miRecSurv Package: Prentice-Williams-Peterson Models with Multiple Imputation of Unknown Number of Previous Episodes

by David Moriña, Gilma Hernández-Herrera and Albert Navarro

Abstract Left censoring can occur with relative frequency when analyzing recurrent events in epidemiological studies, especially observational ones. Concretely, the inclusion of individuals that were already at risk before the effective initiation in a cohort study may cause the unawareness of prior episodes that have already been experienced, and this will easily lead to biased and inefficient estimates. The **miRecSurv** package is based on the use of models with specific baseline hazard, with multiple imputation of the number of prior episodes when unknown by means of the COMPoisson distribution, a very flexible count distribution that can handle over, sub, and equidispersion, with a stratified model depending on whether the individual had or had not previously been at risk, and the use of a frailty term. The usage of the package is illustrated by means of a real data example based on an occupational cohort study and a simulation study.

Introduction

It is not unusual in cohort epidemiological studies that part (or all) of the participants have experienced the event under study at least once before the beginning of the follow-up. This situation is particularly common in the case of observational designs. Under these circumstances, the prior history and prior time at risk of these individuals can be unknown or estimated on the basis of self recall questionnaires, which could lead to modeling issues when the baseline hazard of suffering the event is time-dependent. If the event of interest can only occur once, and it occurred for an individual before the start of the follow-up, the result for this individual is fixed regardless of the duration of the follow-up. Therefore we are in the well-known and well-studied situation of left censoring of a binary variable, for which specific modeling techniques are available. On the other hand, if the event of interest can be suffered several times by the same individual (it is recurrent), and the number of events suffered by the individuals in the cohort before the beginning of the follow-up is unknown, we face a left censoring situation with a discrete censored variable that can define different baseline hazards depending on the episode an individual is at risk of.

This paper introduces the **miRecSurv**, useful when the prior history is unknown for all, or some, of the individuals included in a cohort and the outcome of interest is a recurrent event with event dependence. Specifically, we suppose that we know the moment from where all individuals are at risk, but the number of episodes experienced by the individuals in this time is unknown. This is a realistic situation in practice; for example, it is very probable that in a workforce cohort, we know when a worker started to work (thus, to be at risk of having a sick leave) and, however, and especially for people with ample trajectory, that we do not know whether or not in effect they have already had sick leaves (and in this case, how many). Another situation that might deal with this issue is a study of cohorts with an outcome of incidence of infection from human papillomavirus on adult women. It would be relatively simple to know how long they have been at risk (beginning of active sexual life) or to make a reasonable assumption. However, we will not be able to know the number of infections since most of the time, when they occur, they are asymptomatic (Fernández-Fontelo et al., 2016).

Methods

Theoretical approach

Our proposal starts from the assumption that even though the previous history of all or some of the individuals is unknown, we do know which of these were at risk prior to the beginning of the follow-up and starting when. Further, that is based fundamentally on three considerations: 1) impute k, the number of previous episodes for those subjects at risk before the beginning of the follow-up; 2) treat the subpopulation of subjects "Previously at risk" separately from those "Not previously at risk", and 3) use a frailty term basically to capture the error that will be made when imputing k. Concretely, in the two formulations, "Counting process" (Eq. 1) and "Gap time" (Eq. 2), the ones we

call "Specific Hazard Frailty Model Imputed", in its versions - Counting Process (SHFMI.CP) and Gap Time (SHFMI.GT):

$$\lambda_{ikr}(t) = \nu_i \lambda_{0kr}(t) e^{X_i \beta} \tag{1}$$

$$\lambda_{ikr}(t) = \nu_i \lambda_{0kr}(t - t_{k-1}) e^{X_i \beta}, \tag{2}$$

where k will be the number of previous episodes of individual i in case they are known or their imputed value in case they are not known; r indicates the subpopulation the individual belongs to: "Previously at risk" or "Not previously at risk". In both cases, information corresponding to the time to risk prior to $t = 0$ for each individual is included in $X_i\beta$, which will be zero for all those that start to be at risk as of the beginning of the follow-up and a value different than zero for those that were previously at risk. In practice, this proposal means that we stratify by the interaction between having been at risk or not before the beginning of the follow-up and the number of previous episodes.

