Time to freezing of gait analysis : A time to event analysis

Contents

Executive summary

In this project, we will be using data from kaggle.We aim to analysis time to freezing of gait (FOG). FOG is a pattern occurring in patient with Parkinson disease. It indicates kinetic inability and impairment during gait for instance. Some indicative events like walking hesitation, turning body could be observed and help to detect FOG occurence.

Introduction

We aim at analyzing the time to event in Parkinsonian 44 patients, with median age of 69 years old. Here FOG is the event of interest and it is expected within the execution of a given task, or during daily life activity.We will analyse data from lab tasks. Our main goal is to analyse and explain overall FOG time to event model, the influence/importance of some features to the occurrence of freezing of gait(FOG) from the beginning to the end of the task. To achieve this, we model our data under kaplan-Meier model.

Data preparation, inspection and preprocessing.

let's join all metadata tables before diving into analysis.

- defog_metadata.csv Identifies each series in the tdcsfog dataset by a unique Subject, Visit, Test, Medication condition.
  - `Visit` Lab visits consist of a baseline assessment, two post-treatment assessments for different treatment stages, and one follow-up assessment.
  - `Test` Which of three test types was performed, with 3 the most challenging.
  - `Medication` Subjects may have been either off or on anti-parkinsonian medication during the recording.
- subjects.csv Metadata for each Subject in the study, including their `Age` and `Sex` as well as:
  - `Visit` Only available for subjects in the daily and defog datasets.
  - `YearsSinceDx` Years since Parkinson's diagnosis.
  - `UPDRSIIIOn/UPDRSIIIOff` Unified Parkinson's Disease Rating Scale score during on/off medication respectively. `NFOGQ` Self-report FoG questionnaire score. See: https://pubmed.ncbi.nlm.nih.gov/19660949/
- events.csv Metadata for each FoG event in all data series. The event times agree with the labels in the data series.
  - `Id` The data series the event occured in.
  - `Init` Time (s) the event began.
  - `Completion` Time (s) the event ended.
  - `Type` Whether StartHesitation, Turn, or Walking.
  - `Kinetic` Whether the event was kinetic (1) and involved movement, or akinetic (0) and static.
- tasks.csv Task metadata for series in the defog dataset. (Not relevant for the series in the tdcsfog or daily datasets.)
  - `Id` The data series where the task was measured.
  - `Begin` Time (s) the task began.
  - `End` Time (s) the task ended.
    Task One of seven tasks types in the DeFOG protocol, described on this page.

Remove **Visit** from **subjects** table , as we do not use it

Note: we are going to consider only defog condition subjects.

Data structure

Ad event and tasks duration columns features:

- `eventsDuration <- Completion - Init`
- `tasksDuration <- Begin - end`

Data analysis

Subjects

How many unique subjects are in this dataset ?

There are 44 distinct subjects in this study.

Kinetic / Events

How many trials has missing kinetic/status ?

```
## [1] 0.48
```

48% of trials has missing Kinetic(status).

filter for missing kinetic entries/cases.

```
##
##  new table dimension
```

```
## [1] 2232   19
```

```
## # A tibble: 3 x 19
##   Id          Init Compl~1 Type  Kinetic Begin   End Task  Subject Visit Medic~2
##   <chr>      <dbl>   <dbl> <chr>   <dbl> <dbl> <dbl> <chr> <chr>   <dbl> <chr>
## 1 02ea782681 1377.   1378. Turn        1 1371. 1393. Turn~ ae2d35      2 on
## 2 02ea782681 1377.   1378. Turn        1 1371. 1393. Turn~ ae2d35      2 on
## 3 02ea782681 1466.   1467. Turn        1 1461. 1472. Hots~ ae2d35      2 on
## # ... with 8 more variables: Age <dbl>, Sex <chr>, YearsSinceDx <dbl>,
## #   UPDRSIII_On <dbl>, UPDRSIII_Off <dbl>, NFOGQ <dbl>, eventsDuration <dbl>,
## #   tasksDuration <dbl>, and abbreviated variable names 1: Completion,
## #   2: Medication
```

What is the proportion of observed Kinetic events ?

69 % of events has been observed.

Note : Given that each event is indicative of FOG, we will just gather them and consider that they form one class (Kinetic / events).

Age

What is the median age ?
Age numerical summary.

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   28.00   61.00   69.00   67.44   73.00   82.00
```

At least half of subjects are 69 years old.

Sex

Sex numerical summary

```
## Sex
##    F    M
## 0.45 0.55
```

There is almost 55.1498127% of men in this cohort.

Tasks

What kind of tasks has been performed ?

