STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **7**, June 2019, pp 370–382.
Published online in International Academic Press (www.IAPress.org)

Properties of Non-Parametric Stute Estimators

Didier Alain Njamen Njomen

Department of Mathematics and Computer Science, Faculty of Science, University of Maroua, Cameroon

Abstract In this paper, we use the linear regression model for survival data, explaining that it corresponds to an accelerated time model of lifetime, as described in Kalbfleisch and Prentice [12] and Koul and al. [15]. In this context, we adapt the jumps of the KM estimator as defined in Lopez [16] to the accelerated lifetime model. The introduction of a more restrictive hypothesis allows us to establish a strong consistency property of the Stute [24] estimator obtained by minimizing the sum of the least squares. Using the asymptotic normality of the bivariate distribution estimator proposed by Stute [26] and the Slutsky theorem, we succeed in establishing the asymptotic distribution of the Stute [24] estimator.

Keywords Censored lifetimes; Linear regression; Kaplan-Meier jump; Nonparametric estimator; Asymptotic distribution.

AMS 2010 subject classifications 62N01,62J05, 62F12, 62G20.

DOI: 10.19139/soic.v7i2.392

1. Introduction

The statistical analysis of the lifetimes studies the laws of instants of occurrence of events, based on observations of durations and possibly explanatory variables, made discretely or continuously over time. A priori, we could treat a duration variable like any continuous quantitative random variable, except that it necessarily takes a positive real value. This is not a very discriminating characteristic, since found in other themes of economic analysis, such as that of wages. The usual reference to the normal law then requires a transformation on the data, taking for example the logarithm. Thus, one of the basic laws in wage econometrics is the log-normal law, which consists in making a normality assumption on the log of the variable studied. This distribution, as we shall see, is less central in econometrics of durations. The peculiarity of time data arises from the fact that they can easily be interpreted as resulting from an underlying stochastic process, that is to say from a random path that makes an individual pass between different states. This process thus accounts for the dates of changes in the state of the individual (life And death, employment and unemployment, parenting one child or two children...). The duration of a state is then simply the difference between the start date and the end date of a state. The characteristics of this process then lead to the definition of large classes of probability laws for durations. In many fields of application we have, in addition to the observation of lifetimes, additional information suspected of influencing the durations studied. This additional information, called covariables or explanatory variables, may be different for each individual.This can be a characteristic of the individual (blood group, sex, occupational domain, age ...) or a study-dependent observation (dosage of medical treatment, type of transplant, duration of hospitalization ...). Two major objectives of the survival analysis are the evaluation of the influence of the covariates and the prediction of a survival time. Lifetime analysis is used in many fields, such as medicine, industrial reliability, economics or psychology, and the study of data from these sectors has been developing for several decades. There are many ways of modeling

*Correspondence to: Didier Alain Njamen Njomen (Emails: didiernjamen1@gmail.com / didier.njamen@univ-maroua.cm). Department of Mathematics and Computer Science, Po Box: 814, University of Maroua, Cameroon.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

survival data: in the case of censored data, the Cox model defined by its risk function

$$\lambda(t) \quad = \quad \lambda_0(t)e^{\beta\ X},$$

where $\lambda_0(t)$ is an unparametered basic risk function, and the linear regression model defined by an equality linking the response variable Y to covariables X is

$$Y \quad = \quad \beta^{'} X + \varepsilon,$$

where ε is a random error variable.

Since the founding paper of Cox [6], several works deal with the concept of analysis of survival data as well as linear regression in data censored in particular Buckley and James [5], Aalen ([1], [2]), Andersen and Gill [3], Fleming and Harrington [7], Kaplan and Meier [13], Klein and Moeschberger [14], Miller [17], Susarla and Van Ryzin [27], Kalbfleisch and Prentice [12] and Koul et al. [15].

In addition, some specific probability tools such as survival function or instantaneous risk function or cumulative risk function will play a more decisive role in the analysis than the usual probability density because they have the advantage of being interpreted very simply.

