Uso del programa TestSurvRec para comparar curvas de superviviencia con eventos recurrentes.

C. M. Martínez


Resumen.-

El paquete TestSurvRec implementa las pruebas estadísticas para comparar dos curvas de supervivencia con eventos recurrentes. Este software ofrece herramientas útiles para el análisis de la supervivencia en el campo de la biomedicina, epidemiología, farmacéutica y otras áreas. El paquete TestSurvRec contiene dos conjuntos de datos con eventos recurrentes, un conjunto de datos referido al experimento de Byar que contiene los tiempos de recurrencia de tumores de cáncer de vejiga en los pacientes tratados con piridoxina, tiotepa o considerado como un placebo. Y otro conjunto de datos que contiene los tiempos de rehospitalización después de la cirugía en pacientes con cáncer colorectal. Estos datos provienen de un estudio que se llevó a cabo en el Hospital de Bellvitge, un hospital universitario público en Barcelona (España).

Palabras clave: Pruebas estadísticas, Eventos recurrentes, Lenguaje R, Curvas de supervivencia.

Using TestSurvRec Package for compare survival curves with recurrent events.

Abstract.-

The TestSurvRec package implements statistical tests to compare two survival curves with recurrent events. This software provides useful tools for the analysis of survival in the field of biomedicine, epidemiology, pharmaceutical and other areas. TestSurvRec package contains two sets of data with recurrent events, a data set based on the Byar experiment that shows the time of tumor recurrence in bladder cancer patients treated with pyridoxine, tiotepa or considered as a placebo. Another data set that contains rehospitalization times after surgery in patients with colorectal cancer. These data come from a study conducted at Bellvitge Hospital, a public teaching hospital in Barcelona (Spain).

Keywords: Statistical tests, Recurrent events, R-language, Survival curves.

Recibido: Mayo 2014
Aceptado: Agosto 2014

1. Introduction

The survival analysis is very important in many different research area. Tradicionalmente, the survival analysis is defined as a set of tools that are used for the time modelling to the occurrence of a event. The events can be, for example: failure of a machine, diseases of people, social events, labor accidents, length of remission, death or others. Survival analysis also is called as events history analysis. The study of survival data has focused on predicting the probability of survival of occurrence of an event, comparing of survival curves of populations groups and the estimations of the effects of covariable on the occurrence of the event. Today, the survival analysis research has focused
more on modeling phenomena where events are recurrent and where there are different causes that accelerate or retard the appearance of the same. The latest studies are known as survival analysis with competing risks. So with these studies, we can analyze occurrence of events on time. When the event occurs in the study period, the survival data for this unit is complete. If the event may not be observed on the study period, its produced so-called censored observation. There are three types of censoring. Censored observations can be for the left, for the right or for interval (both left and right). The studies with recurrent events are relative recent. On the last decade, a variate of models, methods and computational algorithms have been proposed for analysis these phenomena. This paper proposes a statististical package to compare survival curves of populations groups. A set of package to analyse survival data with recurrent events have been developed. Emphasize programs made on language as R, $S - plus$, SAS, and others. The package presented in this paper was developed in R language and it was implemented to compare survival curves of groups with recurrent events. This package was called “TestSurvRec”; and it was developed two years ago on CRAN. It was based on a package of CRAN called “survrec” such estimations of survival curves are made using PHS estimator. The survrec package was developed for González et al [1]. The package “TestSurvRec” was implemented to compare groups with a family of weighted tests type logrank. The statistics used on the work were developed on code R.

2. Modeling of survival curves with recurrent events

Many authors of the survival analysis with recurrent events have based theirs models on the Cox model for proportional hazards. For example, the model of Prentice et al. [2] (model PWP), Andersen-Gill [3] (model AG) and Wei et al [4] (model WLW) are some of them. PWP proposed two models for recurrent events. Both PWP models can be considered as extensions of the proportional hazards model. The AG model assumes that all events are independents and of the same type. WLW proposed a model for the analysis of recurrent events. The events may be recurrences of the same kind or events of different natures. Others authors have proposed non-parametric models to estimate the functions of the analysis. Wang-Chang [5] (model WC) proposed an estimator of the common marginal survivor function in the case where within-unit interoccurrence times are correlated. Peña, Strawderman and Hollander [6] (model PSH) developed the product limit estimator valid when the interoccurrence times are assumed to represent an IID sample. For this work, is used the PSH estimator to fit the survival curves of the groups.

