Grabski Franciszek

Załęska-Fornal Agata
Navy University, Gdynia, Poland

Bootstrap methods for the censored data in empirical Bayes estimation of the reliability parameters

Keywords

bootstrap method, resampling method, estimate, bootstrap replicates.

Abstract

Bootstrap and resampling methods are the computer methods used in applied statistics. They are types of the Monte Carlo method based on the observed data. Bradley Efron described the bootstrap method in 1979 and he has written a lot about it and its generalizations since then. Here we apply these methods in an empirical Bayes estimation using bootstrap copies of the censored data to obtain an empirical prior distribution.

1. Introduction

The bootstrap is a data-based method of simulation for assessing statistical accuracy. The term bootstrap derives from the phrase *'to pull oneself up by one's bootstrap'* which can be found in the eighteenth century Adventures of Baron Munchausen by Rudolf Erich Raspe. Efron proposed the method. The main goal of the bootstrap method is a computer-based fulfilling of basic statistical ideas.

2. Bootstrap and resampling copies of the censoring data

The random variable X denotes time to failure of an element. The probability distribution of the time to failure is defined by the cumulative distribution function (*cdf*)

$$F_\theta(x) = P(X \le x) \tag{1}$$

where $\theta \in \Theta$ is true but unknown parameter. To assess this distribution we test *n* identical elements $e_1, e_2, ..., e_n$ through the times $y_1, y_2, ..., y_n$ correspondingly. Suppose, that the numbers $x_1, x_2, ..., x_n$ are the times to failures of the elements mentioned above. A vector $\mathbf{x}_n = (x_1, x_2, ..., x_n)$ of the data is assumed to be the value of the random vector $\mathbf{X}_n = (X_1, X_2, ..., X_n)$, where random variables $X_1, X_2, ..., X_n$ are mutually independent and identically distributed (i.i.d.). That random vector is a sample from the distribution $F_\theta(\cdot)$. A vector $y_n = (y_1, y_2, ..., y_n)$ of the testing times of elements (times of the observations, censoring points) we can treat as the value of the random vector $\mathbf{Y}_n = (Y_1, Y_2, ..., Y_n)$. We assume that $Y_1, Y_2, ..., Y_n$ are mutually independent random variables and they are also independent of X's. Probability distributions of the random variables $Y_1, Y_2, ..., Y_n$ are defined by *cdf*

$$G_i(y_i) = P(Y_i \le y), \quad i = 1,2,...,n \tag{2}$$

Those functions do not depend on parameter $\theta \in \Theta$. In many cases these functions are defined

$$G_i(y) = \begin{cases} 0 & \text{for} \quad y < y_i \\ 1 & \text{for} \quad y \ge y_i \end{cases} \qquad y_i \in [0, \infty] \ .$$

It means that the quantities of $Y_1, Y_2, ..., Y_n$ are determined.
The observations are described by the random variables

$$U_j = \min(X_j, Y_j), \ j = 1, ..., n \tag{3}$$

and

$$\Delta_j = \begin{cases} 1 \text{ for } X_j \le Y_j \\ 0 \text{ for } X_j > Y_j \end{cases} \qquad (4)$$

The sufficient statistic describing observations can be written as the vector $\mathbf{Z}_n = ((U_1, \Delta_1), ..., (U_n, \Delta_n))$, The value of that random vector is the vector $\mathbf{z}_n = ((u_1, \delta_1), ..., (u_n, \delta_n))$, which allows to obtain the vector $\mathbf{z}_{(n)} = (z_{(1)}, z_{(2)}, ..z_{(k)}, z_{(k+1)}, ..., z_{(n)})$, where $z_{(1)}, z_{(2)}, ..., z_{(k)}$ are the instants of the elements failure and $z_{(k+1)}, z_{(k+2)}, ..., z_{(n)}$ are the times observations of the working elements.

