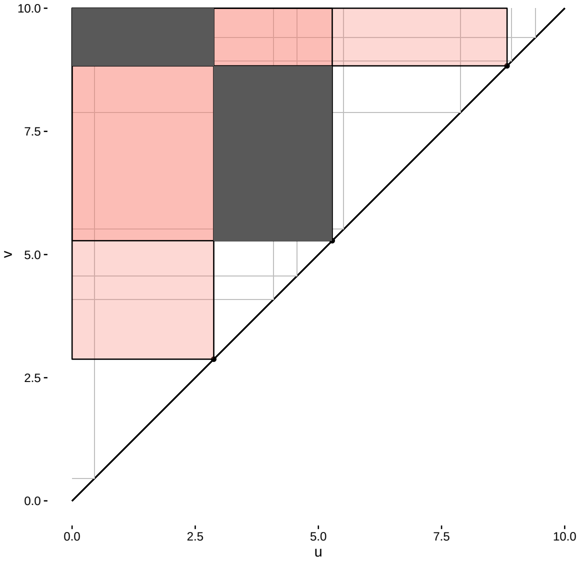
E step



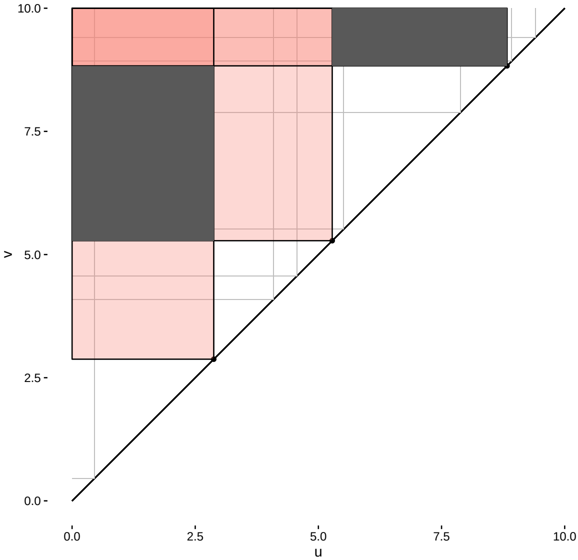
E step



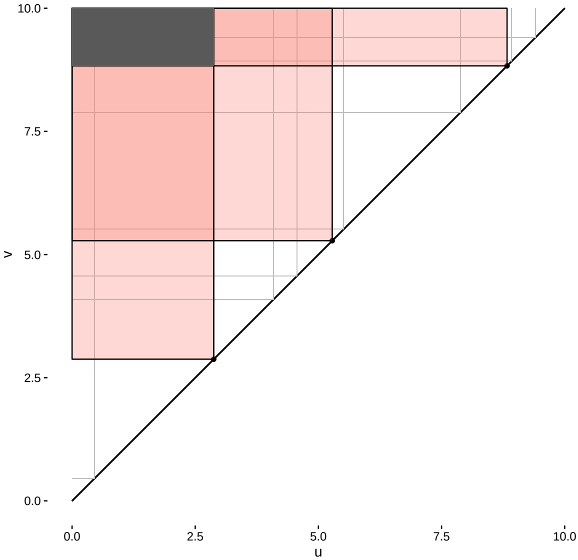
E step



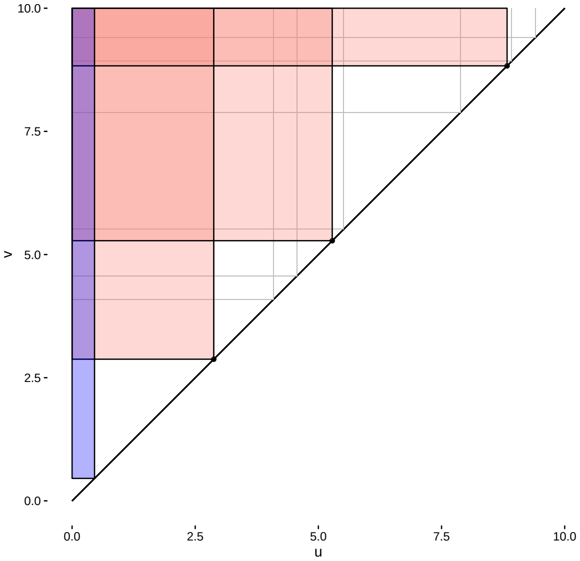
E step



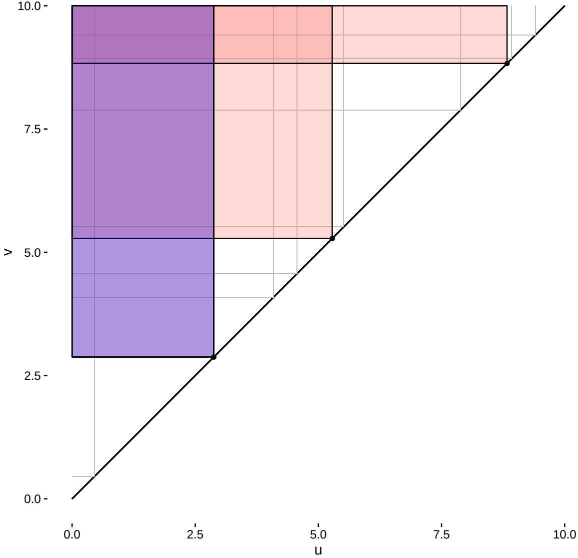
E step