2. Probabilistic tools

2.1. Characteristic functions in lifetime analysis

The statistical analysis of the lifetimes studies the laws of instants of occurrence of events, based on observations of durations and possibly explanatory variables, made discretely or continuously over time.
Thus, we denote by T a positive random variable defined on a probabilized space (Ω, A, P) and representing a duration up to an event of interest, the origin of the times being predefined. In the medical field, this event may be the death, healing, relapse of an individual; in the economic field, loss of employment; in reliability, the moment of first breakdown. Thereafter, the duration T will be called the lifetime. We denote F its distribution function. The law of T can also be characterized by other easily interpretable functions considering T in term of life.

Definition 1
We term survival function S, the probability that the lifetime T is superior to a time t:

$$\forall t \ \in \mathbb{R},\ S(t) = \mathbb{P}(T > t) = 1 - F(t).$$

Note that if the law of T has a density f with respect to the Lebesgue measure, then we have:

$$\forall t \ \in \mathbb{R},\ \ S(t) = \int_t^{+\infty} f(t)dt \ \text{ and } \ f(t) = -S'(t)\ p.p.$$

Definition 2
The instantaneous risk function λ is the function defined for $t \in \mathbb{R}^+$ by

$$\lambda(t) = \begin{cases} \lim\limits_{h\to 0} \dfrac{1}{h}\mathbb{P}\left(t < T \leq t + h \mid T > t\right) & if\ t \text{ such that } \mathbb{P}(T > t) > 0 \\ & +\infty \text{ if not.} \end{cases}$$

The risk function can have very different forms but is necessarily positive on $\mathbb{R}$.
Suppose now that T is a continuous variable, we therefore observe that

$$\forall t \ \in \mathbb{R}^+,\ \lambda(t) = \frac{f(t)}{S(t)} = -\frac{\partial}{\partial t}\ln(S(t)),$$

by posing $c/0 = +\infty$ for all $c > 0$. The definition of λ shows that for h small enough, $h\lambda(t)$ is interpreted as the probability of occurrence of the event of interest in the interval $[t, t + h]$ knowing that this event has not yet occurred in the instant t. This function therefore reflects the evolution over time of the risk of occurrence of the event of interest.

Definition 3
We call the cumulative risk function Λ The function defined for $t \in \mathbb{R}^+$ by:

$$\Lambda(t) = \int_0^t \lambda(s)ds = -\ln(S(t)),$$

which is worth $+\infty$ when $S(t) = 0$.

From the above definitions, it can be deduced that for all $t \in \mathbb{R}^+$, we have the relation:

$$f(t) = \lambda(t) \exp\left(-\Lambda(t)\right).$$

In conclusion, the five previous functions allow us to characterize the law of T and some are deductible from the others. However, it is the interpretation of the instantaneous risk function that will most often guide the choice of a model for lifetime data.

Remark 1
We can therefore characterize the law of duration T by a function of constant instantaneous risk:

$$\forall t \in \mathbb{R}^+,\ \lambda(t) = \lambda,$$

where λ is a strictly positive constant.

This means that in a survival study, assuming that there is no wear or aging effect, the probability of occurrence of the event of interest, knowing that it is not Still occurred, does not change over time.

Example 1
The exponential law was generalized by Weibull in 1939 (Weibull [28]) to the law of the same name by introducing a new parameter, so that the risk function is as follows:

$$\forall t \in \mathbb{R}^+,\ \lambda(t) = \lambda\alpha t^{\alpha-1},$$

where λ and α are two strictly positive constants. The parameter λ is called scale parameter and α, shape parameter. Indeed, λ gives the magnitude of the risk function, and the position of α with respect to 1 defines the monotony of the risk function: if $\alpha = 1$, we find the constant risk function and therefore the exponential law; if $\alpha > 1$ (respectively $\alpha < 1$), λ is increasing (or decreasing) in time and there is therefore a phenomenon of wear, aging (respectively rejuvenation). By expressing the risk function, we obtain the following expressions for t belongs to $\mathbb{R}^+$:

$$f(t) = \lambda\alpha t^{\alpha-1}e^{-\lambda t\alpha},\ S(t) = e^{-\lambda t\alpha} \text{ and } \Lambda(t) = \lambda t^{\alpha}.$$

The Weibull law is widely used in the industrial (reliability) and biomedical (lifetime analysis) fields. Indeed, this law appeared to be the most appropriate choice of model in the description of data concerning the lifetime of manufactured components or the appearance of a tumor in the animal. Its success is also due to the fact that this law has a fairly broad spectrum, covering both the case of a function of increasing risk and that of a decreasing risk function.