$$S(s, t) = \prod_{w \leq t} \left[ 1 - \frac{N(s, \Delta w)}{Y(s, w)} \right]$$

The author defined two counting process, $N$ and $Y$. The notation introduced provides a convenient framework for the treatment of recurrent events. The processes are doubly indexed, with two time scales: calendar time $(s)$ and time intercurrence $(t)$. So, $N(s, t)$ counts the number of observed events that occur over a period of time $[0, s]$ whose time intercurrence times are at most $t$, and $Y(s, t)$ counts the number of events observed in the calendar time $[0, s]$ whose intercurrence times are at least $t$. The expressions (2) and (3) show a more formal definitions of these processes. For more information on these processes refer to the work of Peña et al [6].

\[ N(s, t) = \sum_{i=1}^{n} \sum_{j=1}^{\infty} I \{ S_{ij} = s \} \]

\[ Y(s, t) = \sum_{i=1}^{n} \left[ \sum_{j=1}^{K_{i}(s)} I \{ T_{ij} \geq t \} + I \{ \min(s, \tau_{i}) - S_{ij} \geq t \} \right] \]

Where,

\[ K_{i}(s) = \sum_{j=1}^{n} I \{ S_{ij} \leq s \} \]

2.1. PSH model for survival curves with recurrent events by group

On this work, PSH estimator was extended to define the survival curves with recurrent events by group, see eq. [5]. For identify the group,
we defined a stratification variable called variable of group (denoted by the letter \( r \)). This variable permits identify the estimator the curves of each groups. So,

\[
S_r(t) = \prod_{w \leq t} \left[ 1 - \frac{\Delta N(s,w;r)}{Y(s,w)} \right] \quad \forall \ r = 1, 2
\]

(3)

And, \( S_r(t) \) represents the survival curve of the \( r \)th groups.

3. Statistical tests

The purpose of this package is to present statistical tests for the analysis of recurrent events data. On the year Martínez et al. [7], its publicated one paper where were proposed statistical tests to compare survival curves with recurrent events. The hypothesis is,

\[
\begin{align*}
H_0 : S_1(t) &= S_2(t) \\
H_1 : S_1(t) &\neq S_2(t)
\end{align*}
\]

Where, \( S_1(t) \) and \( S_2(t) \) are survival curves of the groups. The statistic of test is,

\[
Z_r = \frac{\sum_{s \leq t} w_z[\Delta N(s,z;r) - E[\Delta N(s,z;r)]]}{\sqrt{\sum_{s \leq t} w_z^2 Var[\Delta N(s,z;r)]]}} \quad \forall \ r = 1, 2
\]

The statistic \( Z_r \) has a normal asymptotic behaviour, its squared has a chi-square approximate behaviour with degree of freedom. \( \Delta Z \) is approaches zero and \( \Delta N(s,z;1) \) has a hypergeometric behaviour with expected value equal to \( Y(s,z;1) \times \Delta N(s,z)/Y(s,z) \) and variance equal to,

\[
Var \{ \Delta N(s,z;1) \} = \frac{Y(s,z)-Y(s,z;1)}{Y(s,z)-1} Y(s,z;1)pq
\]

with, \( p = \frac{\Delta N(s,z)}{Y(s,z)} \) and \( q = 1 - p \)

Where,

\[
\Delta N(s,z;1) = N(s,z + \Delta z;1) - N(s,z;1)
\]

It was proposed various types of weights \( (w_z) \) for these tests. See expression \( [4] \).

\[
w_z = [S(z)]^\gamma [1 - S(z)]^\beta \frac{[Y(s,z)]^\alpha}{[Y(s,z) + 1]^\beta}
\]

(4)