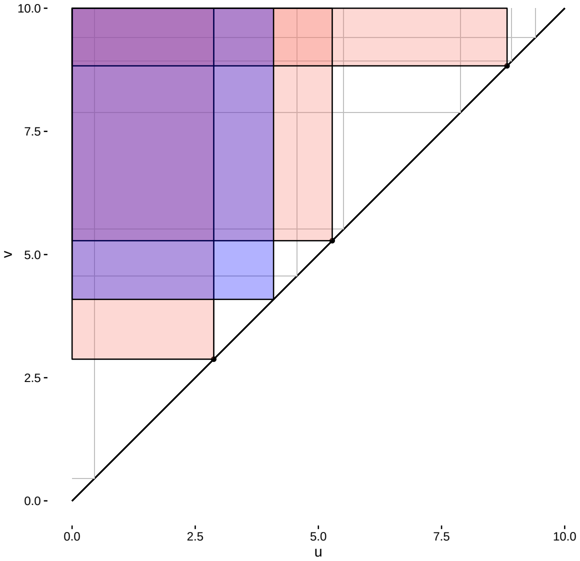
E step



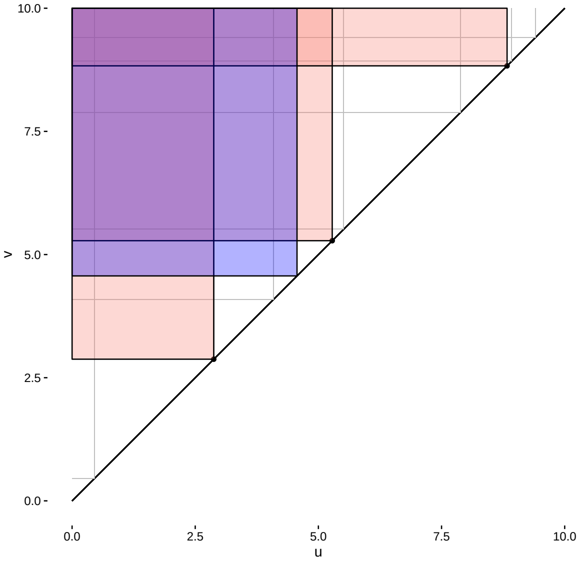
E step



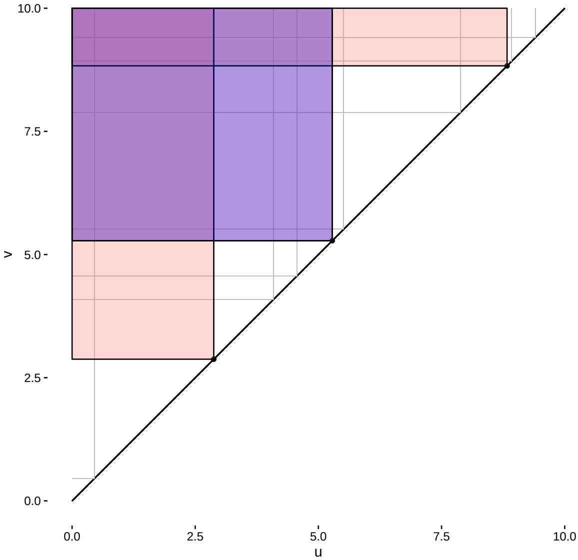
E step



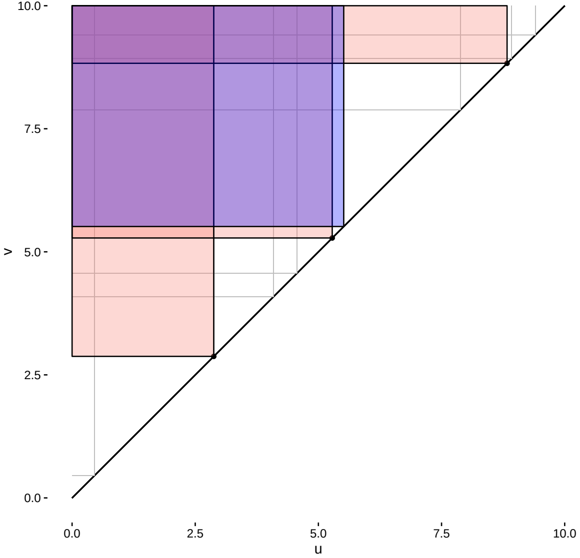
E step



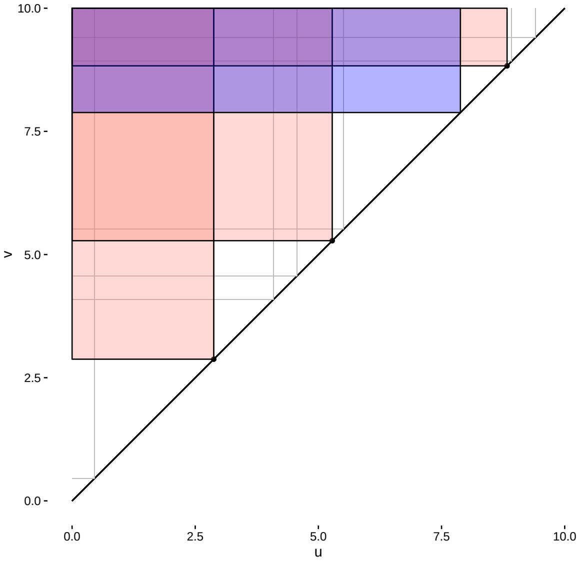
E step



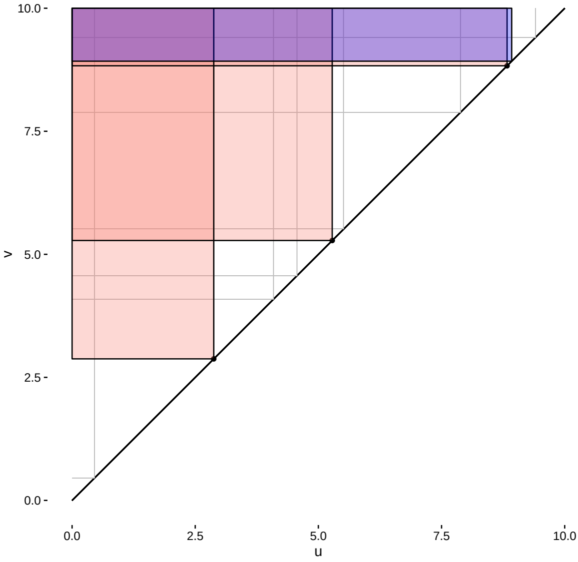
E step



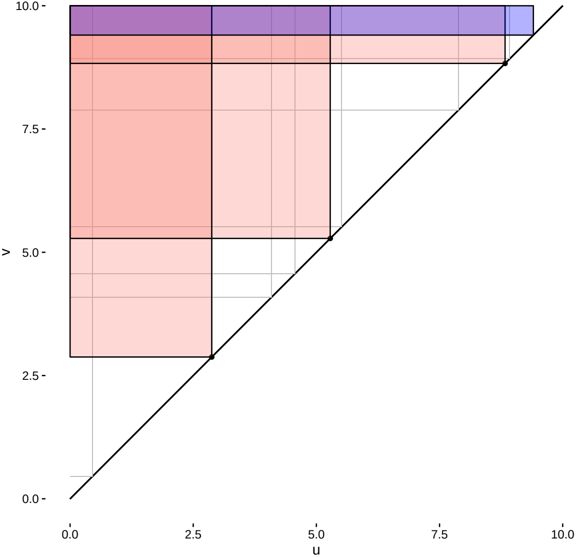
E step



E step



E step



Costly E step

The E step is slow because of two factors:

- Large number of events / individual
- Large number of event time points in the data set

How to make it faster?