3. Taking into account the covariates

In the parametric approach, the interest functions may depend on explanatory covariates that can influence survival. In addition to adjusting survival functions to different factors,this will make it possible to compare survival times (the null hypothesis will be equality of survival distributions).

Let's consider a random lifetime T and a vector of p real explanatory variables $Z = (Z_1, \ldots, Z_p)^{'}$ associated with the survival time T .

Note that these covariates may depend on time, however it is necessary to assume that the value of the covariates does not change between two measures. In order to simplify the writing, it will be assumed in the following that the covariables are fixed over time. We suppose that the covariables modify the risk functions by following a Cox proportionate risk model (other models with proportionate risk are possible). In effect, the multiplicating risk models are defined from a conditional risk function to the covariates Z written as the product of a risk function termed basic λ_0 by covariates positive function $\exp(\beta^{'} x)$:

$$\lambda(t \mid Z) = \lambda_0(t) \exp(\beta' Z),$$

where β' is the vector of the regression coefficients. The survival and density functions corresponding to these risk functions are given by

$$\begin{aligned} S(t \mid Z) &= \exp\left(-\int_0^t \lambda(u \mid Z)du\right) \\ &= \exp\left(-\int_0^t \lambda_0(u)\exp(\beta' Z)du\right) \\ &= S_0(t)^{\exp(\beta' Z)}, \end{aligned}$$

and

$$\begin{aligned} f(t \mid Z) &= -S(t \mid Z) \\ &= \lambda(t \mid Z)\exp\left(-\int_0^t \lambda(u \mid Z)du\right) \\ &= \lambda_0(u)\exp(\beta' Z) \times S_0(t)^{\exp(\beta' Z)}, \end{aligned}$$

with $S_0(t) = \exp\left(-\int_0^t \lambda_0(u)du\right)$.

3.1. Comparison of two groups

Let's consider the situation where we want to compare the survival times of two groups A and B. We introduce the following covariate:

$Z = 0$ if the individual belongs to the group $A \Longrightarrow \lambda_A(t) = \lambda_0(t)$;

$Z = 1$ if the individual belongs to the group $B \Longrightarrow \lambda_B(t) = \lambda_0(t)\exp(\beta)$.

To compare the two groups, we estimate the regression coefficient β and we test the null hypothesis $H_0 : \beta = 0$ i.e. $H_0 : \lambda_A = \lambda_B$. We can use the tests of the likelihood ratio of Wald or of the score that follows asymptotically a law of $\chi^2(1)$, under H_0.

Example 2
Consider a basic risk according to a Weibull law $W(\theta, \nu)$, so, we have respectively

$$\begin{aligned} \lambda_0(t) &= \nu\left(\frac{1}{\theta}\right)^{\nu} t^{\nu-1},\ t \geq 0 \text{ and } \theta, \nu > 0; \\ S_0(t) &= \exp\left(-\left(\frac{1}{\theta}\right)^{\nu}\right) \text{ and} \\ f_0(t) &= \nu\left(\frac{1}{\theta}\right)^{\nu} t^{\nu-1}\exp\left(-\left(\frac{1}{\theta}\right)^{\nu}\right). \end{aligned}$$

From the beginning of this section, the risk, survival and density functions in the case of covariates are defined by:

$$\begin{aligned}
\lambda(t \mid Z) &= \nu \left(\frac{1}{\theta}\right)^{\nu} t^{\nu-1} \times \exp(\beta' Z)\ t \geq 0 \text{ and } \theta, \nu > 0; \\
S(t \mid Z) &= \exp\left(-\left(\frac{1}{\theta}\right)^{\nu}\right)^{\exp(\beta' Z)} \text{ and} \\
f(t \mid Z) &= \nu \left(\frac{1}{\theta}\right) t^{\nu-1} \times \exp(\beta' Z) \times \exp\left(-\left(\frac{1}{\theta}\right)^{\nu}\right)^{\exp(\beta' Z)}
\end{aligned}$$