Therefore, the use of the term individual random error ν_i intends to capture the error that will be made when imputing and the effect of any variable that having a potential effect, would not have been considered in the analysis. Stratifying by the number of prior episodes intends to safeguard the event dependence. Doing it as an interaction with the fact that it is an individual previously at risk, or not, separates the two subpopulations to not mix times that are not comparable, on the same scale. More details on the proposed methodology can be found in Hernández-Herrera et al. (2021).

The imputation of the number of previous events in individuals at risk before the beginning of the follow-up is done through the COMPoisson generalized distribution (using the log link function), that allows adjusting a regression model using the Conway-Maxwell Poisson (COMPoisson) distribution (Shmueli et al., 2005) considering the dispersion of the data (sub, equi, or overdispersion). This imputation is carried out through multiple imputation calculating its parameters directly from the observed data in two phases: Firstly, a generalized linear model (GLM) is fitted using the number of episodes observed during the follow-up as the response variable and the covariates selected by the user as explanatory, based on the COMPoisson distribution. Imputed values are randomly sampled from this distribution with the parameters obtained in the previous step, including random noise generated from a normal distribution. In order to produce a proper estimation of uncertainty, the described methodology is included in a multiple imputation framework, according to the well known Rubin's rules (Rubin, 1987) and based on the following steps in a Bayesian context:

1. Fit the COMPoisson count data model and find the posterior mean and variance $\hat{\beta}$ and $V(\hat{\beta})$ of model parameters β.
2. Draw new parameters β^* from $N(\hat{\beta}, V(\hat{\beta}))$.
3. Compute predicted scores p using the parameters obtained in the previous step.
4. Draw imputations from the COMPoisson distribution and scores obtained in the previous step.

The COMPoisson random number generation is based on the `rcom` function included in the archived package **compoisson** (Dunn, 2012). The overperformance of the COMPoisson distribution in this context compared to alternative discrete distributions (Poisson, negative binomial, Hermite and zero-inflated versions) is discussed in Hernández-Herrera et al. (2020) by means of a comprehensive simulation study.

The miRecSurv package

The main function of the **miRecSurv** is `recEvFit`, which allows the user to fit recurrent events survival models. A call to this function might be

```
recEvFit(formula, data, id, prevEp, riskBef, oldInd,
         frailty=FALSE, m=5, seed=NA, ...)
```

The description of these arguments can be summarized as follows:

- `formula`: a formula object, with the response on the left of a $\sim$ operator and the terms on the right. The response must be a survival object as returned by the `Surv` function.
- `data`: a data.frame in which to interpret the variables named in the formula.
- `id`: subject identifier.
- `prevEp`: known previous episodes.
- `riskBef`: indicator for new individual in the cohort (`riskBef=FALSE`) or subject who was at risk before the start of follow-up (`riskBef=TRUE`).

- `oldInd`: time an individual has been at risk prior to the follow-up. This time can be positive or negative (time origin as the start of follow-up).
- `frailty`: should the model include a frailty term. Defaults to `FALSE`.
- `m`: number of multiple imputations. The default is `m=5`.
- `seed`: an integer that is used as argument by the `set.seed` function for offsetting the random number generator. Default is to leave the random number generator alone.
- `...`: extra arguments to pass to `coxph`.

The output of this function is a list with seven elements:

- `fit`: a list with all the `coxph` objects fitted for each imputed dataset.
- `coeff`: a list with the vectors of coefficients from the models fitted to each imputed dataset.
- `loglik`: a list with the loglikelihood for each model fitted.
- `vcov`: a list with the variance-covariance matrices for the parameters fitted for each of the imputed datasets.
- `AIC`: a list with the AIC of each of the models fitted.
- `CMP`: summary tables of the fitted COMPoisson models used for imputing missing values.
- `data.impute`: the original dataset with the multiple imputed variables as final columns.

In order to facilitate the interpretation, a pooled summary table is available via `summary` method, formatted in a very similar way to the summary tables of very well-known functions as `coxph`, as can be seen in the next section. As in some cases the multiple imputation process might be computationally expensive and take some time, the function `recEvFit` provides the user with a progress bar.