Suppose that we are able to estimate a parameter $\theta \in \Theta$ by using estimator $\hat{\boldsymbol{\theta}}_n = T(\mathbf{Z}_n)$ (or $\hat{\boldsymbol{\theta}}_n = \tilde{T}(\mathbf{Z}_{(n)})$). The numbers $\hat{\theta}_n = T(\mathbf{z}_n)$ (or $\hat{\theta}_n = \tilde{T}(\mathbf{z}_{(n)})$) are their values. After that we can use the distribution $F_{\hat{\theta}_n}(\cdot)$ to simulate so-called *bootstrap copies*

$$\mathbf{z}_{(n)}^{*(b)} = (z_{(1)}^{*(b)}, z_{(2)}^{*(b)}, ..., z_{(n)}^{*(b)}), \; b = 1,2,...., B$$

of data $\mathbf{z}_{(n)} = (z_{(1)}, z_{(2)}, ..., z_{(n)})$. The bootstrap copies of data are the values of the random vectors $\mathbf{Z}_{(n)}^{*(b)} = (Z_{(1)}^{*(b)}, Z_{(2)}^{*(b)}, ..., Z_{(n)}^{*(b)}), \; b = 1,2,...., B$ that are called the *bootstrap samples*. The function $F_{\hat{\theta}_n^b}(\cdot)$ is a cumulative probability distribution of the independent random variables $Z_1^{*(b)}, Z_2^{*(b)}, ..., Z_n^{*(b)}$.

If we have a vector of observation $\mathbf{z}_{(n)} = (z_{(1)}, z_{(2)}, ..., z_{(n)})$ of size n, we can define the empirical cumulative distribution function $\hat{F}$ as

$$\hat{F}(z; \mathbf{z}_{(n)}) = \frac{\#\{z_{(i)} : z_{(i)} \le z\}}{n}$$

that is equivalent to the discrete distribution

$$\hat{p}_k = \frac{n_k}{n}, \quad k = 1,2,..., l,$$

where $n_k = \#\{i : z_{(i)} = z_{(k)}\}$.

This distribution can be expressed as a vector of frequencies $\hat{\mathbf{p}} = (\hat{p}_1, \hat{p}_2, ..., \hat{p}_l)$.

Vectors of the data

$$\mathbf{z}_n^{\circ(r)} = (z_1^{\circ(r)}, z_2^{\circ(r)}, ..., z_n^{\circ(r)}), \; r = 1,2,...., R$$

coming from distribution $\hat{F}(z; \mathbf{z}_{(n)})$ are said to be *resampling copies* of the data

$$\mathbf{z}_{(n)} = (z_{(1)}, z_{(2)}, ..., z_{(n)}).$$

In other words a resampling copy of the data $\mathbf{z}_n^{\circ(r)} = (z_1^{\circ(r)}, z_2^{\circ(r)}, ..., z_n^{\circ(r)})$ is generated by randomly sampling $n-$times with replacement from the original data points $\mathbf{z}_{(n)} = (z_{(1)}, z_{(2)}, ..., z_{(n)})$. The randomly sampling means the random choice of an element from among $z_{(1)}, z_{(2)}, ..., z_{(n)}$ in each of n drawings. The resampling copy of the data is composed of the elements of the original sample, some of them can be taken zero times, some of them can be taken ones or twice etc. Notice that in $\mathbf{z}_n^{\circ(r)} = (z_1^{\circ(r)}, z_2^{\circ(r)}, ..., z_n^{\circ(r)})$ - the resampling copy, the elements are repeated as a rule.

The typical number of the bootstrap B or resampling copies of the data, range from 50 to 1000.

3. Bootstrap estimators

Let $\mathbf{Z}_n^* = (Z_1^*, Z_2^*, ..., Z_n^*)$ be a bootstrap sample for the given vector of data $\mathbf{z}_n = (z_1, z_2, ..., z_n)$.

A random variable $\boldsymbol{\theta}_n^* = T(\mathbf{Z}_n^*)$ is said to be a bootstrap estimator of the parameter θ.