For $\nu = 1$; we find the exponential distribution $\varepsilon(\frac{1}{\theta})$. Thus, in the case of a risk according to an exponential law with covariates, we obtains respectively:

$$\begin{aligned}
\lambda(t \mid Z) &= \frac{1}{\theta} \times \exp(\beta' Z), \theta > 0 \\
S(t \mid Z) &= \exp\left(-\frac{1}{\theta}\right)^{\exp(\beta' Z)} \text{ and} \\
f(t \mid Z) &= \frac{1}{\theta} \exp(\beta' Z) \times \exp\left(-\frac{1}{\theta}\right)^{\exp(\beta' Z)}.
\end{aligned}$$

4. Accelerated Failure Time model

Among the regression models, accelerated life models are often regarded in terms of reliability. These models can be defined in two ways. The first representation of accelerated life models is given by the accelerated survival function:

$$S(t \mid Z) = S_0\left(te^{\beta' Z}\right),$$

where Z is a vector of covariate and β the vector of the regression coefficients.

Indeed,

$$\begin{aligned}
S(t \mid Z = z) &= \mathbb{P}(T > t \mid Z = z) \\
&= \mathbb{P}\left(\ln(T) > \ln(t) \mid Z = z\right) \\
&= \mathbb{P}\left(\varepsilon > \ln(t) - \beta' Z \mid Z = z\right) \\
&= \mathbb{P}(\exp(\varepsilon) > t \exp(-\beta' z)) \\
&= \mathbb{P}\left(T > t \exp(-\beta' z) \mid Z = 0\right) \\
&= S_0\left(t \exp(-\beta' z)\right).
\end{aligned}$$

The term $e^{\beta' Z} Z$ is an acceleration factor because a change in the covariates modifies the time scale. We can obtain the following expression of the risk function:

$$\begin{aligned}
\lambda(t \mid Z) &= \left[-\ln(S(t \mid Z))\right]' \\
&= -\frac{[S(t \mid Z)]'}{S(t \mid Z)} \\
&= \frac{-e^{\beta' Z} \times \lambda_0(te^{\beta' z}) \times S_0(te^{-\beta' z})}{S_0(te^{\beta' z})} \\
&= e^{\beta' z} \lambda_0(te^{\beta' Z}).
\end{aligned}$$

Indeed, we have the following equations:

$$S(t \mid Z) = S_0(t\, e^{\beta' z}) = \exp(-\Lambda_0(t\, e^{tbeta' z})) = \exp\left[-\int_0^t \lambda_0(ue^{\beta' Z})du\right],$$

where λ_0 and Λ_0 are the instantaneous and cumulative risk functions of T when the covariates are set equal to 0. In the same way, when we know the covariates $Z \in \mathbb{R}^p$, so we have

$$\Lambda(t \mid Z) = \Lambda_0(t \exp(-\beta' Z)).$$

The density $f(t|Z)$ can then be written as a function of the basic risk function:

$$f(t|Z) = \lambda_0(t\, e^{-\beta' X})e^{-\beta' Z} \exp\left(-\Lambda_0(t \exp(-\beta' Z))\right).$$

This model of acceleration or deceleration of time is commonly used in industry, where multiplicative time scales are common.

On the other hand, assuming that $S_0(t)$ is the survival function of the variable $\exp(\mu + \varepsilon)$, then $S_0(t) = \mathbb{P}\left(e^{\mu+\varepsilon} > t\right)$. Thus, one obtains that is the survival function of the variable X where $\log(X) = \mu - \beta' Z + \varepsilon$. Considering the change of variable $\alpha = -\beta$, we obtain the second representation by a log-linear regression model for the duration of survival

$$\log(X) = \mu\ - \beta' Z + \varepsilon,$$

where X is the survival time (not always observed because $T = \min(X, C)$) and ε is a random variable (in the case of several observations, the ε_i are i.i.d.).

Several laws are possible for varepsilon variables, for example:

1. $\varepsilon\ \sim$ Law to extreme values $(f_\epsilon(y) = \exp(y - e^y))$
2. $\varepsilon\ \sim$ log-logistic
3. $\varepsilon\ \sim$ log-normal
4. $\varepsilon\ \sim$ generalized gamma

One can safely deduce that the law of X and the parameter estimates are obtained by maximizing the likelihood.