Examples

Simulation study

To illustrate our proposal, we use simulated data generated with the parameters estimated in a worker cohort, where the outcome is the occurrence of sick leave due to any cause. Table 1 shows the characteristics of each episode in this population, estimated in a cohort study described in Navarro et al. (2012). The maximum number of episodes that a subject may suffer was not fixed, although the baseline hazard was considered constant when $k \geq 4$. X_1, X_2, and X_3 are covariates that represent the exposure, with $X_i \sim Bernoulli(0.5)$, $i = 1, 2, 3$, and $\beta_1 = 0.25$, $\beta_2 = 0.5$, and $\beta_3 = 0.75$ being their parameters that represent effects of different magnitudes, set independently of the episode k to which the worker is exposed. All the simulations were conducted using the R package **survsim** (Moriña and Navarro, 2014), and all the code to reproduce these analyses is available as supplementary material.

Episode	Distribution	β_0	Ancillary
1	Log-logistic	7.974	0.836
2	Weibull	7.109	0.758
3	Log-normal	5.853	1.989
4	Log-normal	5.495	2.204

Table 1: Characteristics of the simulated population.

To illustrate the usage of the **miRecSurv** package, the results corresponding to a sample from the first scenario can be obtained by means of

```
library(survsim)
library(miRecSurv)
d.ev    <- c('llogistic','weibull','weibull','weibull')
b0.ev   <- c(5.843, 5.944, 5.782, 5.469)
a.ev    <- c(0.700, 0.797, 0.822, 0.858)
d.cens  <- c('weibull','weibull','weibull','weibull')
b0.cens <- c(7.398, 7.061, 6.947, 6.657)
a.cens  <- c(1.178, 1.246, 1.207, 1.422)
```

```
set.seed(1234)
sample1 <- rec.ev.sim(n=1500, foltime=1095,
                      dist.ev=d.ev, anc.ev=a.ev, beta0.ev=b0.ev,
                      dist.cens=d.cens, anc.cens=a.cens, beta0.cens=b0.cens,
                      beta=list(c(-.25,-.25,-.25,-.25), c(-.5,-.5,-.5,-.5), c(-.75,-.75,-.75,-.75)),
                      x=list(c("bern", .5), c("bern", .5), c("bern", .5)),
                      priskb=.1, max.old=5475)
sample1$old2 <- -sample1$old
sample1$old2[is.na(sample1$old)] <- 0

### Shared frailty
ag_s1 <- coxph(Surv(start2,stop2,status)~as.factor(x)+as.factor(x.1)+as.factor(x.2)+old2+
               strata(as.factor(risk.bef))+frailty(nid), data=sample1)

### Counting process
shfmi.cp_s1 <- recEvFit(Surv(start2, stop2, status)~x+x.1+x.2, data=sample1,
                        id="nid", prevEp = "obs.episode",
                        riskBef = "risk.bef", oldInd = "old", frailty=TRUE, m=5, seed=1234)

### Gap time
shfmi.gt_s1 <- recEvFit(Surv(stop2-start2, status)~x+x.1+x.2, data=sample1,
                        id="nid", prevEp = "obs.episode",
                        riskBef = "risk.bef", oldInd = "old", frailty=TRUE, m=5, seed=1234)
```

The generated cohort, including the estimated number of previous events (multiple imputed), can be obtained as

```
head(shfmi.cp_s1$data.impute)
  nid real.episode obs.episode       time status    start     stop      time2   start2
1   1            1           1 119.673502      1   0.0000 119.6735 119.673502 124.5052
2   1            2           2 389.637244      1 119.6735 509.3107 389.637244 244.1787
3   1            3           3 244.130315      1 509.3107 753.4411 244.130315 633.8160
4   1            4           4 144.476145      0 753.4411 897.9172 144.476145 877.9463
5   2            1           1 107.943850      1   0.0000 107.9438 107.943850 681.4178
6   2            2           2   6.472291      1 107.9438 114.4161   6.472291 789.3617
      stop2 old risk.bef long z x x.1 x.2 EprevCOMPoissDef1 EprevCOMPoissDef2
1  244.1787   0    FALSE   NA 1 1   0   0                 0                 0
2  633.8160   0    FALSE   NA 1 1   0   0                 1                 1
3  877.9463   0    FALSE   NA 1 1   0   0                 2                 2
4 1022.4224   0    FALSE   NA 1 1   0   0                 3                 3
5  789.3617   0    FALSE   NA 1 1   1   0                 0                 0
6  795.8340   0    FALSE   NA 1 1   1   0                 1                 1
  EprevCOMPoissDef3 EprevCOMPoissDef4 EprevCOMPoissDef5
1                 0                 0                 0
2                 1                 1                 1
3                 2                 2                 2
4                 3                 3                 3
5                 0                 0                 0
6                 1                 1                 1
```