The appropriate selection of weights for survival analysis depends of the curves behavior. With the selection of the values of the parameters \( (\alpha, \beta, \gamma \) and \( \eta \) on the proposal, this statistical is able of adjustment to this behavior. With the proposal, we are able to make studies on survival analysis with recurrent events and of generate alternative tests for the analysis, as for example, the classical tests types logrank, Gehan, Peto-Peto, Fleming-Harrington and so on. If all the parameters are zero imply that \( w_z = 1 \), generates the test type logrank. If, \( \alpha = 1 \) and the other parameters are zero \( w_z = Y(s,z) \), generates the test type Gehan. If, \( \gamma = 1 \) and the other parameters are zero \( w_z = S(z) \), generates the test of Peto-Peto. If, \( \gamma = 1, \eta = 1 \) and the others parametres are zero, generates the family of tests type Fleming-Harrington. We can demonstrate that with this statistic test, they are available for recurrent events and are able to generate others tests similar as for example the classical survival analysis.

4. TestSurvRec Package

In this paper, its have been considered events that may occur more than once over the follow-up time for a given subject called recurrent events. TestSurvRec package permits adjusts and compares the survival functions of populations groups with events of these types, see Martínez [8]. The estimations are make through of PSH estimator using right-censored data. TestSurvRec use other package to adjusted the survival curves of the groups. This package is the called survival, see González [1]. To compare the survival curves of the groups, TestSurvRec uses CMrec tests that are tests to generalize the classical logrank tests for recurrent events with right-censored data. CMrec tests were proposed for Martínez [7]. The tests are implemented in the TestSurvRec package and are available on CRAN. The package contains a set of useful functions for survival analysis with recurrent events and are used two databases. One dataset refers to Byar’s experiment that contains the recurrence of bladder cancer tumors Byar [9], and other contains the rehospitalization times after surgery in patients with colorectal cancer González et al [10]. Both data sets are used to explain as implemented the functions of the
4.1. Datasets of the TestSurvRec package

In this package two datasets are used, both as an example of making the analysis estimations. The first dataset corresponding to Byar’s experiment refers to the recurrence of bladder cancer tumours in patients treated with pyridoxine, thiotepa and placebo and other dataset concern to the rehospitalations of patients with colorectal cancer. This database can be found in several works by authors who work with recurring events, as for example WLW. The datasets used here, are from medicine area. The purpose is to increase the number of datasets on the package to explore applications other area, such as: engineering, social science, among others. The first dataset provided a study of bladder cancer tumours that was conduced by veterans administration cooperative urological of research group in a hospital of United States. On the experiment the patients with superficial bladder tumours were included in the trial and were treated randomized. The patients with presence of tumours, these were removed. Our aim in the package was proposal statistical tests to compare survival curves of the groups. The second dataset contains times after surgery in patients with colorectal cancer. The study took place in the hospital of Bellvitge in Barcelona (Spain). Four hundred and three patients with colon and rectum cancer have been included in the study. The variables considered were: sex, age, tumour site (rectum, colon), tumuor stage (Dukes classifications: AB, C and D), type treatment (chemotherapy, radiotherapy) and distance from living place to hospital (less than 30 km, more than 30 km). Both datasets are used to explainas works the functions of the package and they are available on CRAN. The dataset names for both experiments are shows as fellow,

- TBCplapyr
- TBCplathi
- TBCpyrthi
- DataColonDukesABvsD
- DataColonDukesCvsD
- DataColonDukesABvsC

This datasets can be read from “TestSurvRec”package. For example, the DataColonDukesABvsC file can be read from R using the code:

```r
> require(TestSurvRec)
> data(DataColonDukesABvsC)
> XL<-dataframe(DataColonDukesABvsC)
> print(XL[1 : 11])
```

DataColonDukesABvsC data contains rehospitalizations time after surgery of patients with colorectal cancer with Duke tumoral stage type AB and type C. The code before produced the following output:

