Analysis Example and R Codes

This tutorial provides brief information on the implementation of the estimation procedure for the Bayesian joint model of trivariate processes using sample datasets for 2000 subjects. The sample datasets (`dat.long.rdata` – data on repeated measurements, `dat.terminal.rdata` – data on terminal event, and `dat.recurrent.rdata` – data on recurrent events) and the following R codes for fitting the model are publicly available at https://github.com/esrakurum/Bayes-Trivariate:

- `prepare_data.R` – prepares the data to be fit via the joint models,
- `jags.R` – includes the R code for joint modeling via JAGS,
- `Fit.R` – includes the main code; i.e., loading packages and the data, running the code to prepare the data, running the joint model in JAGS, and saving the posterior samples.

In order to run the Bayesian joint model of trivariate processes,

1. Download all the files and place them in one folder,
2. Set this folder as the working directory using the `setwd(...)` specified in the `Fit.R`,
3. Run the code in `Fit.R`.

The following function within `Fit.R` file fits the joint models using the Bayesian estimation method presented in the main paper.

```
BTJM_fit(dat.long, dat.terminal, dat.recurrent, ncX, ncZ, n.iter, n.burn, n.thin,
n.chains, parallel)
```

The function `BTJM_fit` requires a separate data frame corresponding to each component of the joint model: longitudinal (`dat.long`), terminal event (`dat.terminal`), and recurrent events (`dat.recurrent`). The number of primary and secondary covariates are passed on via the arguments `ncZ` and `ncX`, respectively. The arguments `n.iter`, `n.burn`, and `n.thin` are the total number of iterations (including burn-in period), number of iterations during burn-in, and number of iterations for thinning, respectively. The last argument `parallel` indicates whether the chains should be run in parallel; defaults to `TRUE`. The function returns the posterior samples for all parameters defined in the trivariate model (equation on page 4 of the main manuscript), which is stored in `post.samples` data frame and obtained from the number of chains indicated using the `n.chains` argument. Note that the notation used to store the parameters follows the notation in the main manuscript.

Detailed information on data structures are illustrated in each sample data:

```
> load("dat.long.rdata")
> load("dat.terminal.rdata")
> load("dat.recurrent.rdata")
> names(dat.long)
[1] "subj.id" "t" "Y" "x1" "z1"
> names(dat.terminal)
[1] "subj.id" "T.terminal" "event.t" "x1" "z1"
> names(dat.recurrent)
[1] "subj.id" "T.gap" "event.r" "r.obs.id" "x1" "z1",
```

where

- `subj.id` – subject id
- `t` – longitudinal measurement time

- `Y` – longitudinal outcome
- `x1` and `z1` – secondary covariate (risk factors) and primary covariate (primary "exposure" variables or main risk factors), respectively
- `T.terminal` and `event.t` – terminal event time and terminal event indicator, respectively
- `T.gap` and `event.r` – recurrent event gap time and recurrent event indicator, respectively, at the `r.obs.id`th recurrent event for subject i, where `r.obs.id` is the index for recurrent event time ($1, 2, \ldots, r_i, r_i + 1$; r_i: number of recurrent events for the ith subject).

Note that

- the `dat.terminal.rdata` only stores information on terminal event time and terminal event indicator; therefore, in this data set, there should be one row per subject.
- in the `dat.recurrent.rdata`, `T.gap` is the time between two consecutive recurrent event times. For subjects with at least one recurrent event, `T.gap` value in their first row is the time to their first recurrent event and `T.gap` value in their last row is the time between terminal event time (`T.terminal`) and the last recurrent event time with `event.r` recorded as zero for this row. For subjects with no recurrent events, `T.gap` value is equal to their terminal event time (`T.terminal`) and `event.r` is recorded as zero. More specifically, for each subject, there should be $r_i + 1$ rows (r_i: number of recurrent events for the ith subject).