```
##  [1] "TUG-DT"     "TUG-C"      "Turning-DT" "Hotspot2"   "Hotspot2-C"
##  [6] "Hotspot1"   "Turning-ST" "4MW"        "4MW-C"      "TUG-ST"
## [11] "Turning-C"  "Hotspot1-C" "MB10"       "MB11"       "MB13"
## [16] "MB12"
```

how many tasks had been performed ?

There had been 16 task performed.

Visit

How many rounds of visit did the patient had ?

```
## [1] 2
```

What is the proportion of patient within each number of visit round ?

```
## Visit
##    1    2
## 0.37 0.63
```

37% of patients had one Visit, while the remaining had 2.

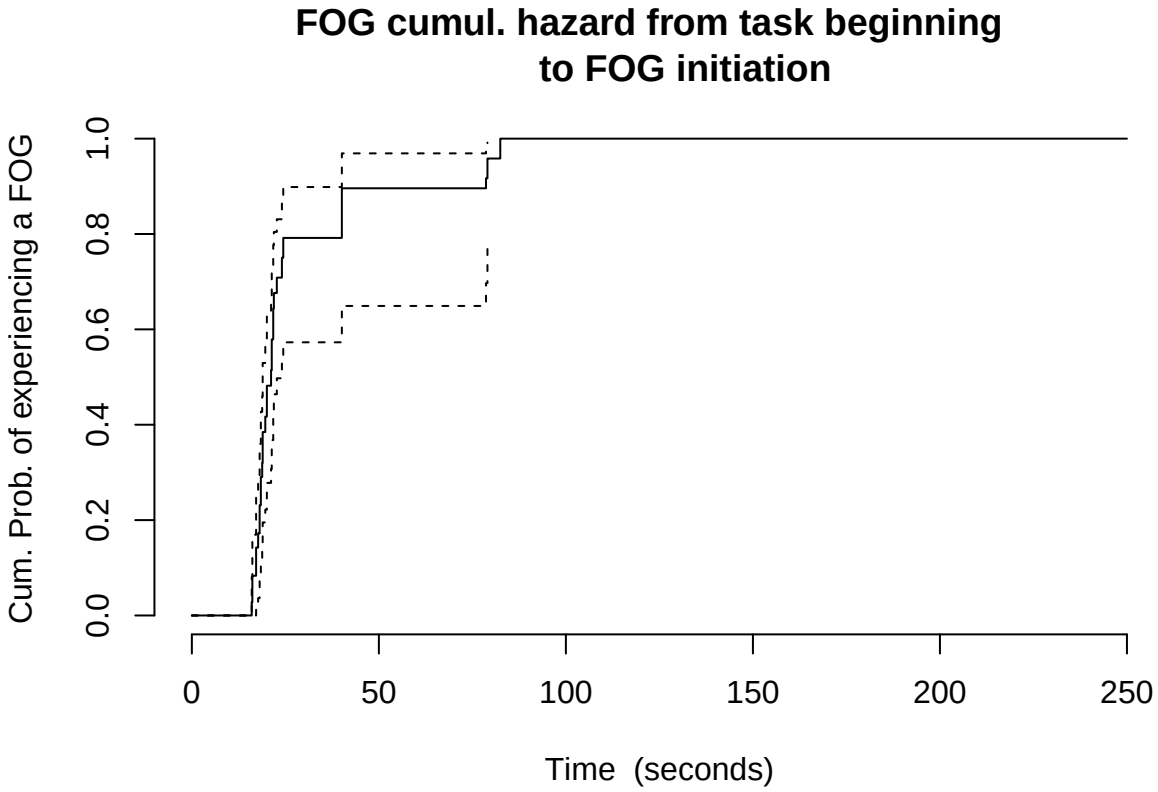
Medication

How many subjects are under parkinsonian-medication ?