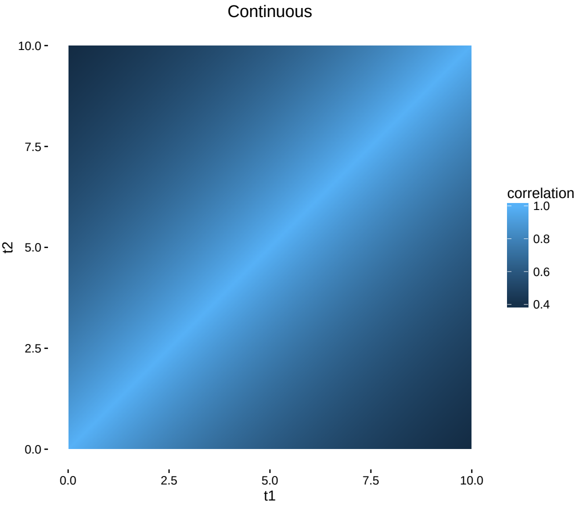
Costly E step

The E step is slow because of two factors:

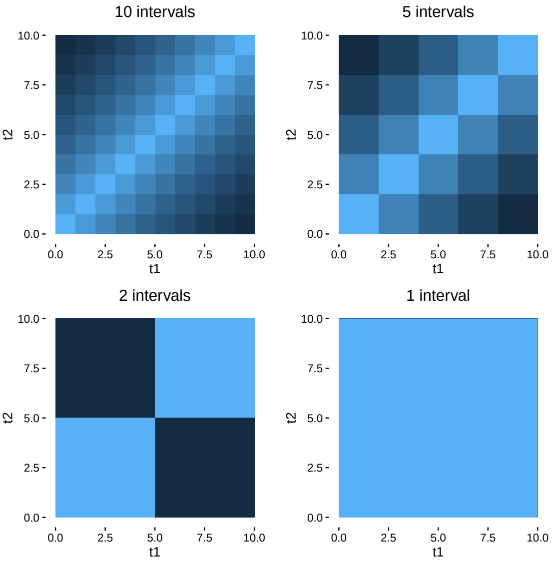
- Large number of events / individual
- Large number of event time points in the data set

How to make it faster?
Our proposal: piecewise-constant frailty, $Z(t)$ constant on pre-defined intervals

Costly E step

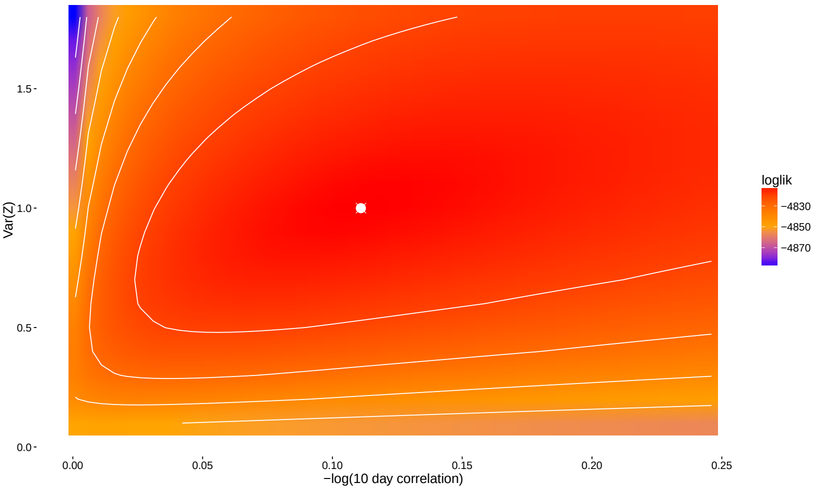


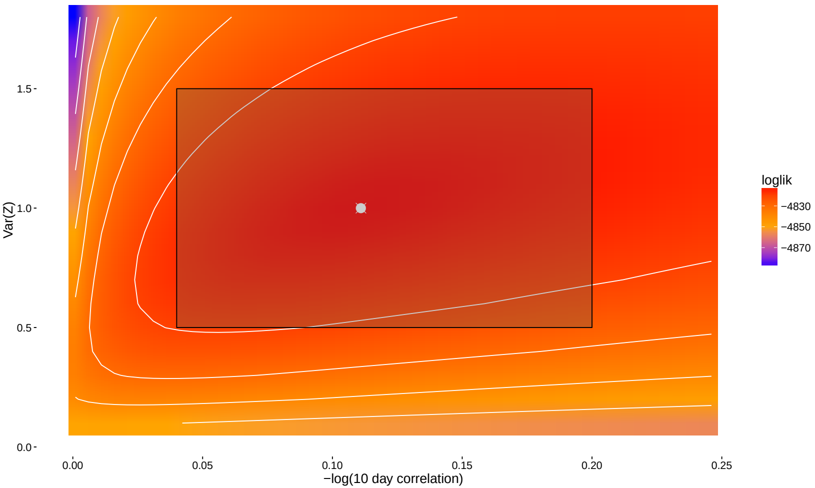
Costly E step



- Dataset featured in Duchateau & Janssen - *The frailty model* (2008)
- 232 individuals randomized to placebo or drug
- average 6.9 events / individual, total of 561 distinct event time points
- administrative censoring at 600 days after randomization

