Final Project Proposal

yuqingxia

2022-10-28

Data set description: This data set documents the party leadership succession in 23 parliamentary democracies. There are 25 columns and 4559 rows in the data, it includes the country, party information, name, sex, and term information about the leaders, and it also includes a status vector which use one to indicate the leader is still in office and 0 to indicate that they are out of office. There are, however, many missing values in the data set due to the lack of information for some countries.

Read in the data:

```
leaders <- read.csv("Karabulut_PartyLeadersData.csv")
```

Data set overview:

```
head(leaders)
```

```
##   country_id country election_year continent party_id
## 1          1 Albania          1991    Europe        1
## 2          1 Albania            NA    Europe        2
## 3          1 Albania            NA    Europe       NA
## 4          1 Albania            NA    Europe       NA
## 5          1 Albania            NA    Europe       NA
## 6          1 Albania            NA    Europe       NA
##                    party_name party_abbr party_position party_founded
## 1  Party of Labour of Albania        PLA                         1941
## 2 Democratic Party of Albania         PD                         1990
## 3                                                                  NA
## 4                                                                  NA
## 5                                                                  NA
## 6                                                                  NA
##   party_dissolved    leader_name  X sex age cochair in_year out_year tenure
## 1              NA Ramiz Tafë Alia     M  60              1985     1991    6.0
## 2              NA    Sali Berisha  I  M  47              1991     1992    1.0
## 3              NA    Sali Berisha II  M  53              1997     2013   17.0
## 4              NA   Eduart Selami     M  NA              1992     1995    3.0
## 5              NA    Tritan Shehu     M  46              1995     1997    2.0
## 6              NA    Lulzim Basha     M  39              2013       NA    8.5
##   status sysofgov sysofgov1 X.1 X.2 X.3 X.4
## 1      0        0         0  NA  NA
## 2      0        0         0  NA  NA
## 3      0        0         0  NA  NA
## 4      0        0         0  NA  NA
## 5      0        0         0  NA  NA
## 6      1        0         0  NA  NA
```

```r
dim(leaders)
```

```
## [1] 4559    25
```

Load Libraries:

```r
library(survival)
library(survminer)
```

```
## Loading required package: ggplot2
```

```
## Loading required package: ggpubr
```

```
##
## Attaching package: 'survminer'
```

```
## The following object is masked from 'package:survival':
##
##     myeloma
```

Survival Model Fit

```r
leaders$status <- ifelse(leaders$status ==  0, 1, 0)
fit <- survfit(Surv(tenure, status) ~ 1, data = leaders)
fit
```

```
## Call: survfit(formula = Surv(tenure, status) ~ 1, data = leaders)
##
##    901 observations deleted due to missingness
##          n events median 0.95LCL 0.95UCL
## [1,] 3658   3071      4       4       5
```
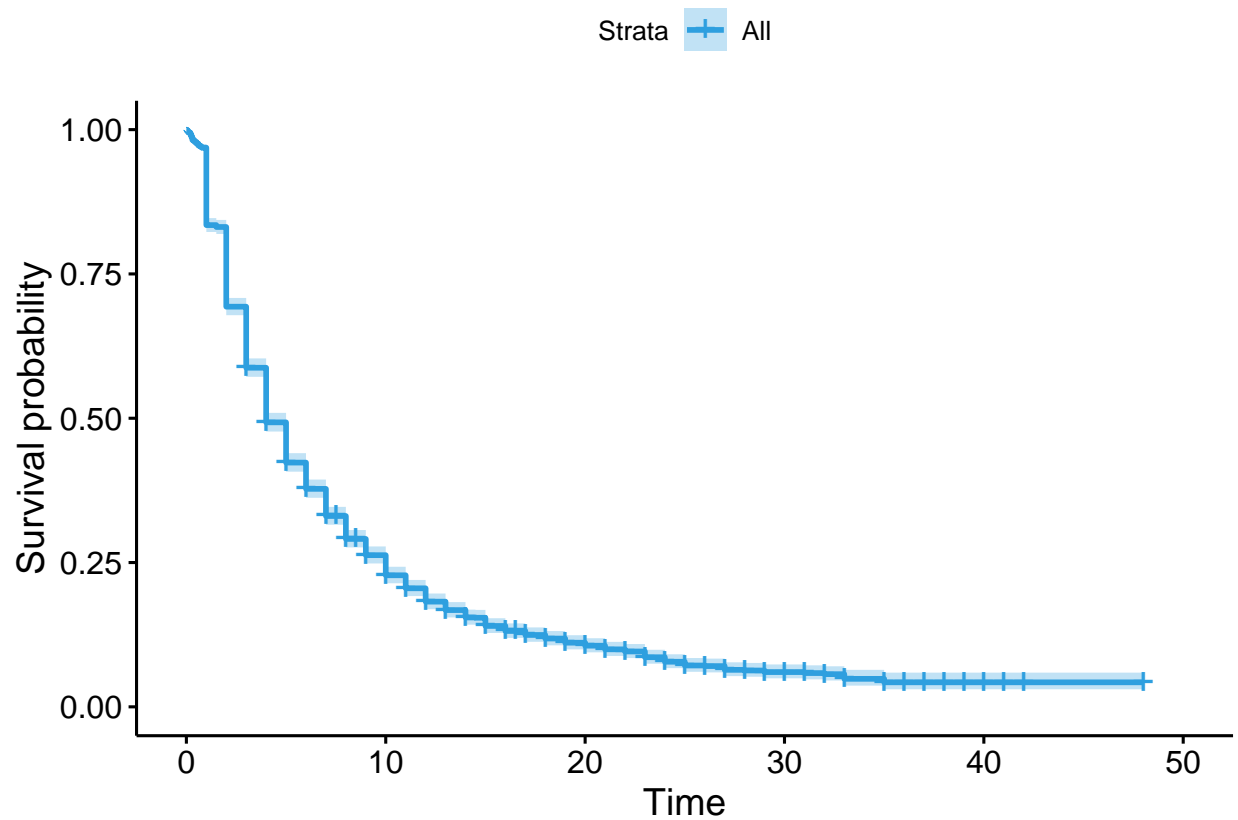
```r
# Drawing curves
ggsurvplot(fit, palette ="#2E9FDF")
```

Statistical tools: Besides the basic Kalpan Meier model, cox proportional hazard model, and testing functions such as anova and summary, I would like to use cluster function and recurrent event model to see if the year (as a continuous covariate) affect the tenure of a political party leader, and to identify correlated groups of observations that reveals some patterns from the data.

Citation: Horiuchi, Y., Laing, M., & Hart, P. 't. (2015). Hard acts to follow: Predecessor effects on party leader survival. Party Politics, 21(3), 357–366. https://doi.org/10.1177/1354068812472577