

Survival and longitudinal data analysis

Exercice 1

Time-dependent covariates: `bmt` data

Le jeu de données `bmt` (package `KMsurv`) contient des données pour 137 patients qui ont subi une greffe de moelle osseuse (https://fr.wikipedia.org/wiki/Transplantation_de_moelle).

Le but est d'expliquer le temps, noté DFS (disease free survival), jusqu'à la rechute ou la mort (variable `t2` avec indicatrice `d3`). Pour cela, nous allons considérer 10 variables statiques (variables `z1` à `z10`) et une variable dépendant du temps `platelet_recov` qui sera codée à partir des variables `tp` et `dp`.

1. Codage et transformation des variables statiques.

- Renommer les variables `z1` à `z10` en `agep, aged, genderp, genderd, cmvp, cmvd, waiting, FAB, hospital, MTW`, la variable `t2` en `DFS` et `d3` en `DFSstatus`.
- Renommer les valeurs prises par la variable `group` en `ALL, Low, High`
- Translater les âges `agep` et `aged` de -28 .
- Retirer les variables `t1, d1, d2, ta, da, tc, dc`. On notera `bm2` le jeu de données obtenu après ces transformations.

2. Variable dépendant du temps `platelet_recov`

- Pour les individus 1 et 14, donner les valeurs que prend cette variable en fonction du temps.
- Ecrire à la main un code qui permet de transformer les données `bm2` en données longitudinales (format `start/stop`) en calculant la variable `platelet_recov`
- Refaire de même en utilisant la fonction `tmerge` du package `survival`:

```
bmt2_merge <- tmerge(bmt2, bmt2, id=id, tstop=DFS)
bmt2_merge <- tmerge(bmt2_merge, bmt2, id=id, platcovery=tdc(tp))
#adds platelet recovery as time dependent covariate
```

3. Construction d'un modèle de Cox

- Construire un modèle de Cox
- Faire une sélection de variable
- Interpréter le modèle sélectionné.