The pooled coefficients table can be obtained as

```
summary(shfmi.cp_s1)
Call:
recEvFit(formula = Surv(start2, stop2, status) ~ x + x.1 + x.2,
    data = sample1, id = "nid", prevEp = "obs.episode", riskBef = "risk.bef",
    oldInd = "old", frailty = TRUE, m = 5, seed = 1234)

Coefficients:
                 coef exp(coef)  se(coef)        Chisq     Pr(>|z|)
x        2.766059e-01 1.3186465 0.3161915  76.49190798 1.401017e-17
x.1      4.337341e-01 1.5430085 0.3010033 174.86750236 2.970463e-39
x.2      7.489812e-01 2.1148444 0.5144217 436.60796033 8.875830e-95
old     -2.584687e-05 0.9999742 0.4976880   0.03005029 7.830972e-01
frailty            NA        NA        NA  77.64460457 6.481409e-02
```

```
AIC:  34867.6

COM-Poisson regression model for imputing missing values
              Estimate     SE   z.value  Pr(>|z|)
(Intercept)    -6.3611 0.0348 -182.5394 0.000e+00
x               0.2356 0.0240    9.8376 7.755e-23
x.1             0.3563 0.0258   13.7869 3.054e-43
x.2             0.6446 0.0305   21.1139 5.927e-99
S:(Intercept)  -0.5029 0.0315  -15.9637 2.288e-57
```

Simulated data include four scenarios of $n = 1500$ subjects, with a maximum follow-up time of 3 years and a maximum time at risk prior to the beginning of the cohort of 15 years. The first scenario has a 10% of subjects at risk prior to the beginning of the cohort (i.e., we do not know the number of previous episodes in 10% of the subjects), whilst the second, third, and fourth samples have 25%, 50%, and 100%, respectively. Samples based on these settings were generated 100 times and the estimates were averaged across simulations for each sample.

To compare the results obtained, we also estimate the shared frailty model (Eq 3). This model has a common hazard baseline (i.e., does not consider the event dependence) and incorporates an individual frailty term.

$$\lambda_{ikr}(t) = \nu_i \lambda_0(t) e^{X_i \beta} \tag{3}$$

Results of fitting the described models in scenario 1 are summarized in the Table 2:

Model	$\hat{\beta_1}$	$SE(\hat{\beta_1})$	$\hat{\beta_2}$	$SE(\hat{\beta_2})$	$\hat{\beta_3}$	$SE(\hat{\beta_3})$
Shared Frailty	0.292	0.043	0.558	0.043	0.843	0.043
SHFMI.CP	0.244	0.034	0.470	0.036	0.706	0.039
SHFMI.GT	0.218	0.030	0.419	0.031	0.632	0.034

Table 2: Average estimates obtained on scenario 1 (10% of subjects at risk prior to the beginning of the cohort).

Below are presented the results for the second scenario, Table 3:

Model	$\hat{\beta_1}$	$SE(\hat{\beta_1})$	$\hat{\beta_2}$	$SE(\hat{\beta_2})$	$\hat{\beta_3}$	$SE(\hat{\beta_3})$
Shared Frailty	0.287	0.039	0.554	0.039	0.832	0.040
SHFMI.CP	0.242	0.032	0.472	0.035	0.703	0.040
SHFMI.GT	0.217	0.028	0.425	0.030	0.633	0.034

Table 3: Average estimates obtained on scenario 2 (25% of subjects at risk prior to the beginning of the cohort).