On this output shows the data of eleven observations of patients treated with Duke tumoral stage type AB and type C. Full data frame is compased of 655 observations for one total of four hundred and three patients and a total of ten variable (columns). This data frame contains the following columns: Obs variable is an identificator of the observation number. Iden is an identificator of each subject. Too, the Id variable is an identificator of the subject. If, the observation is recurrent then the Iden and Id indentificator are repeat. Tinicio variable is the initial time of observation just before each recurrence. Tinicio variable is the start time of each interval. Tinicio variable is zero if \( j = 0 \) and \( Tinicio = t_{ij} \) if \( j = 1, 2, ..., n \). The time variable is the time of each interval or observation time, \( time = t_{ij} \). Tcal variable is the end time of each interval. Tcal variable is \( time \) if \( j = 0 \) and if \( j = 1, 2, ..., n \). Tcal = Tinicio + time on each observation. Tcal variable is 1 for recurrence and 0 for everything else (censored data). chemoter, dukes and distance are study variables of the dataset.

4.2. Functions of TestSurvRec package

This section illustrated, how you can use “TestSurvRec”package functions and obtain estimations of the \( p \)-values on the statistic tests for...
Tabla 1: DataColonDukesABvsC data contains rehospitalizations time after surgery of patients with colorectal cancer with Duke tumoral stage type AB and type C.

<table>
<thead>
<tr>
<th>Obs</th>
<th>Iden</th>
<th>id</th>
<th>Tinicio</th>
<th>time</th>
<th>Tcal</th>
<th>event</th>
<th>chemoter</th>
<th>dukes</th>
<th>distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10767</td>
<td>2</td>
<td>0</td>
<td>489</td>
<td>489</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>10767</td>
<td>2</td>
<td>489</td>
<td>693</td>
<td>1182</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>15843</td>
<td>3</td>
<td>0</td>
<td>15</td>
<td>15</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>15843</td>
<td>3</td>
<td>15</td>
<td>768</td>
<td>783</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>19655</td>
<td>4</td>
<td>0</td>
<td>163</td>
<td>163</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>19655</td>
<td>4</td>
<td>163</td>
<td>125</td>
<td>288</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>19655</td>
<td>4</td>
<td>288</td>
<td>350</td>
<td>638</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>19655</td>
<td>4</td>
<td>638</td>
<td>48</td>
<td>686</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>19655</td>
<td>4</td>
<td>686</td>
<td>1362</td>
<td>2048</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>25109</td>
<td>5</td>
<td>0</td>
<td>1134</td>
<td>1134</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>25109</td>
<td>5</td>
<td>1134</td>
<td>10</td>
<td>1144</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

recurrent events. With this package, we can obtain the graphs of the curves for both groups whose estimates are obtained with the PSH estimator, see expression (1).

- Dif.Surv.Rec Function
- Print.Summary Function
- Gen.Data.RE Function
- Plot.Data.Events Function
- Plot.Event.Rec
- Plot.Surv.Rec

4.2.1. Dif.Surv.Rec Function

Dif.Surv.Rec function calculates and computes the statistical difference between two survival curves using the expressions (6) and (9). The survival curves are computes using the expression (5). The following is a example of the instructions in R language to get the results.

- require(TestSurvRec)
- data(TBCplapyr)
- Dif.Surv.Rec(TBCplapyr,"all",0,0,0,0).

The output is as follow,

<table>
<thead>
<tr>
<th>Nomb.Est</th>
<th>Chi.square</th>
<th>p.value</th>
</tr>
</thead>
<tbody>
<tr>
<td>LRrec</td>
<td>0.3052411</td>
<td>0.5806152</td>
</tr>
<tr>
<td>Grec</td>
<td>1.4448446</td>
<td>0.2293570</td>
</tr>
<tr>
<td>TWrec</td>
<td>0.9551746</td>
<td>0.3284056</td>
</tr>
<tr>
<td>PPrecc</td>
<td>1.1322772</td>
<td>0.2872901</td>
</tr>
<tr>
<td>PMrec</td>
<td>1.1430319</td>
<td>0.2850126</td>
</tr>
<tr>
<td>PPrecc</td>
<td>1.1834042</td>
<td>0.2766641</td>
</tr>
<tr>
<td>HFrecc</td>
<td>0.3052411</td>
<td>0.5806152</td>
</tr>
<tr>
<td>CMrec</td>
<td>0.3052411</td>
<td>0.5806152</td>
</tr>
<tr>
<td>Mrec</td>
<td>1.5298763</td>
<td>0.2161310</td>
</tr>
</tbody>
</table>

In this example, we use data from patients with bladder cancer who received treatment, (placebo or pyridoxine). The results of statistical tests indicate that the hypothesis of equal survival curves for both groups can not be rejected. That is, the supply of pyridoxine with bladder cancer have no significant effect on tumor recurrence compared to patients treated with placebo. Also, you can see another interesting result, note that when all the test parameters are zero, the same result on CMrec, LRrec and HFrec is obtained.

