

Bootstrapping Median Survival with Recurrent Event Data

Juan R. González*

Cancer Prevention and Control Unit, Catalan Institute of Oncology, Avda.
Gran via s/n, km. 2.7, Hospitalet de Llobregat 08907, Spain.

and

Edsel A. Peña†

Department of Statistics, The University of South Carolina, LeConte
College, Columbia, SC 29208, USA.

SUMMARY. In this paper, several resampling schemes to estimate the sampling distributions of median estimators of the inter-event time of a recurrent event are introduced and studied through simulations. Two types of recurrent event models are considered: first is a model where the inter-event times are independent and identically distributed, and second is a model where the inter-event times are associated, with the association arising from a gamma frailty model. The procedures studied are anchored on estimators proposed in Peña, Strawderman and Hollander (2001, *Journal of the American Statistical Association* **96**, 1299-1315) and Wang and Chang (1999, *Journal of the American Statistical Association* **94**, 146-153). The resampling procedures

* *email:* jrgonzalez@ico.scs.es

† *email:* pena@stat.sc.edu

are then employed to analyze a data set pertaining to hospital readmission of patients with colorectal cancer.

KEY WORDS: Gamma frailty model; resampling schemes; sum-quota data accrual.

1. Introduction

Recurrent event data are ubiquitous in longitudinal studies arising in a wide variety of settings, such as biomedicine, psychiatry, engineering, or sociology. Some examples of recurrent events are repeated hospitalization due to a chronic disease, epileptic seizures, small bowel motility in gastroenterology, depression, breakdown of a mechanical or electronic system, stoppage of a nuclear power plant, or auto insurance claims.

Statistical inference in the presence of recurrent event data has been considered by several authors such as Gill (1981), Vardi (1982a), Vardi (1982b), McClean and Devine (1995), Soon and Woodroffe (1996), Wang and Chang (1999) (WC), and Peña et al. (2001) (PSH). A main aspect with this type of data is the sum-quota accrual scheme which leads to an informative stopping rule as well as an informative censoring mechanism. Except in PSH (2001), most papers have used restrictive data accrual and censoring schemes for recurrent event data to avoid the two difficulties mentioned above. In PSH (2001) it was assumed that the interoccurrence times represent independent and identically distributed (i.i.d.) observations from an unknown continuous distribution F , and that each subject is observed for a possibly random period of time. As a consequence, the number of event occurrences for a subject or unit is a random variable whose distribution depends on, hence informative about, F . Moreover, the last observation for each subject is al-

ways right-censored, with the censoring variable depending on the length of the observation period and on the previous interoccurrence times for that subject, rendering the censoring mechanism to become informative.

This model is quite reasonable in engineering and reliability settings, but in the biomedical context it is somewhat restrictive because in biomedical settings (i) the distribution of the time of the first event may differ from the interoccurrence distribution for succeeding events; (ii) the interoccurrence times may be correlated; and (iii) the interoccurrence times may depend on relevant covariates. For problem (i) the WC (1999) and PSH (2001) estimators circumvented this problem by assuming that the initial occurrence of the event is also the criterion for admission into the study. For problem (ii) the WC (1999) estimator or another estimator proposed in PSH (2001) under the case where the within-subject interoccurrence times follow a gamma frailty model may be utilized. Finally, extensions of survival models based on the Cox proportional hazards approach may be employed to take problem (iii) into account (cf., Andersen and Gill (1982), Wei et al. (1989), and Prentice et al. (1981), among others).

The major goal of this paper is to study bootstrapping schemes for estimating the sampling distribution of estimators of the median of the event interoccurrence time distribution in the presence of recurrent event data. In particular, this will enable us to estimate standard errors of the estimators, and thereby construct bootstrap confidence intervals for the median. Ultimately, this will provide a mechanism for comparing the median survival times for different groups of subjects, the grouping possibly arising from the values of covariates.

We have organized this paper as follows. In section 2 we introduce notation and present the WC (1999) and PSH (2001) estimators. Different bootstrapping schemes are described in section 3 under the the i.i.d. model and a correlated interoccurrence times model. In section 4 a simulation is used to compare and discuss the statistical properties of median survival time estimated using the different bootstrap plans. Section 5 describes the software developed to estimate the bootstrap confidence intervals of median survival time. Finally, in Section 6 we apply these procedures to a data set of hospitality readmissions in patients with colorectal cancer. Our conclusions are given in Section 7.

2. Estimation of survival function

2.1 *Mathematical Setting*

We suppose that n independent subjects (e.g., units) are available in the study. For the i th subject, we denote the successive interoccurrence times of the recurrent event of interest by $\{T_{ik}, k = 1, 2, \dots\}$. We will first assume that the event interoccurrence times are i.i.d. nonnegative random variables with a common absolutely continuous distribution function $F(t) = P\{T_{ij} \leq t\}$. We assume that monitoring of the i th subject ceases at a possibly random time τ_i , where $\tau_1, \tau_2, \dots, \tau_n$ are i.i.d. with a common distribution function $G(w) = P\{\tau_i \leq w\}$. We also assume that τ_i and T_{ij} are mutually independent. We denote by $G_n(t)$ the empirical distribution of the τ_i 's.