Results for scenario 3, Table 4:

Model	$\hat{\beta_1}$	$SE(\hat{\beta_1})$	$\hat{\beta_2}$	$SE(\hat{\beta_2})$	$\hat{\beta_3}$	$SE(\hat{\beta_3})$
Shared Frailty	0.277	0.033	0.542	0.034	0.818	0.035
SHFMI.CP	0.243	0.030	0.474	0.036	0.710	0.042
SHFMI.GT	0.217	0.025	0.421	0.029	0.632	0.033

Table 4: Average estimates obtained on scenario 3 (50% of subjects at risk prior to the beginning of the cohort).

Finally, Table 5 shows the estimates when 100% of the subjects are at risk prior to the beginning of the follow-up:

The average relative bias of the estimates produced by each method is shown in Figure 1. It can be seen that the estimates produced by the **miRecSurv** are less biased than those based on the common baseline hazard model.

Model	$\hat{\beta_1}$	$SE(\hat{\beta_1})$	$\hat{\beta_2}$	$SE(\hat{\beta_2})$	$\hat{\beta_3}$	$SE(\hat{\beta_3})$
Shared Frailty	0.262	0.026	0.527	0.027	0.797	0.029
SHFMI.CP	0.248	0.029	0.495	0.040	0.742	0.053
SHFMI.GT	0.208	0.022	0.416	0.027	0.626	0.034

Table 5: Average estimates obtained on scenario 4 (100% of subjects at risk prior to the beginning of the cohort).

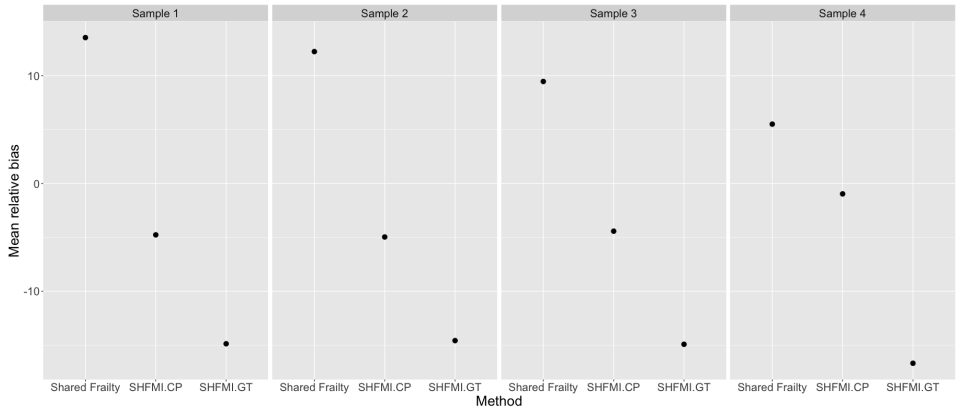

Figure 1: Average relative bias of the estimates produced by method (Shared Frailty, SHFMI.CP, SHFMI.GT) in each scenario (extreme left: 10% of subjects at risk prior to the beginning of the cohort; left: 25%; right: 50%; extreme right: 100%).

Example

The real example corresponds to a selected sample of the HC-UFMG cohort (Reis et al., 2008), which involves workers of a public hospital in Belo Horizonte, Brazil. Specifically, we selected all workers present on January 1st, 2004 ($n = 512$) and all of them whose contract with the hospital was signed from that date until the end of the follow-up, December 31st, 2007 ($n = 884$). The outcome of interest was the occurrence of sick leave for any diagnosis, with three covariates: *contract* (civil servant; non civil servant), *type of work* (healthcare professionals; non-healthcare) and *sex* (female; male). The incidence rate was 11.7 sick leaves per 100 worker-months, and the median of follow-up per worker was 24.3 months. Figure 2 represents a random subsample of 30 workers from this cohort, and it can be seen that some of the workers were at risk of suffering the event of interest before the start of the follow-up, and these episodes cannot be seen and must be estimated. Time 0 is the start of the follow-up of each worker, being January 1st, 2004 for all those who were contracted in the hospital before that date, or the starting date of their contract among those who were contracted later.