4.2.2. Print.Summary Function

This function is used to print results of statistics tests to compare survival curves of groups with recurrent events. Print.Summary function
computes and plots survival curves for recurrent event data using PSH estimator. This function calculates the statistical difference between two survival curves. The estimations of survival curves are made in (5) and the statistical difference is estimated with CMrec tests, see (6). The following R code produces the output that is shows above.

```r
> data(TBCplapyr)
> Print.Summary(TBCplapyr)
```

It produces the output

<table>
<thead>
<tr>
<th>Group = 0</th>
<th>time</th>
<th>n.event</th>
<th>n.risk</th>
<th>surv</th>
<th>std.error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>127</td>
<td>0.984</td>
<td>0.0110</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>124</td>
<td>0.913</td>
<td>0.0243</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>14</td>
<td>113</td>
<td>0.800</td>
<td>0.0340</td>
<td></td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>1</td>
<td>0.244</td>
<td>0.0422</td>
<td></td>
</tr>
<tr>
<td></td>
<td>31</td>
<td>1</td>
<td>0.225</td>
<td>0.0427</td>
<td></td>
</tr>
<tr>
<td></td>
<td>35</td>
<td>1</td>
<td>0.200</td>
<td>0.0439</td>
<td></td>
</tr>
<tr>
<td></td>
<td>42</td>
<td>1</td>
<td>0.236</td>
<td>0.0582</td>
<td></td>
</tr>
<tr>
<td></td>
<td>44</td>
<td>1</td>
<td>0.189</td>
<td>0.0599</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group = 1</th>
<th>time</th>
<th>n.event</th>
<th>n.risk</th>
<th>surv</th>
<th>std.error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>84</td>
<td>0.964</td>
<td>0.0199</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>81</td>
<td>0.893</td>
<td>0.0327</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>12</td>
<td>73</td>
<td>0.746</td>
<td>0.0447</td>
<td></td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>1</td>
<td>0.283</td>
<td>0.0514</td>
<td></td>
</tr>
<tr>
<td></td>
<td>42</td>
<td>1</td>
<td>0.236</td>
<td>0.0582</td>
<td></td>
</tr>
<tr>
<td></td>
<td>44</td>
<td>1</td>
<td>0.189</td>
<td>0.0599</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 group</td>
</tr>
<tr>
<td>1 Pooled Group</td>
</tr>
<tr>
<td>2 1er Group</td>
</tr>
<tr>
<td>3 2do Group</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Nomb.Est</th>
<th>Chi.square</th>
<th>p.value</th>
</tr>
</thead>
<tbody>
<tr>
<td>LRrec</td>
<td>0.3052411</td>
<td>0.5806152</td>
</tr>
<tr>
<td>Grec</td>
<td>1.4448446</td>
<td>0.2293570</td>
</tr>
<tr>
<td>TRec</td>
<td>0.9551746</td>
<td>0.3284056</td>
</tr>
<tr>
<td>PPrec</td>
<td>1.1322772</td>
<td>0.2872901</td>
</tr>
<tr>
<td>PMrec</td>
<td>1.1430319</td>
<td>0.2850126</td>
</tr>
<tr>
<td>PFrec</td>
<td>1.1834042</td>
<td>0.2766641</td>
</tr>
<tr>
<td>HFrec</td>
<td>0.3052411</td>
<td>0.5806152</td>
</tr>
<tr>
<td>CMrec</td>
<td>0.3052411</td>
<td>0.5806152</td>
</tr>
<tr>
<td>Mrec</td>
<td>1.5298763</td>
<td>0.2161310</td>
</tr>
</tbody>
</table>

The output shows the estimates and graphics of survival curves of the treated groups including the survival curve of pooled group. The tests results indicates that the curves are not different. On the output, too are illustrate medians survival times. The first group is the placebo one and the second group is the treated group with pyridoxine. The graphic shows that the placebo group had lightly better survival experence than pyridoxine group. The 50% of the placebo patients experiment the recurrence tumour on 9 months, whereas about 38% of the pyridoxine patients treated, they have the recurrence in six months.