For each $i = 1, \dots, n$, let $S_{i0}=0$ and $S_{ij} = \sum_{l=1}^j T_{il}$, $j = 1, 2, \dots$. The number of event occurrences for the i th unit is

$$K_i = \max\{k \in \{0, 1, \dots\} : S_{ik} \leq \tau_i\}, \quad (1)$$

and the observable random variables for the i th unit are

$$(K_i, \tau_i, T_{i1}, T_{i2}, \dots, T_{iK_i}, \tau_i - S_{iK_i}). \quad (2)$$

However, as pointed out in the Introduction, the i.i.d. assumption is restrictive in biomedical settings, so we also need to consider other models. A specific type of model that results in correlated within-subject interoccurrence times is a multiplicative frailty model (cf., Andersen et al. (1993); Murphy (1995)). In this model it is postulated that there exists for each subject an unobservable positive-valued frailty Z_i such that, conditionally on $Z_i = z_i$, the interoccurrence times T_{i1}, T_{i2}, \dots are i.i.d. with common conditional survivor function

$$\bar{F}(t | Z_i = z) = [\bar{F}_0(t)]^z = \exp\left(-z \int_0^t \lambda_0(u) du\right) \quad (3)$$

where $\lambda_0(\cdot)$ is the hazard function associated with a baseline survivor function $\bar{F}_0(\cdot)$. The frailties Z_1, Z_2, \dots, Z_n are assumed to be i.i.d. from an unknown distribution function H . In general, the Z 's are not observed, so we are interested in estimating the marginal survivor of T_{ij} which under this model is given by

$$\bar{F}(t) = \mathbf{E} \{ \exp(-Z_1 \Lambda_0(t)) \} \quad (4)$$

where $\Lambda_0(t) = -\log[\bar{F}_0(t)]$ is the cumulative hazard function of \bar{F}_0 .

A common choice of the unknown frailty distribution H is a gamma distribution with shape and scale parameters both equal to an unknown parameter α . In this case, the common marginal survivor function \bar{F} in (4) becomes

$$\bar{F}(t) = \left[\frac{\alpha}{\alpha + \Lambda_0(t)} \right]^\alpha. \quad (5)$$

The parameter α controls the degree of association between interoccurrence times within subject. In particular, as α increases (decreases), association between interoccurrence times decreases (increases). Letting $\alpha \rightarrow \infty$, we obtain a model with independent interoccurrence times in which the T_{ij} has a common survivor function of \bar{F}_0 .

2.2 PSH estimator of \bar{F}

Peña et al. (2001) developed a nonparametric maximum likelihood estimator of the inter-event time survivor function under the assumption of i.i.d. model. This generalizes the product-limit estimator to the situation where the event is recurrent. To describe this estimator, we first need to introduce some notation. For a given calendar time s and a gap time t , we define by

$$\begin{aligned} K_i(s) &= \sum_{j=1}^{\infty} I\{S_{ij} \leq s\}; \\ N(s, t) &= \sum_{i=1}^n \sum_{j=1}^{K_i(s)} I\{T_{ij} = t\}; \\ Y(s, t) &= \sum_{i=1}^n \left[\sum_{j=1}^{K_i(s-)} I\{T_{ij} \geq t\} + I\{\min(s, \tau_i) - S_{iK_i(s-)} \geq t\} \right]. \end{aligned}$$

The PSH (2001) generalized product-limit estimator of the common survivor function \bar{F} of the event interoccurrence times is given by

$$\hat{\bar{F}}(s, t) = \prod_{\{w \leq t\}} \left[1 - \frac{N(s, \Delta w)}{Y(s, w)} \right]. \quad (6)$$

If $s = \infty$, then a more simplified form is obtained for in this case $K_i(s-) = K_i$. The resulting simplified estimator is the one utilized in the simulations and numerical illustrations.

PSH (2001) also proposed an estimator (referred to as FRMLE in their paper) of the common marginal distribution of the interoccurrence time distribution in the case of correlated interoccurrence times induced by a gamma

frailty model. They showed that the estimation of α and Λ_0 of (5) can be obtained via the maximization of the marginal likelihood function of α and $\Lambda_0(\cdot)$ and with an implementation of the expectation-maximization (EM) algorithm [see, for details, PSH (2001)]. This estimator of (5) is of form

$$\tilde{\bar{F}}(s, t) = \left[\frac{\hat{\alpha}}{\hat{\alpha} + \hat{\Lambda}_0(s, t)} \right]^{\hat{\alpha}} \quad (7)$$

where $\hat{\Lambda}_0(s, t)$ is an estimator of the marginal cumulative hazard function $\Lambda_0(t)$.