```
## Medication
##   off    on
## 68.89 31.11
```

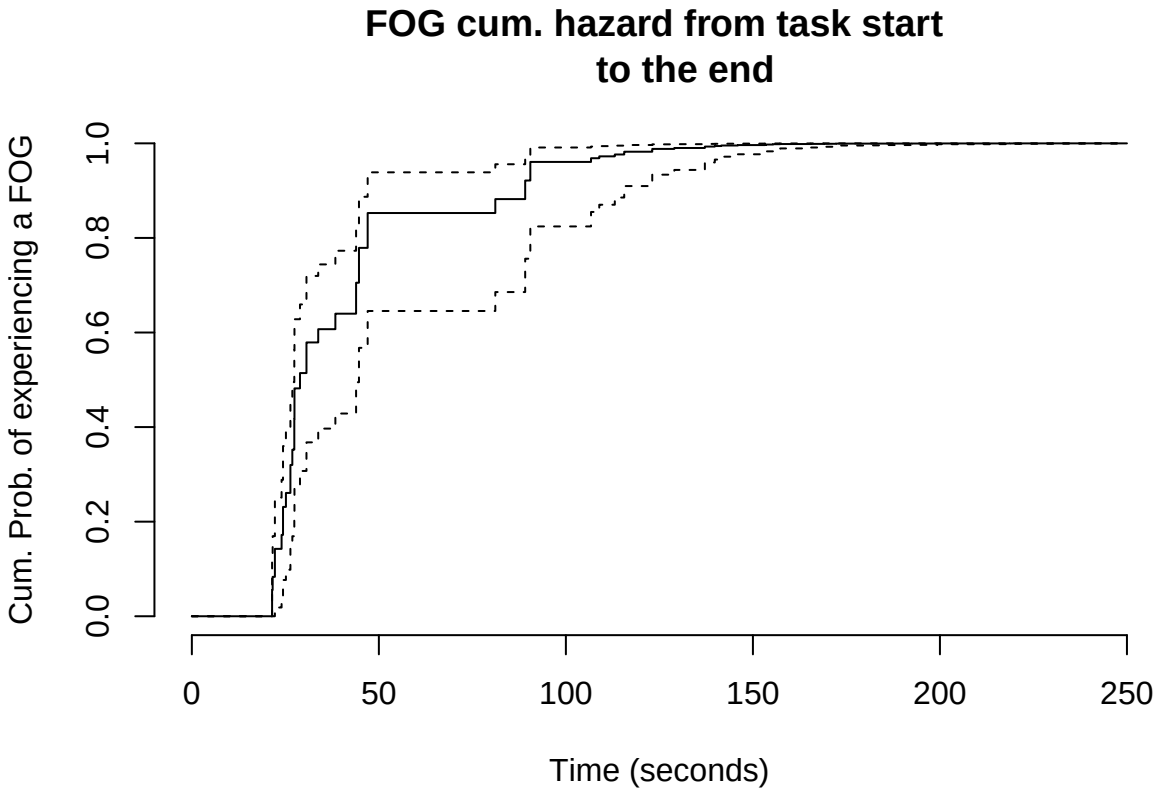
31.11% are under medication.

Modelling and Analysis : overall time to FOG.




```
## Call: survfit(formula = Surv(Begin, Init, Kinetic) ~ 1, data = fog)
##
##      records n.max n.start events median 0.95LCL 0.95UCL
## [1,]    2232   130      36   1530   21.2    18.9    24.1
```

When considering time to FOG initiation , half of patients experiment FOG within 21.183 s.




```
## Call: survfit(formula = Surv(Begin, End, Kinetic) ~ 1, data = fog)
##
##      records n.max n.start events median 0.95LCL 0.95UCL
## [1,]    2232   190      36   1530   28.9    27.4    44.7
```

When considering time to the end of the task, half of patients experiment FOG within 28.911 s.

Univariate analysis.

Sex

p-value is 0.15 .There is no effect of sex on time to FOG.

Medication

P-value is 0« 0.05. there is huge effect of Medication in experiencing FOG holding other variables fixed.

Task type

```
## Warning in agreg.fit(X, Y, istrat, offset, init, control, weights = weights, :
## Loglik converged before variable 6 ; beta may be infinite.
```

```
## # A tibble: 3 x 3
##   term                   estimate p.value
##   <chr>                     <dbl>   <dbl>
## 1 factor(Task)Hotspot1       1.03 0.0127
## 2 factor(Task)TUG-C          1.08 0.0186
## 3 factor(Task)Turning-DT    -1.24 0.00165
```

Three levels have significant effect on time to fog,holding other variables fixed.

events Type

```
## # A tibble: 0 x 3
## # ... with 3 variables: term <chr>, estimate <dbl>, p.value <dbl>
```

There is no effect of Task type on FOG occurrence.

visit

```
## # A tibble: 1 x 3
##   term           estimate   p.value
##   <chr>             <dbl>     <dbl>
## 1 factor(Visit)2    0.254 0.0000442
```

p-value « 0.05. Having more than one (1) round visit to hospital, has an influence on FOG occurrence.

yearSinceDx : year since diagnosis.

p-value : 0.5104841 is > 0.05. There is no effect of Year since diagnosis on FOG occurrence.

Multivariate analysis

```
## # A tibble: 2 x 5
##   term         estimate std.error statistic  p.value
##   <chr>           <dbl>     <dbl>     <dbl>    <dbl>
## 1 Medicationon    0.740    0.0702     10.5  5.84e-26
## 2 Visit           0.240    0.0630      3.81 1.36e- 4
```

Medication and Visit present a p value < 0.05.Being on medication and having visit to hospital has an effect on the occurrence of FOG in this population.

Results

Discussion

Conclusion

References

To be continued !!!!