To analyze the association of sex, type of contract, and the kind of job developed by these workers over the risk of suffering sick leaves, the package **miRecSurv** can be used in the following way.

```
mod1 <- recEvFit(Surv(start, stop, status)~Male+Non.civil.servant+
                                          Non.healthcare, data=example,
                 id="nid", prevEp="num", riskBef="prev2", oldInd="days_prev",
                 frailty=TRUE, m=5, seed=1234)
summary(mod1)
Call:
recEvFit(formula = Surv(start, stop, status) ~ Male + Non.civil.servant +
    Non.healthcare, data = example, id = "nid", prevEp = "num",
    riskBef = "prev2", oldInd = "days_prev", frailty = TRUE,
    m = 5, seed = 1234)

Coefficients:
                          coef exp(coef)  se(coef)       Chisq     Pr(>|z|)
Male              -0.2966733104 0.7432868 0.2033092   24.386339 3.999350e-07
Non.civil.servant -0.1788549829 0.8362272 0.1320313    9.221709 2.391622e-03
Non.healthcare    -0.1315412853 0.8767431 0.1318741    5.568327 1.703076e-02
days_prev         -0.0002383671 0.9997617 0.2027526    6.457974 7.089736e-03
```

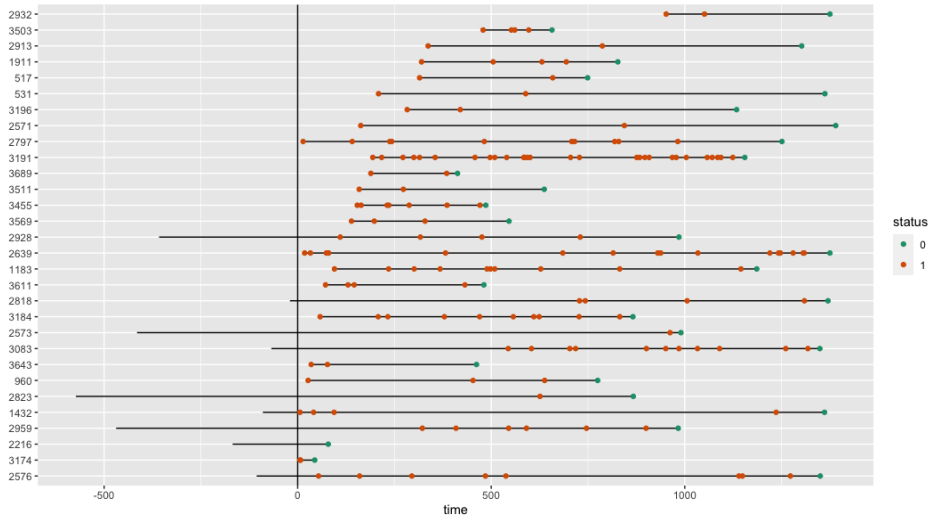

Figure 2: Evolution in time of a subsample of 30 workers from the HC-UFMG cohort. The vertical line represents the start of follow-up.

```
frailty                      NA         NA         NA 1157.231797 5.664063e-35

AIC:  29699

COM-Poisson regression model for imputing missing values
                  Estimate     SE   z.value  Pr(>|z|)
(Intercept)        -6.7602 0.0407 -166.0266 0.000e+00
Male               -0.1665 0.0261   -6.3766 1.810e-10
Non.civil.servant  -0.0132 0.0246   -0.5354 5.924e-01
Non.healthcare     -0.0917 0.0246   -3.7297 1.917e-04
S:(Intercept)      -1.4043 0.0709  -19.8059 2.646e-87
```

According to these results, it can be seen that all three considered explanatory variables are significantly associated with the risk of suffering sick leaves (see Coefficients table), although the type of contract is not significant on the model used to impute the number of episodes suffered before the start of the follow-up.

Conclusion

Left censoring, when analyzing recurrent events, is a situation that can occur with a certain frequency in cohort studies. For example, it will occur in every cohort that follows subjects that were already at risk of experiencing the outcome of interest before the beginning of the follow-up. If the recurrent event presents event dependence, not knowing the number of episodes that each individual has had prior to the effective initiation of the follow-up, and analyzing this through "classical" methods, is an important problem, as it will lead to biased and inefficient estimates.

The presented proposal seems to work reasonably well, outperforming the alternative to a common hazard model. It is important to carry out a comprehensive study to evaluate the performance of the presented proposal. It is also worth noticing that although in this particular example, the average relative biases produced by the shared frailty model are relatively low, this is not always the case, as can be seen in a simulation study like Navarro et al. (2017).