2.3 WC estimator of \bar{F}

Wang and Chang (1999) proposed an estimator of the common marginal survivor function in the case where within-unit interoccurrence times are correlated. They consider a correlation structure which is quite general, and includes as special cases both the i.i.d. and gamma frailty models. Setting all their weights to be equal to 1, their estimator is described below. For the i th unit, define

$$K_i^* = I\{K_i = 0\} + K_i I\{K_i > 0\}$$

and define the processes

$$\begin{aligned} d^*(t) &= \sum_{i=1}^n \frac{1}{K_i^*} \sum_{j=1}^{K_i} I\{T_{ij} = t\}; \\ R^*(t) &= \sum_{i=1}^n \frac{1}{K_i^*} \left[\sum_{j=1}^{K_i} I\{T_{ij} \geq t\} + I\{\tau_i - S_{iK_i} \geq t\} I\{K_i = 0\} \right], \end{aligned}$$

and with \mathcal{T} denoting the set of distinct observed complete interoccurrence times for the n units. The WC estimator of \bar{F} is given by

$$\hat{S}(t) = \prod_{\{T_k \in \mathcal{T}; T_k \leq t\}} \left[1 - \frac{d^*(T_k)}{R^*(T_k)} \right]. \quad (8)$$

This estimator possesses less bias than the generalized product-limit estimator when interoccurrence times are correlated within subjects. For more discussions concerning these estimators and the comparisons of their properties, refer to PSH (2001).

3. Bootstrapping Schemes

We will estimate the sampling distribution of estimators of the median survival time according to several competing bootstrap schemes described below. Arguably, the new contribution of the present paper is the examination of the question of how to do bootstrapping in the presence of recurrent event data arising from a sum-quota data accrual scheme. In the schemes below, the number of bootstrap replications is denoted by B .

Plan I: (Bootstrapping the observed data)

Obtain B i.i.d. samples of form

$$\{(K_i^*, \tau_i^*, T_{i1}^*, T_{i2}^*, \dots, T_{iK_i}^*, \tau_i^* - S_{iK_i}^*), i = 1, 2, \dots, n\},$$

with replacement, from the observed sample

$$\{(K_i, \tau_i, T_{i1}, T_{i2}, \dots, T_{iK_i}, \tau_i - S_{iK_i}), i = 1, \dots, n\}.$$

For each sample, compute the generalized PLE \hat{F} of \bar{F} , and compute the resulting estimator of the median, i.e., $\text{median}(\hat{F})$. From these B median estimates, a bootstrap estimate of the sampling distribution of the median estimator is obtained. Consequently, a bootstrap estimate of the standard error of the median estimator can be obtained, and a bootstrap confidence interval could also be constructed.

Plan II: (Bootstrapping T_{ij}^* 's from the PSH estimator)

Let \hat{F} be the generalized PLE estimator of \bar{F} . For $i = 1, \dots, n$, a bootstrap sample is generated as follows:

Step 1. Take $\tau_i^* = \tau_i$;

Step 2. From the distribution \hat{F} , continue generating an i.i.d sequence of T_{ij}^* 's until K_i^* where

$$\sum_{j=1}^{K_i^*} T_{ij}^* \leq \tau_i^* < \sum_{j=1}^{K_i^*+1} T_{ij}^*.$$

Step 3. The bootstrap sample for the i th unit is

$$(K_i^*, \tau_i^*, T_{i1}^*, T_{i2}^*, \dots, T_{iK_i^*}^*, \tau_i^* - S_{iK_i^*}^*)$$

where $S_{ij}^* = \sum_{l=1}^{K_i^*} T_{il}^*$.

Step 4. For this bootstrap sample, compute the generalized PLE \hat{F} of \bar{F} , and compute the associated median estimate.

Repeat Steps 1–4 a total of B times. The B estimates of the median provide data for obtaining the bootstrap estimates of the sampling distribution, standard error, and for constructing a bootstrap confidence interval.

Plan III: (Bootstrapping T_{ij}^* 's from \hat{F} and τ_i^* 's from G_n)

This scheme is analogous to that of Plan II, except that for each bootstrap sample, $\tau_i^*, i = 1, 2, \dots, n$, is an i.i.d. sample from the empirical distribution G_n .

Plan IV: (Bootstrapping T_{ij}^* 's from WC estimator)

Let \hat{S} be the WC estimator of \bar{F} . For $i = 1, \dots, n$, a bootstrap sample is generated as follows:

Step 1. Take $\tau_i^* = \tau_i$;

Step 2. From the distribution \hat{S} , continue generating an i.i.d sequence of T_{ij}^* 's until K_i^* where

$$\sum_{j=1}^{K_i^*} T_{ij}^* \leq \tau_i^* < \sum_{j=1}^{K_i^*+1} T_{ij}^*.$$

Step 3. The bootstrap sample for the i th unit is

$$(K_i^*, \tau_i^*, T_{i1}^*, T_{i2}^*, \dots, T_{iK_i^*}^*, \tau_i^* - S_{iK_i^*}^*)$$

where $S_{ij}^* = \sum_{l=1}^{K_i^*} T_{il}^*$.

Step 4. For this bootstrap sample, compute the WC estimator \hat{S} of \bar{F} , and compute the associated median estimate.

Repeat Steps 1–4 a total of B times. The B estimates of the median provide data for obtaining the bootstrap estimates of the sampling distribution, standard error, and for constructing a bootstrap confidence interval.