- Artificially censored all individuals after 5 events
- Fit the dynamic frailty model with gamma distribution, piecewise-constant frailty with 20 intervals
- For each combination of frailty parameters (λ, θ) employ an EM algorithm


loglik
−4830
−4850
−4870
Var(Z)
1.5
1.0
0.5
0.0
0.00
0.05
0.10
0.15
0.20
0.25
−log(10 day correlation)



loglik
−4830
−4850
−4870
Var(Z)
1.5
1.0
0.5
0.0
0.00
0.05
0.10
0.15
0.20
0.25
−log(10 day correlation)

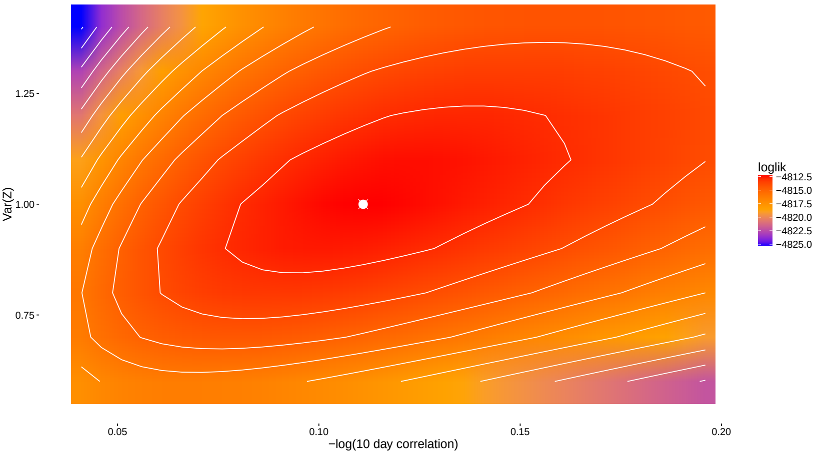
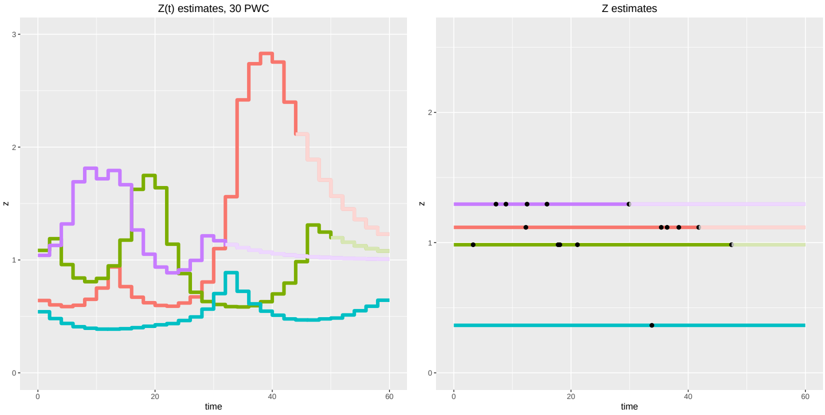

loglik
−4812.5
−4815.0
−4817.5
−4820.0
−4822.5
−4825.0
Var(Z)
1.25
1.00
0.75
0.05
0.10
0.15
0.20
−log(10 day correlation)

Table: Some results

dynamic	
frailty variance θ	0.98
10-day correlation	0.89
30-day correlation	0.71
50-day correlation	0.56
100-day correlation	0.33
$\beta_{\mathrm{treatment}}$	-0.16 (0.022)
static	
frailty variance θ	0.33
x-day correlation	1
$\beta_{\mathrm{treatment}}$	-0.15 (0.067)

Empirical Bayes frailty estimates



Summary

- Piecewise-constant model "in between" dynamic frailty & static frailty
- Lots of flexibility in terms of distribution & correlation structure
- Works well for medium-sized data sets
- Many potential uses; e.g. prediction
- Still working on:
  - Score test for the "dynamic" character of the frailty
  - Monte-Carlo E step

Selection of frailty intervals