The distribution of the statistics $\boldsymbol{\theta}_n^* - \hat{\boldsymbol{\theta}}_n$ for the bootstrap sample with the fixed values data is close to the distribution of the statistics $\hat{\boldsymbol{\theta}}_n - \theta$.

From that rule it follows that the shapes of the distributions of the statistics $\boldsymbol{\theta}_n^*$, $\hat{\boldsymbol{\theta}}_n$ are similar.

To obtain empirical distribution of the random variable $\boldsymbol{\theta}_n^*$ we have to simulate bootstrap copies

$$\mathbf{z}_n^{*(b)} = (z_1^{*(b)}, z_2^{*(b)}, ..., z_n^{*(b)}), \; b = 1,2,...., B$$

of data $\mathbf{z}_n = (z_1, z_2, ..., z_n)$. After that we calculate the values of statistics

$$\theta_n^{*(b)} = T(\mathbf{z}_n^{*(b)}), \; b = 1,2,...., B$$

We can use a nonparametric kernel estimator to obtain the estimate of probability density of the bootstrap estimator $\boldsymbol{\theta}_n^*$. The value of this estimator with Gaussian kernel is given by

$$\hat{g}(\vartheta) = \frac{1}{B\,h}\sum_{b=1}^{B} K\left(\frac{\vartheta - \theta_n^{*(b)}}{h}\right)$$

where

$$K(\vartheta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\vartheta^2}{2}}, \ \vartheta \in (-\infty, \infty),$$

and

$h = 1.06\, s\, B^{-0.2}$, s - standard deviation of $\theta_n^{*(b)},\ b = 1,2,....,B$.

4. The bootstrap estimate of standard error

$$\mathbf{z}_n^{*(b)} = (z_1^{*(b)}, z_2^{*(b)},..., z_n^{*(b)}),\ b = 1,2,....,B$$

are the bootstrap replication of the statistics values

$$\theta_n^{*(b)} = T(z_n^{*(b)}),\ b = 1,2,....,B \tag{5}$$

and they correspond to the bootstrap censoring data. The bootstrap estimate of the standard error of $\hat{\theta}$ is defined by the following formula

$$se_{\hat{\theta}^*} = \sqrt{\frac{\sum_{b=1}^{B}\left(\hat{\theta}^{*(b)} - \overline{\theta}^*\right)^2}{B-1}}, \tag{6}$$

where $\overline{\theta}^* = \frac{\sum_{b=1}^{B}\hat{\theta}^{*(i)}}{B}$.

The bootstrap algorithm for estimating standard errors is as follows:

- Get B independent bootstrap samples $\mathbf{z}_n^{*(b)} = (z_1^{*(b)}, z_2^{*(b)},..., z_n^{*(b)}),\ b = 1,2,....,B$ (for estimating a standard error, the number of B should be in the range 30-200).
- Compute the bootstrap replication correspond each bootstrap sample, $\theta_n^{*(b)} = T(\mathbf{z}_n^{*(b)}),\ b = 1,2,....,\ B$.
- Compute the standard error $se_{\hat{\theta}^*}$ by the sample standard deviation of B replications according to (6).

5. Empirical Bayes estimation

The recent work deal with empirical Bayes estimation has been stimulated by the work of Robbins (1955). It is well known that the value of Bayes estimator $\hat{\theta}_B$ of the parameter θ under the squared-loss function is an expectation in posterior distribution. If $\hat{\theta}$ is a value of sufficient statistics for parameter θ, than the value of Bayes estimator $\hat{\theta}_B$ of the parameter θ is

$$\hat{\theta}_B = E(\theta \mid \hat{\theta}) = \frac{\int_{\Theta} \theta\, \tilde{\mathbf{f}}(\hat{\theta} \mid \theta) g(\theta) d\nu(\theta)}{\int_{\Theta} \tilde{\mathbf{f}}(\hat{\theta} \mid \theta) g(\theta) d\nu(\theta)} \tag{7}$$

where ν denotes a discrete counting measure or Lebesgue measure and $g(\theta)$ is a prior density function of the parameter θ with respect to the measure ν.