Plan V: (Bootstrapping T_{ij}^* 's from \hat{S} and τ_i^* 's from G_n)

This scheme is analogous to that of Plan IV, except that for each bootstrap sample, $\tau_i^*, i = 1, 2, \dots, n$, is an i.i.d. sample from the empirical distribution G_n .

Plan VI: (Semiparametric Bootstrap)

Let \tilde{F} be the FRMLE estimator of \bar{F} .

Step 1. Given the data, estimate $\hat{\alpha}$, the frailty parameter, and $\hat{\Lambda}_0$, the cumulative hazard function associated with $\bar{F}_0(t)$. Then, estimate the \bar{F}_0 distribution using

$$\hat{\bar{F}}_0(t) = \prod_{\{j: t_j \leq t\}} \left[1 - \Delta \hat{\Lambda}_0(t_j) \right]. \quad (9)$$

Step 2. Generate $Z_1^*, Z_2^*, \dots, Z_n^*$ according to a $Gamma(\hat{\alpha}, \hat{\alpha})$

For $i = 1, \dots, n$, a bootstrap sample is generated as follows:

Step 3. Take $\tau_i^* = \tau_i$;

Step 4. From $\hat{F}_0^{Z_i^*}$, continue generating an i.i.d sequence of T_{ij}^* 's until

K_i^* where

$$\sum_{j=1}^{K_i^*} T_{ij}^* \leq \tau_i^* < \sum_{j=1}^{K_i^*+1} T_{ij}^*.$$

Step 5. The bootstrap sample for the i th unit is

$$(K_i^*, \tau_i^*, T_{i1}^*, T_{i2}^*, \dots, T_{iK_i^*}^*, \tau_i^* - S_{iK_i^*}^*)$$

where $S_{ij}^* = \sum_{l=1}^{K_i^*} T_{il}^*$.

Step 6. For this bootstrap sample, compute FRMLE $\tilde{\tilde{F}}$ of \bar{F} , and compute the associated median estimate.

Repeat Steps 3–6 a total of B times. The B estimates of the median provide data for obtaining the bootstrap estimates of the sampling distribution, standard error, and for constructing a bootstrap confidence interval.

Plan VII: (Semiparametric bootstrap and bootstrapping τ_i^* 's from G_n)

This scheme is analogous to that of Plan VI, except that for each bootstrap sample, $\tau_i^*, i = 1, 2, \dots, n$, is an i.i.d. sample from the empirical distribution G_n .

4. Simulation Results

To assess the finite-sample performance of the proposed bootstrap schemes a simulation was performed. The data were generated under two scenarios: i.i.d. and gamma frailty models. To simulate the samples under the i.i.d. model, we first generate the monitoring time of each subject, τ_i , using

$G(t|\nu) = 1 - \exp(-t/\nu)$, and then we simulate the interoccurrence times, T_{ij} , through $F(t|\theta) = 1 - \exp(-t/\theta)$. To simulate the samples under a gamma frailty model we also generate the monitoring times using the same G distribution and $F_0(t|\theta) = 1 - \exp(-t/\theta)$.

For each sample, median survival time has been estimated as we have described for each of the resampling schemes. The true median survival time under the i.i.d. model is $-\theta \log(0.5)$ and under the gamma frailty model is

$$\frac{\theta\alpha(1 - 0.5^{1/\alpha})}{0.5^{1/\alpha}}.$$

We have simulated 2,000 samples and 500 bootstrap replicates ($B=500$). For each sample, the mean square error (MSE) and the 95% bootstrap percentile confidence interval (BPCI) have been calculated. In addition, for each BPCI the coverage percentage was estimated by the proportion of times the BPCI covered the true median survival time in the 2,000 samples. Mean, median, and variance of the length of the BPCI bootstrap intervals have also been calculated. Samples were generated using $n \in \{15, 50, 80\}$, $\theta \in \{1/3, 1/6\}$ and $\nu = 1$, and for the correlated case $\alpha \in \{6, 2\}$. The simulation was carried out with Fortran90. DRNUN subroutine from numerical libraries has been used as a random number generator. Tables obtained are reported for $\theta=1/3$ because results for $\theta=1/6$ show similar patterns.

[Table 1 about here.]

The results of the simulation are summarized in Table 1 and 2. Table 1 gives the results for the i.i.d. model except for the plans VI and VII, because the results for these schemes showed poor coverages (less than 80%) and large

biases (around 30% of the MSE). Table 2 gives the results for the correlated case except for the plans I, II and III, since these plans also present large biases (around 20%) and poor coverages (less than 80%). Figure 1 shows the observed distribution of the median survival time under an i.i.d. model and under a gamma frailty model, respectively.

In all simulations, as the sample size increases we obtain better coverage, less bias and less MSE, as is intuitively expected. From Table 1 we see that, in terms of MSE, the best schemes for the i.i.d. case are plans I, II and III. However, plan I has a poorer coverage than both plans II and III. Regarding the length of the BPCI, the three plans show similar average size, but both plans II and III have the smallest variance. These conclusions are the same for all sample sizes and for both values of θ . When we examine the observed distribution of the median survival under the i.i.d. model (Figure 1, bottom panels), we immediately notice that plans I and III have less variance than plan V. We can also see that the three plans obtain a sample distribution centered at the true median survival. Similar results are obtained for sample sizes set equal to 15 and 80.