Survival curves of patients treated with placebo and pyridoxine. TestSurvRec package

### 4.2.3. Tests.Bonf.K.Groups.RE Function

This function is used to compare two or more survival curves of populations groups with recurrent events. This test is based on Bonferroni’s method and it was developed by Martínez [11]. The idea is applied to this methodology with a sequential procedure to control type I error on the multiple tests. This type I error occurs when the null hypothesis is true, but is rejected. The hypothesis test is,

\[ H_0 : S_1(z) = S_2(z) = \ldots = S_k(z) \]
\[ H_1 : \text{At least one } S_r(z) \text{ is different with } r = 1, 2, \ldots, k \]

Where, \( S_r(z) \) is the survival curve of the \( r \)th group and its estimation is calculated by the Generalized Product Limit Estimator (GPLE) too called PSH estimator. The procedure consist in use the sequential to make multiple contrasts. The null hypothesis of the test on each step is,

\[ H_0 : S_r(z) = S_{r'}(z) \]
\[ H_1 : S_r(z) \neq S_{r'}(z) \]

The procedure is as follow:

- Start by comparing the groups in pairwise. This method have a total number of test, \( q = k \times (k - 1)/2 \). Use an order of comparison. The pair \( (r,r') \) indicates that \( r \) groups contrasted...
with \( r' \) group. Use tests proposed by Martínez [7].

- Order the p-values obtained: \( p_{01} \leq p_{02} \leq \ldots \leq p_{0q} \).
- Compare \( p_{0i} \) with Bonferroni’s critical value, \( \alpha/q \).
- if \( p_{0i} \) is greater than \( \alpha/q \), the null hypotheses \( (H_0_1 \) and \( H_0_0 \) are accepted and the procedure is finished.
- if \( p_{0i} \) is less than \( \alpha/q \), the null hypotheses \( (H_0_0 \) and \( H_0_0 \) are rejected.
- The global null hypothesis is rejected and the procedure is finished.

To use Test.Bonf.K.Group.RE function, you must use the following syntax,

\[
\text{Test.Bonf.K.Group.RE(yy, power, type, alfa, use the following syntax,}
\]

- \( yy \rightarrow \) Data type recurrent events. Examples: TBCplapyr
- \( power \rightarrow \) Power of the test
- \( type \rightarrow \) Type of test for recurrent events
- \( alfa \rightarrow \) parameter of the test
- \( beta, gamma, eta \rightarrow \) parameter of the test

Below an example of using the R code is shown,

```r
Gen.Data.RE(100,5,0.5,"rnorm",30,10,"runif",0.01,70,3,1)
```

The output is as shown below,

<table>
<thead>
<tr>
<th>( j )</th>
<th>( id )</th>
<th>( time )</th>
<th>( Tcal )</th>
<th>( event )</th>
<th>( group )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0.00</td>
<td>23.98</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>23.98</td>
<td>8.17</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>0.00</td>
<td>2.87</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>0.00</td>
<td>22.32</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>22.32</td>
<td>26.78</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>49.07</td>
<td>5.13</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>258</td>
<td>100</td>
<td>0.00</td>
<td>21.6</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>259</td>
<td>100</td>
<td>21.6</td>
<td>14.4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>260</td>
<td>100</td>
<td>36.0</td>
<td>32.3</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

### 4.2.4. Gen.Data.RE Function

Gen.Data.RE function permits generates survival data with recurrent events. The survival data with and without censoring are independents. Users can define the distribution type which can adjust the occurrences time of the event and censoring times. To use the function, you must use the following syntax,

\[
\text{Gen.Data.RE(N,ERmax,p,dist1,p11,p12,dist2,exact,p21,p22,p23)}
\]