We suppose that a prior density of mentioned above parameter is unknown. In classical empirical Bayesian procedure a prior distribution is assessed from the *past data*. Very often the only data we have is the small sample $\mathbf{z} = (z_1, z_2,..., z_n)$. In those cases instead of the past data, we can use the vectors $\mathbf{z}_n^{*(b)} = (z_1^{*(b)}, z_2^{*(b)},..., z_n^{*(b)}),\ b = 1,2,....,B$, that are values of the *bootstrap samples* corresponding to an unknown distribution $F_\theta(\cdot)$ of a random variable X, which denotes (for example) a time to failure. The bootstrap copies for the censored data are generated from the distribution $F_{\hat{\theta}}(\cdot)$, where $\hat{\theta} = T(\mathbf{z}_{(n)})$. To estimate the unknown parameter θ we have to calculate the values of the bootstrap statistics $\theta^{*(b)} = T(z_n^{*(b)}),\ b = 1,2,....,B$ of that one.

As a prior density we propose a discrete density function

$$g(\theta) = \frac{m_i}{m}\delta(\theta, \theta^{*(i)}), \quad i \in \{j_1, j_2,..., j_w\} \subseteq \{1,..., B\} \tag{8}$$

where

$$m_i = \#\{k : \ \theta^{*(k)} = \theta^{*(i)}\}$$

denotes number observations equal to $\theta^{*(i)}$.

$$\delta(\theta, \theta^{*(i)}) = \begin{cases} 1 & \text{for} \quad \theta = \theta^{*(i)} \\ 0 & \text{for} \quad \theta \neq \theta^{*(i)} \end{cases}$$

and

$$m = \sum_{i=1}^{w} m_{j_i} = B$$

From (7), for counting measure ν and for the density function defined by (8) we obtain

$$\hat{\theta}_B = E(\theta \,|\, \hat{\theta}) = \frac{\sum_{i=1}^{w} m_i \theta^{*(i)} \tilde{f}(\hat{\theta} \,|\, \theta^{*(i)})}{\sum_{i=1}^{w} m_i \tilde{f}(\hat{\theta} \,|\, \theta^{*(i)})} = \frac{\sum_{i=1}^{B} \theta^{*(i)} \tilde{f}(\hat{\theta} \,|\, \theta^{*(i)})}{\sum_{i=1}^{B} \tilde{f}(\hat{\theta} \,|\, \theta^{*(i)})}. \quad (9)$$

Let

$$\mathbf{f}_\theta(\mathbf{z}_{(n)}^{*(b)}) = l(\mathbf{z}_{(n)}^{*(b)}; \theta)$$

be a likelihood function for the bootstrap sample

$$\mathbf{z}_{(n)}^{*(b)} = (z_{(1)}^{*(b)}, z_{(2)}^{*(b)}, \ldots, z_{(n)}^{*(b)})$$

with unknown parameter $\theta \in \Theta$. The function is defined by the formula

$$l(\mathbf{z}_{(n)}^{*(b)}, \theta) = \prod_{i=1}^{k_b} f_\theta(z_{(i)}^{*(b)}) \prod_{i=k_b+1}^{n} [1 - F_\theta(z_{(i)}^{*(b)})]. \quad (10)$$

Notice that a prior distribution is constructed on the basis on the bootstrap samples. Since, a value of bootstrap empirical Bayes estimator has the form of (9).

6. Examples

Example 1.