[Table 2 about here.]

From Table 2 we see that the best schemes for the correlated case in terms of MSE are both semiparametric bootstrap schemes (plans VI and VII). These plans have also the shortest BPCIs and smallest variances. Evidently, the performance of all plans degrades as the level of association among the within-unit interoccurrence times increases. These conclusions are the same for all sample sizes and for both values of θ . Figure 1 (top panels) shows the

observed distribution of the median survival under a gamma frailty model. Examining these graphs, we see that resampling plan III outperforms plan V in in the i.i.d. model, whereas plan VII is best under the gamma frailty model. The performance of the resampling plan using the WC estimator seems intermediate between those based on the PSH and the FRMLE under the i.i.d. and the gamma frailty model, so in a sense this scheme may provide a robust procedure when uncertain about the model that generated the data. And this robustness property was the intent of Wang and Chang's (1999) proposing this estimator.

[Figure 1 about here.]

5. Software developed

González et al. (2002) have created an R package which calculates WC and PSH estimators called *survrec*. We have also written a function which is now included in this package which implements the bootstrap plans mentioned above (*survdiff* function). Version 1.1-1 of *survrec* which is available at <http://www.r-project.org/> contains this function. This function allows the calculation of the normal, studentized, percentile, and bias-corrected accelerated (BCa) confidence intervals. For this, we use the *boot* package included in the R project.

6. Application to a Hospital Readmission Data

The median survival time of data from a study concerning hospital readmissions of patients with colorectal cancer have been compared. To do so, we will use some of the bootstrap plans discussed in the preceding sections. The aim of the analyzed data is to investigate whether there are differences regarding the time of the recurrent hospitalizations due to social-demographic

or clinical outcomes. Four hundred and three patients with colon and rectum cancer have been included in the study. Information about their sex (male or female), age (≤ 60 , 60-74 or ≥ 75), and tumoral stage using Dukes classification (A-B, C, or D) have been recorded. All patients included in the study have been operated between January 1996 and December 1998. For each patient, we have considered this date as the beginning of the observational period. All patients were followed until June 2002. Consequently, the length of the monitoring period can differ for each patient, depending on its surgery date. The first interoccurrence time has been considered as the time between the surgical intervention and the first hospitalization related to cancer. The following interoccurrence times have been considered as the difference between the last hospitalization and the current one. Only readmissions related to cancer have been considered. This data can be obtained upon request from the first author.

[Table 3 about here.]

Table 3 shows hospital readmission distribution for patients included in the study. We can observe that most of the patients have none or one readmission and only about five percent of subjects have more than 5 readmissions. We can also see that male patients have more readmissions than women, the number of hospitalizations decreases when age increases, and when the tumoral stage becomes more severe, the number of readmissions increases.

We started our analysis of this data set by first employing the different bootstrap schemes for the sex variable. Table 4 shows a comparison among males and females. The agreement with the simulation results can

be seen. Plans II and III show the narrowest bootstrapped confidence intervals, followed by plan I. Both plans IV and V show the longest bootstrapped confidence intervals. If we can accept an i.i.d. model, plans II and III are the most appropriate for making decisions. In our example, under this assumption we could conclude that there are differences between the median of readmission times for males and females because confidence intervals do not overlap. We would obtain the same conclusion regarding plan I. However, we will conclude that no sex differences exists if we utilized either plans IV or V. Because of this differing conclusions, it is imperative that we determine if the i.i.d. model is viable.

[Table 4 about here.]

Peña et al. (2001) suggested that since formal statistical methods for checking this i.i.d. assumption are not yet available, a graphical method may be employed to assess the viability of the i.i.d. model by comparing the agreement among the PSH, WC, and FRMLE estimators. The idea is if the model is viable, then these three estimators should not differ too much from each other. Employing this idea we estimated the distribution function of readmission using WC, PSH, and FRMLE estimators in order to compare them. The resulting estimates of the readmission time distribution are presented in Figure 2. We have displayed the estimates of the distribution function instead of the survival function because in this study, the investigator is interested in analyzing the probability of readmission instead of the probability of not visiting the hospital. A considerable difference between these three estimates is obviously evident. The difference is clear between PSH and both WC and

FRMLE estimators. Thus, basing on the idea of Peña, Strawderman and Hollander's idea, we may conclude that the i.i.d. model is not appropriate for this readmission data set. In making practical conclusions, it behooves therefore to use the inference obtained from the gamma frailty model.

[Figure 2 about here.]

Table 5 shows the comparison among the variables analyzed using both FRMLE and WC estimators. We observe that the median survival time of readmission for patients diagnosed with stage D is smaller than for patients diagnosed with stage A-B or C, since in both cases the BPCI do not overlap. However, there are no differences in median time to readmission across gender, nor with respect to the age of the patients. The results using FRMLE or WC estimators agree in this context.

[Table 5 about here.]

