Modeling the cumulative incidence function of clustered competing risk data: computational and numerical aspects of a multinomial GLMM approach

Henrique Ap. Laureano

Wagner H. Bonat

http://leg.ufpr.br/~henrique | http://leg.ufpr.br/~wagner

December/2020

Context: clustered competing risk data

Idea: causes competing by the occurence of an event such the

confiability analysis
failure of an industrial or
electronic component

survival analysis
failure or progress of a patient
or some biological process

Idea: causes competing by the occurence of an event such the

confiability analysis

failure of an industrial or
electronic component

survival analysis

failure or progress of a patient
or some biological process

A typical data set consists of

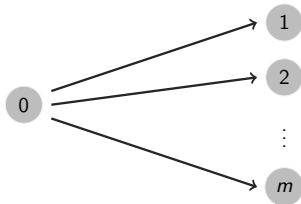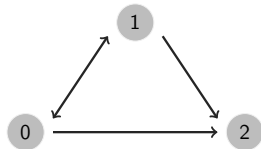| Group	ID	Cause 1	Cause 2	Censorship	Time	Feature
1	1	1	0	0	10	A
1	2	0	0	1	8	A
2	1	0	0	1	7	B
2	2	0	1	0	5	A

Survival data designs

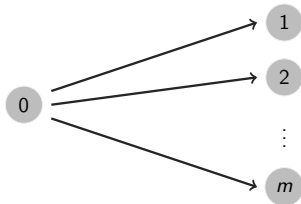Failure time process   Competing risk process   Multistate process
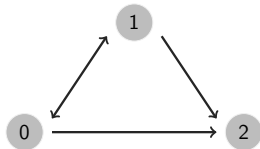
Survival data designs

Failure time process

Competing risk process

Multistate process

Survival modeling framework

What we do?

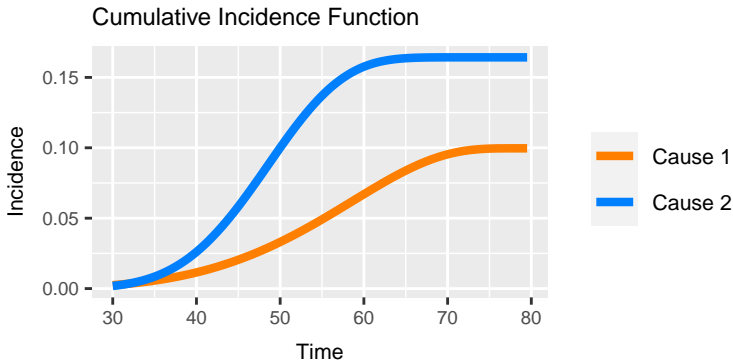We model the risk of each competing cause along the time and taking into account the possible within-group dependence

... all this in terms of a

We model the risk of each competing cause along the time and taking into account the possible within-group dependence

... all this in terms of a

Cumulative Incidence Function

Main focus application: cancer incidence in twins

Clustered competing risks data
↳ Clusters? Families
↳ Family studies
↳ Twins data

Main focus application: cancer incidence in twins

Clustered competing risks data
↳ Clusters? Families
↳Family studies
↳ Twins data

» Taking into account the within-family dependence may reflect both disease **heritability** and the impact of shared **environmental effects**
» A complication is that we have little information to track that **dependence** since each 'family' consists of only a pair of twins

Challenges

Besides the small size groups, the data is very simple . . .

- » we just know if the event occured (1 or 0) and the time
  - » with this, we have to be able to construct the
    cumulative incidence curves
- » and we have to accommodate the within-family dependency
  - » that can happen in different ways and with different intensities

---

to accomplish all this a powerful modeling framework is made necessary

. . . with this,
computational and numerical challenges has also to be overcome

Thank you

http://leg.ufpr.br/~henrique

@hap_laureano