Suppose that we wish to estimate a failure rate $\theta = \lambda$ in the exponential distribution given by *pdf*

$$f_\theta(x) = \lambda e^{-\lambda x}, \quad x \geq 0, \quad \lambda > 0 \,. \quad (11)$$

Assume that we have data, which is the vector

$$\mathbf{z}_{(n)} = (z_{(1)}, z_{(2)}, .. z_{(k)}, z_{(k+1)}, \ldots, z_{(n)}),$$

where $z_{(1)}, z_{(2)}, \ldots, z_{(k)}$ are times to failure of the tested elements and $z_{(k+1)}, z_{(k+2)}, \ldots, z_{(n)}$ are times of the working elements observations. In that case a likelihood function is

$$l(\mathbf{z}_{(n)}, \lambda) = \prod_{i=1}^{k} f_\theta(z_{(i)}) \prod_{i=k+1}^{n} [1 - F_\theta(z_{(i)})] = \prod_{i=1}^{k} \lambda e^{-\lambda\, z(i)} \prod_{i=k+1}^{n} [e^{-\lambda\, z(i)}] = \lambda^k e^{-\lambda \sum_{i=1}^{n} z(i)}. \quad (12)$$

The number

$$\tau = \sum_{i=1}^{n} z_{(i)} \quad (13)$$

is the value of some sufficient statistics for the unknown parameter λ. By substitution we obtain the likelihood function

$$l(\tau, \lambda) = \lambda^k e^{-\lambda \tau},$$

which depends on τ. To find the value of the maximum likelihood estimator we have to solve an equation

$$\frac{\partial \ln l(\tau, \lambda)}{\partial \lambda} = 0.$$

The solution of it is

$$\hat{\lambda} = \frac{k}{\tau} = \frac{k}{\sum_{i=1}^{n} z_{(i)}}. \quad (14)$$

The same way, using formula (7) for the bootstrap samples $\mathbf{z}_{(n)}^{*(b)} = (z_{(1)}^{*(b)}, z_{(2)}^{*(b)}, \ldots, z_{(n)}^{*(b)}),\ b = 1,2,\ldots.,B$ we obtain the values of the maximum likelihood estimator of λ

$$\lambda^{*(b)} = \frac{k^{*(b)}}{\tau^{*(b)}} = \frac{k^{*(b)}}{\sum_{i=1}^{n} z_i^{*(b)}}, \quad b = 1,2,\ldots.,B$$

The function (9) in this case is given by the formula

$$\hat{\lambda}_B = E(\lambda \,|\, \hat{\lambda}) = \frac{\sum_{i=1}^{w} m_i \lambda^{*(i)} \tilde{f}(\hat{\lambda} \,|\, \lambda^{*(i)})}{\sum_{i=1}^{w} m_i \tilde{f}(\hat{\lambda} \,|\, \lambda^{*(i)})},$$

where

$$\tilde{f}(\hat{\lambda} \,|\, \lambda^{*(i)}) = (\lambda^{*(i)})^k e^{-\frac{k\lambda^{*(i)}}{\hat{\lambda}}}.$$

Finally we obtain

$$\hat{\lambda}_B = \frac{\sum_{i=1}^{w} m_i \lambda^{*(i)} (\lambda^{*(i)})^k e^{-\frac{k\lambda^{*(i)}}{\hat{\lambda}}}}{\sum_{i=1}^{w} m_i (\lambda^{*(i)})^k e^{-\frac{k\lambda^{*(i)}}{\hat{\lambda}}}} = \frac{\sum_{j=1}^{B} (\lambda^{*(j)})^{k+1} e^{-\frac{k\lambda^{*(j)}}{\hat{\lambda}}}}{\sum_{j=1}^{B} (\lambda^{*(j)})^k e^{-\frac{k\lambda^{*(j)}}{\hat{\lambda}}}}$$

where

$$\hat{\lambda} = \frac{k}{\sum_{i=1}^{n} z_{(i)}}, \quad \lambda^{*(b)} = \frac{k^{*(b)}}{\sum_{i=1}^{n} z_i^{*(b)}}, \quad b = 1,2,....,B$$

By repetition we can obtain a sequence of values of a Bayes estimator that we can use to construct its empirical distribution.

Example 2.