7. Concluding Remarks

In this paper we have studied several bootstrapping schemes to estimate the sampling distribution of median survival time estimators in the presence of recurrent event data and in consideration of the sum-quota data accrual which induces informative stopping and censoring. We proposed several resampling plans under the i.i.d. model and a correlated interoccurrence times model. From this study, we conclude that the best bootstrapping scheme to estimate the median survival sample distribution under an i.i.d. model are plans I, II or III. For a correlated interoccurrence times models, both semi-parametric plans (VI and VII) are the best ones. Plan IV, which is anchored in using the Wang and Chang (1999) estimator of the inter-event survivor

function appears to offer a robust procedure when uncertain about the model that generated the data. Based on the simulation studies, it appears that bootstrapping from the empirical distribution of the monitoring times do not provide improvements. The results of this study provides impetus to further study resampling schemes in this recurrent event setting from a theoretical context, and this will be attempted in future research.

ACKNOWLEDGEMENTS

To Centre de Supercomputació de Catalunya (CESCA) for providing its resources to carry out the simulations. The second author's research was partially supported by US NSF Grant DMS 0102870.

REFERENCES

- Andersen, P. and Gill, R. (1982). Cox's regression model for counting processes: a large sample study. *Annals of Statistics* **10**, 1100–1120.
- Andersen, P. K., O, B., Gill, D, R. and Keilding, N. (1993). *Statistical models based on counting process*. Springer-Verlag, New York.
- Gill, R. (1981). Testing with replacement and the product-limit estimator. *Annals of Statistics* **9**, 853–860.
- González, J., Peña, E. and Strawderman, R. (2002). *The survrec Package*. The Comprehensive R Archive Network:, <http://cran.r-project.org>.
- McClellan, S. and Devine, C. (1995). A nonparametric maximum likelihood estimator for incomplete renewal data. *Biometrika* **85**, 605–618.
- Murphy, S. (1995). Asymptotic theory for the frailty model. *The Annals of Statistics* **23**, 182–198.

- Peña, E., Strawderman, R. and Hollander, M. (2001). Nonparametric estimation with recurrent event data. *Journal of the American Statistical Association* **96**, 1299–1315.
- Prentice, R., Williams, B. and Peterson, A. (1981). On the regression analysis of multivariate failure time data. *Biometrika* **68**, 373–379.
- Soon, G. and Woodroffe, M. (1996). Nonparametric estimation and consistency for renewal processes. *Journal of Statistical Planning and Inference* **53**, 171–195.
- Vardi, Y. (1982a). Nonparametric estimation in the presence of length bias. *The Annals of Statistics* **10**, 616–620.
- Vardi, Y. (1982b). Nonparametric estimation in renewal processes. *The Annals of Statistics* **10**, 772–785.
- Wang, M. and Chang, S. (1999). Nonparametric estimation of a recurrent survival function. *Journal of the American Statistical Association* **94**, 146–153.
- Wei, L., Lin, D. and Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *Journal of the American Statistical Association* **84**, 1065–1073.

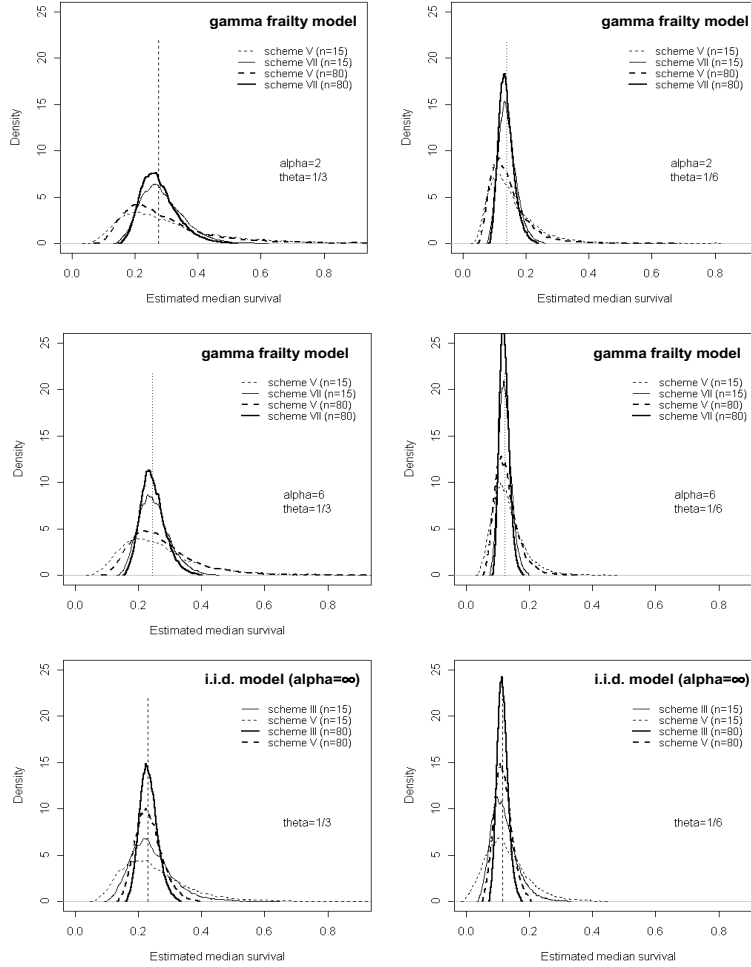


Figure 1. Observed distribution of the median survival for a i.i.d. model and a gamma frailty model in 1,000,000 replications, for selected bootstrap plans. Each panel shows the observed distribution for all combinations of θ and α that we have simulated. Vertical lines represent the true median survival time.