This function plot data with recurrent events. To use the function, you must use the following syntax,

Plot.Data.Events(yy, paciente, inicio, dias, censored, especiales,colevent, colcensor)

Its arguments are,

- \( yy \rightarrow \) Data type recurrent events. Examples: TBCplapyr
- \( paciente \rightarrow \) Vector of number of units on the data base
- \( inicio \rightarrow \) Vector, its assumed that the units are observed from one time equal to zero.
- \( dias \rightarrow \) Vector of the periods of observations of the study units
- \( censored \rightarrow \) vector of times of censorship for each unit
- \( especiales \rightarrow \) Three-column matrix containing the identification of the units, times of occurrence of the event and type of event.
- \( colevent \rightarrow \) Color event identifier.
- \( colcensor \rightarrow \) Color censored data identifier.

Below an example of using the R code is shown,
It produces one plot for data with recurrent events

4.2.6. Plot.Event.Rec Function

The counting processes are powerful tools in survival analysis. These processes consider a time scale, a calendar time, and a gap time. This idea originally provides from Gill [12] and the concept was extended by Peña et al [6]. This function is used for plot the recurrence of an event on a two scales time, one scale with the gap times and other scale with the calendar times. This authors designed a special graphic, that allows to count the occurrence of events per unit time. Doubly indexed processes illustrate for an case. To use the function, you must use the following syntax,

\[
\text{Plot.Event.Rec}(yy, xy, xf)
\]

Its arguments are,

- \(yy\) -> Object type recurrent events data. Example: TBCplapyr
- \(xy\) -> Identification of the unit to plotted. "1" is defect value.
- \(xf\) -> Argument to plot the occurrence events of the unit \(xf\). "1" is defect value.

It produces the output that shows above,

Plot for data with recurrent events

The graphic shows a case followed during 24,01 months. This patient presents four recurrences at months 7, 10, 16 and 24 from the beginning of study. This fact implies that interoccurrence times are 7, 3, 6, 8 and the censored time correspond to 0,01 months. Let us assume that we are interested in computing the single processes, \(N(t)\) and \(Y(t)\) for a selected interoccurrence time \(t = 5\). In this case \(N(t = 5) = 1\) and \(Y(t = 5) = 3\). For the calendar time scale, \(s = 20\), we have \(N(s = 20) = 3\) and \(Y(s = 20) = 1\). Now, let us assume that we would like to know double-indexed processes for both selected interoccurrence and calendar times. Using both time scales we observe that \(N_{14}(s = 20, t = 5) = 1, Y_{14}(s = 20, t = 5) = 2\) and \(\Delta N_{14}(s = 20, t = 6) = 1\).

4.2.7. Plot.Surv.Rec Function

This function survival curves for both groups are plotted. Both curves are estimated using survrec package, it is available in R language. This important say, that the PSH estimator is of valid use when it assumed that the inter-occurrence times are IID. Its obvious that this assumption is restrictive in biomedical applications and its use is more valid on the field of engineering. We describe here the argument that needed this function. To use the function, you must use the following syntax,

\[
\text{Plot.Surv.Rec}(XX)
\]

The argument is,

- \(XX\) -> Object type recurrent events data. Indicates the name of the data file. Example: TBCplapyr
It produces the output that shows above,

![Survival Curves of the groups](image)

Survival curves for both groups

5. Results.

A new version of the TestSurvRec package has been presented through this paper, which is available from CRAN, we illustrate various new functions for the survival analysis with recurrent events, specifically for comparison tests. Also, the article shows how to build database for the survival analysis with recurrent events and right-censored data. By developing the TestSurvRec package in R, we hope to have provided a useful, and pertinent tool for many research areas.

Acknowledgement

This work was approved and presented on the Medical Research Council Conference on Biostatistics 2014 on the poster session, in celebration of the MRC Biostatistics Unit’s Centenary year on Cambridge (UK). It was supported by the CDCH-UC, (Consejo de Desarrollo Científico y Humanístico de la Universidad de Carabobo), and the author would like to thank for your finance to present this research.

Referencias