We wish to estimate a value of an exponential reliability function

$$R_\theta(x) = e^{-\lambda x},\ x \ge 0,\ \lambda > 0,\ \theta = \lambda. \tag{15}$$

At a fixed moment x_0 the number

$$r = R_\theta(x_0) = e^{-\lambda x_0}$$

is a value of the reliability function. Hence

$$\lambda = \frac{\ln r}{x_0}. \tag{16}$$

There is a given vector

$$\mathbf{z}_{(n)} = (z_{(1)}, z_{(2)}, ..z_{(k)}, z_{(k+1)}, ..., z_{(n)})$$

the coordinates of which have the same meaning as in Example 1. Let τ is described by (13). A likelihood function of the parameter λ for $\mathbf{z}_n$ is

$$l(\tau, \lambda) = \lambda^k e^{-\lambda \tau}.$$

Substituting the value of λ and $r = e^{\ln r}$ we get the form of the likelihood function

$$l(\tau, \lambda) = \mathbf{f}(\tau \mid r) = \left(-\frac{\ln r}{x_0}\right)^k e^{\left(\frac{\ln r}{x_0}\right)\tau} = \left(-\frac{\ln r}{x_0}\right) r^{\frac{\tau}{x_0}}. \tag{17}$$

The likelihood equation

$$\frac{\partial \ln l(\tau, \lambda)}{\partial \lambda} = 0$$

is carried out to the following form

$$\frac{k}{r \ln r} + \frac{\tau}{r x_0} = 0.$$

A root of the equation is a value of the maximum likelihood estimate of r and it has a form of

$$\hat{r} = e^{-\left(\frac{k x_0}{\tau}\right)}. \tag{18}$$

Using the bootstrap samples

$$\mathbf{z}_{(n)}^{*(b)} = (z_{(1)}^{*(b)}, z_{(2)}^{*(b)}, ..., z_{(n)}^{*(b)}),\ b = 1,2,....,B$$

we obtain the values of the maximum likelihood estimator of r and it is defined by

$$r^{*(b)} = e^{\left(-\frac{k^{*(b)} x_0}{\sum_{i=1}^{n} z_i^{*(b)}}\right)}, \quad b = 1,2,..., B.$$

As

$$\tilde{f}(\hat{r} \mid r^{*(i)}) = \ln(r^{*(i)}, \tau) = \left(-\frac{\ln r^{*(i)}}{x_0}\right)^k (r^{*(i)})^{\frac{\tau}{x_0}},$$

then the value of the Bayes empirical estimate of r computed on the basis on

$$\hat{r}_B = E(r \mid \hat{r}) = \frac{\sum_{i=1}^{w} m_i r^{*(i)} \tilde{f}(\hat{r} \mid r^{*(i)})}{\sum_{i=1}^{w} m_i \tilde{f}(\hat{r} \mid r^{*(i)})}.$$

has the following form

$$\hat{r}_B = \frac{\sum_{i=1}^{w} m_i r^{*(i)} \left(-\frac{\ln r^{*(i)}}{x_0} \right)^k (r^{*(i)})^{\frac{\tau}{x_0}}}{\sum_{i=1}^{w} m_i \left(-\frac{\ln r^{*(i)}}{x_0} \right)^k (r\ )^{\frac{\tau}{x_0}}} .$$

7. Conclusions

In that paper we present the possibility of applying the bootstrap methods in empirical Bayes estimation. The bootstrap copies of the given data are used to construct an empirical prior distribution function.

References

[1] Belyaev, Yu. K. (2001). Resampling and bootstrap methods in analysis of reliability data. *Safety & Reliability. ESREL 2001*, 1877-1882.

[2] Grabski, F. & Jadźwiński, J. (2001). *Metody bayesowskie w niezawodności i diagnostyce*. WKŁ, Warszawa 2001.

[3] Efron, B. & Tibshirani, R. J. (1993). *An introduction to the Bootstrap, Chapman & Hall*. New York, London.

[4] Koronacki, J. & Mielniczuk, J.(2001). *Statystyka dla studentów kierunków technicznych i przyrodniczych*. Wydawnictwa Naukowo-Techniczne, Warszawa 2001.

[5] Savĉuk, W.P. (1989). Bayesian methods of the statistical assessment. Nauka, Moskva (Russian).