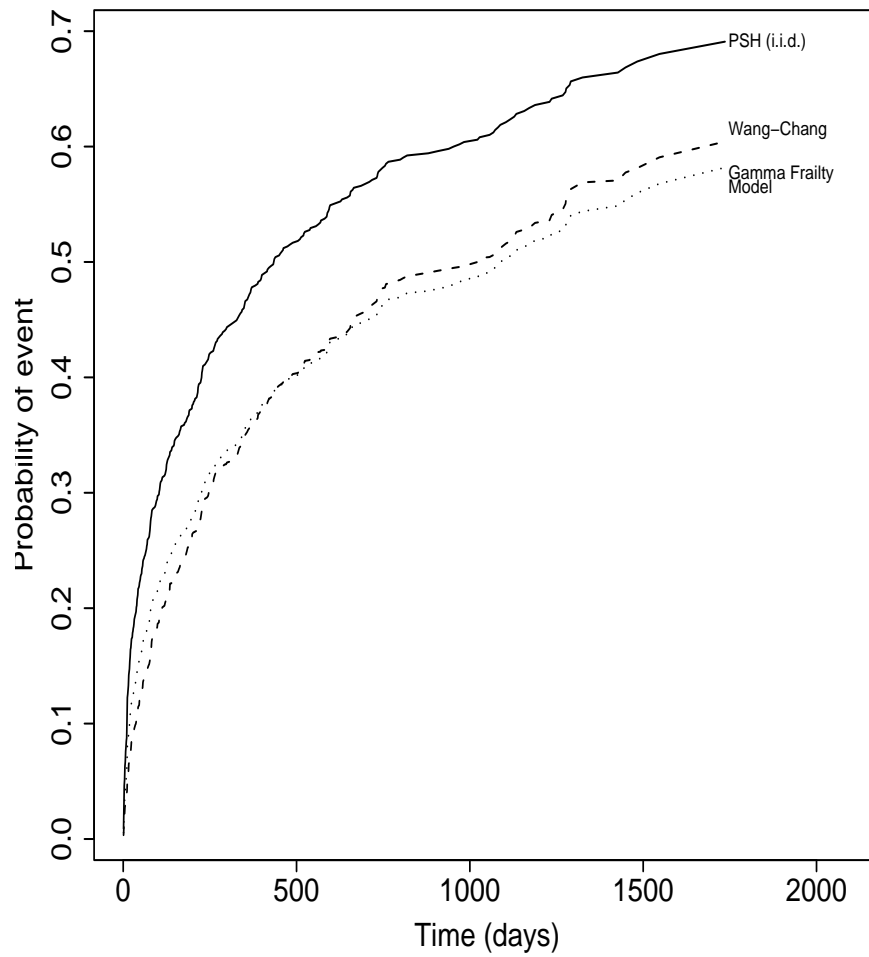


Figure 2. Plots of the three distribution function estimators (PSH,WC and FRMLE) for the hospitality readmission data set.

	95% bootstrap percentile confidence interval					
	MSE ($\times 10^6$)	% due to Bias	% Emp. Cov.	Length		Var. ($\times 10^6$)
				Mean	Median	
n=15						
Plan I	2,836	8.1	88.4	0.19	0.17	10,625
Plan II	2,914	11.3	94.3	0.21	0.19	10,760
Plan III	2,916	11.9	95.2	0.22	0.20	11,748
Plan IV	7,037	10.3	93.6	0.32	0.28	33,831
Plan V	6,879	10.4	94.2	0.32	0.28	35,398
n=50						
Plan I	667	3.1	93.3	0.10	0.10	772
Plan II	662	3.7	94.6	0.10	0.10	653
Plan III	662	3.7	94.8	0.10	0.10	668
Plan IV	1418	3.2	94.9	0.15	0.15	2050
Plan V	1425	3.2	94.5	0.15	0.15	2026
n=80						
Plan I	391	2.1	94.1	0.08	0.08	357
Plan II	385	2.5	95.4	0.08	0.08	293
Plan III	387	2.6	95.2	0.08	0.08	290
Plan IV	847	1.9	95.3	0.12	0.12	903
Plan V	847	1.9	95.4	0.12	0.12	941

Table 1

Simulation results for 2,000 samples and 500 bootstrap replicates under the i.i.d. model. Mean square error (MSE) ($\times 10^6$) and proportion of MSE due to bias. Coverage and mean, median and variance of the length ($\times 10^6$) of 95% bootstrap percentile confidence intervals. Results for the first five bootstrap schemes, varying sample sizes, $\theta = 1/3$ and $\nu=1$.