Unavailability of the number of previous episodes that an individual has had should not be a justification for the use of models with a common baseline hazard which on the large majority of occasions show a higher bias and less coverage than specific baseline hazard models, as shown by the results of this work.

Bibliography

J. Dunn. *compoisson: Conway-Maxwell-Poisson Distribution*, 2012. URL https://CRAN.R-project.org/package=compoisson. R package version 0.3. [p]

A. Fernández-Fontelo, A. Cabaña, P. Puig, and D. Moriña. Under-reported data analysis with INAR-hidden Markov chains. *Statistics in Medicine*, 35(26):4875–4890, nov 2016. ISSN 10970258. doi: 10.1002/sim.7026. [p]

G. Hernández-Herrera, A. Navarro, and D. Moriña. Regression-based imputation of explanatory discrete missing data. *arXiv preprint*, 2020. URL http://arxiv.org/abs/2007.15031. [p]

G. Hernández-Herrera, D. Moriña, and A. Navarro. Left-censored recurrent event analysis in epidemiological studies: a proposal when the number of previous episodes is unknown. *arXiv preprint*, 2021. URL http://arxiv.org/abs/2102.11279. [p]

D. Moriña and A. Navarro. The R package survsim for the simulation of simple and complex survival data. *Journal of Statistical Software*, 59(2):1–20, 2014. URL http://www.jstatsoft.org/v59/i02/. [p]

A. Navarro, D. Moriña, R. Reis, F. B. Nedel, M. Martín, and S. Alvarado. Hazard functions to describe patterns of new and recurrent sick leave episodes for different diagnoses. *Scandinavian journal of work, environment & health*, 38(5):447–55, sep 2012. ISSN 1795-990X. doi: 10.5271/sjweh.3276. URL http://www.ncbi.nlm.nih.gov/pubmed/22286954. [p]

A. Navarro, G. Casanovas, S. Alvarado, and D. Moriña. Analyzing recurrent events when the history of previous episodes is unknown or not taken into account: proceed with caution. *Gaceta sanitaria*, 31(3):227–234, 2017. [p]

R. J. Reis, P. d. F. La Rocca, L. Basile, A. Navarro, and M. Martín. Cohort profile: The hospital das clínicas cohort study, belo horizonte, minas gerais, brazil. *International Journal of Epidemiology*, 37(4): 227–234, 2008. [p]

D. B. Rubin. *Multiple Imputation for Nonresponse in Surveys*. John Wiley, 1987. ISBN 047108705X. [p]

G. Shmueli, T. P. Minka, J. B. Kadane, S. Borle, and P. Boatwright. A useful distribution for fitting discrete data: revival of the Conway–Maxwell–Poisson distribution. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 54(1):127–142, jan 2005. ISSN 1467-9876. doi: 10.1111/J.1467-9876.2005.00474.X. URL https://rss.onlinelibrary.wiley.com/doi/full/10.1111/j.1467-9876.2005.00474.x. [p]

David Moriña
Department of Econometrics, Statistics and Applied Economics - Riskcenter (IREA), Universitat de Barcelona (UB)
Avinguda Diagonal, 690 (08034) Barcelona
Spain
Centre de Recerca Matemàtica (CRM)
ORCID: 0000-0001-5949-7443
dmorina@ub.edu

Gilma Hernández-Herrera
Instituto de Investigaciones Médicas, Facultad de Medicina, Universidad de Antioquia
Carrera 51 D No. 62-29. Edificio Manuel Uribe Ángel, Medellín
Colombia
PhD program in Methodology of Biomedical Research and Public Health. Autnomous University of Barcelona (UAB)
Avinguda de Can Domènech, S/N (08193) Cerdanyola del Vallès
Spain
gilma.hernandez@udea.edu.co

Albert Navarro
Research group on Psychosocial Risks, Organization of Work and Health (POWAH), Autonomous University of Barcelona (UAB)
Biostatistics Unit, Faculty of Medicine, Autonomous University of Barcelona (UAB)
Avinguda de Can Domènech, S/N (08193) Cerdanyola del Vallès
Spain
albert.navarro@uab.cat