		95% bootstrap confidence interval					
		MSE ($\times 10^6$)	% due to Bias	% Emp. Cov.	Length		Var. ($\times 10^6$)
					Mean	Median	
$\alpha=2$							
n=15	Plan IV	25,760	15.2	93.4	0.57	0.45	169,858
	Plan V	23,622	15.6	93.8	0.58	0.46	169,358
	Plan VI	18,764	7.2	92.1	0.45	0.34	131,154
	Plan VII	18,858	7.3	92.3	0.45	0.34	135,060
n=50	Plan IV	4,569	5.4	94.2	0.26	0.24	15,422
	Plan V	4,569	5.5	94.2	0.26	0.23	16,193
	Plan VI	2,582	0.1	93.8	0.20	0.19	5,187
	Plan VII	2,563	0.1	94.1	0.20	0.19	5,262
n=80	Plan IV	2,653	5.6	94.8	0.20	0.19	4,262
	Plan V	2,679	5.8	95.1	0.20	0.19	4,346
	Plan VI	1,676	0.2	94.4	0.16	0.15	2,069
	Plan VII	1,684	0.2	94.8	0.16	0.15	2,009
$\alpha=6$							
n=15	Plan IV	12,034	12.7	92.9	0.40	0.33	71,607
	Plan V	11,549	12.9	93.1	0.41	0.33	70,850
	Plan VI	9,276	16.7	92.2	0.31	0.26	38,316
	Plan VII	9,217	16.9	92.7	0.31	0.26	46,671
n=50	Plan IV	2,093	4.0	93.7	0.18	0.17	3,192
	Plan V	2,085	4.1	93.9	0.18	0.17	3,523
	Plan VI	1,274	2.0	93.7	0.14	0.13	1,537
	Plan VII	1,271	2.0	93.5	0.14	0.13	1,514
n=80	Plan IV	1,208	3.7	95.3	0.14	0.14	1,579
	Plan V	1,217	3.9	95.4	0.14	0.14	1,584
	Plan VI	727	1.0	95.3	0.11	0.10	788
	Plan VII	725	1.0	95.2	0.11	0.10	777

Table 2

Simulation results for 2,000 samples and 500 bootstrap replicates under a gamma frailty model with shape and scale parameter set equal to $\alpha=2$ and $\alpha=6$. Mean square error (MSE) ($\times 10^6$) and proportion of MSE due to bias.

Coverage and mean, median, and variance of the length ($\times 10^6$) of 95% bootstrap percentile confidence intervals. Results for the last four bootstrap plans, varying sample sizes, $\theta = 1/3$ and $\nu=1$.

	Number of Hospitalizations						mean
	0	1	2	3	4	≥ 5	
Sex							
Male	112(46.9)	57(23.8)	34(14.2)	13(5.4)	10(4.2)	13(5.4)	2.3
Female	87(53.0)	48(29.3)	11(6.7)	8(4.9)	5(3.0)	5(3.0)	1.9
Age							
<60	47(42.3)	32(28.8)	11(9.9)	7(6.3)	8(7.2)	6(5.4)	2.4
60-74	98(50.5)	44(22.7)	27(13.9)	12(6.2)	7(3.6)	6(3.1)	2.1
≥ 75	54(55.1)	29(29.6)	7(7.1)	2(2.0)	0(0.0)	6(6.1)	1.8
Dukes							
A-B	103(57.2)	43(23.9)	16(8.9)	8(4.4)	7(3.9)	3(1.7)	1.8
C	67(45.3)	40(27.0)	20(13.5)	7(4.7)	6(4.1)	8(5.4)	2.2
D	29(38.7)	22(29.3)	9(12.0)	7(3.2)	4(1.8)	8(3.7)	2.7

Table 3

Distribution of the variables analyzed in hospital readmission data set.

	Estimator	Male		Female	
		Median Survival	Bootstrap CI95%	Median Survival	Bootstrap CI95%
Plan I	PSH	343	(219,483)	748	(504,1437)
Plan II	PSH	343	(230,436)	748	(468,1288)
Plan III	PSH	343	(242,436)	748	(462,1268)
Plan IV	WC	909	(524,1230)	1222	(731,2175)
Plan V	WC	909	(523,1171)	1222	(721,2175)
Plan V	FRMLE	799	(539,1171)	1427	(755,2175)
Plan V	FRMLE	799	(539,1171)	1427	(755,2175)

Table 4

Median survival comparison among males and females using the bootstrap plans mentioned in section 3 for the hospital readmission data.

	Semiparametric (plan VII)			T_{ij}^* from WC (plan V)	
	α	Median (days)	CI95%	Median (days)	CI95%
Sex					
Male	0.99	799	(539,1171)	909	(524,1230)
Female	1.50	1427	(755,2175)	1222	(721,2175)
Age					
<60	1.22	799	(415,983)	718	(474,1134)
60-74	1.05	1230	(597,1427)	1104	(646,1547)
≥ 75	0.94	1188	(551,2175)	1188	(510,2175)
Dukes					
A-B	1.11	2175	(1188, ∞)	1736	(1188,2175)
C	1.45	1073	(450,1288)	1028	(489,1325)
D	2.19	199	(109,297)	199	(161,350)

Table 5

Median survival time and 95% bootstrap percentile confidence interval (CI95%) of readmission time for the covariate analyzed using both semiparametric bootstrap and bootstrapping T_{ij}^* from WC schemes.