Remark 2
It may be noted that in the case of accelerated life models, for a covariate $Z > 0$, a negative α regression coefficient results in a smaller survival time. Whereas in the Cox semi-parametric model, a negative α regression coefficient results in a lower risk of event and therefore greater survival.

The reader interested in a more detailed development on the accelerated time model can refer to Bagdonavicius and Nikulin [4]. In addition, there are many works in which the linear regression model plays a central role in modern statistics, for example: Jorgensen [11], Rao and Toutenburg [21], Searle [23], Rencher and Schaalje [22].

5. Motivation

In this section, we use the linear regression model for survival data, explaining that it corresponds to an accelerated time model. The interpretation of this type of model is given in terms of the usual functions characterizing the law. We then introduce into the general frame of survival time censored right the jumps of the Kaplan-Meier estimator Kaplan and Meier [13] and the asymptotic results obtained by Stute [24] and Stute [26].

5.1. Random censorship

We are working in a straight type I random censorship mechanism. The interested reader may refer to Klein and Moeschberger [14] for full censorship on Type II censorship, or Type I censorship on the left or by Intervals.

Consider now a random lifetime T^0. Let us introduce a random variable C independent of T^0 to value in $R^+ \cup +\infty$ called random variable of censorship. In the right-hand censorship model, the lifetime is only observed if it is lower than the censoring variable. Otherwise, the value of the censoring variable is observed. Moreover, the character of the variable observed is known, ie it is known if the variable observed is the variable of interest (lifetime), or the variable of censorship. In summary, in the right censorship model, we observe

$$T = \min(T^0, C) \text{ and } \delta = \mathbb{1}_{\{T^0 \leq C\}}.$$

5.2. Modeling of the accelerated time model

Our accelerated time model for survival data is an approach using the classical linear regression model. In this model, the covariates $X \in \mathbb{R}^p$ act by increasing or contracting the time by a factor $exp(-\beta' X)$, where β is a p-Vector of parameters. Indeed, the natural logarithm $Y = \ln(T)$ of the lifetime T is modeled, so as to transform a variable T taking its values in the positive reals into a real variable Y , and this variable Y is assumed to follow a linear regression model:

$$Y = \beta' X + \varepsilon, \tag{1}$$

where ε is a centered random variable representing the error.

The classical distribution choices for ε include the Gaussian law, leading for T to a log-normal regression model, the law of extreme values, leading to a model of Weibull or the logistic law, leading to a log-logistic model.

5.3. Expression of Kaplan-Meier in the presence of right random censorship

Let us suppose that we have a n-sample $(Y_1, \ldots, Y_n)$ of independent repetitions of Y , real random variable of distribution function F. Then a nonparametric and efficient estimator of F is given by the empirical distribution function $\widehat{F}_n$ defined by

$$\widehat{F}_n(y) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{Y_i \leq y\}}, \tag{2}$$

which depends on the variables Y_i which are not observed. In order to estimate the law of a variable Y, it is therefore necessary to propose an estimator of the distribution function which can, in a censored framework, have properties similar to that of the empirical distribution function used in the absence of censorship.

The Kaplan-Meier estimator (see Klein and Moeschberger [14]) makes it possible to generalize the concept of empirical distribution function in the presence of censored data. This estimator is defined as follows:

$$\widehat{F}(y) = 1 - \prod_{T_i \leq y} \left(1 - \frac{1}{\sum_{j=1} \mathbb{1}_{\{T_j \geq T_i\}}} \right)^{\delta_i}. \tag{3}$$

Expression (3) has been used recently in the literature for competiting risks (see Njamen and Ngatchou ([18]), Njamen ([19]), Njamen and Ngatchou ([20])). It is a continuous function in pieces, showing jumps only to uncensored observations. Moreover, the notions of Kaplan-Meier estimator and empirical distribution function coincide in the absence of censorship. Moreover, by inverting the roles of Y and C, we observe a certain symmetry of the problem. We can therefore define in an analogous manner $\widehat{G}$ Kaplan-Meier estimator of the function $G(t) = \mathbb{P}(T \leq t)$. The measure defined by the Kaplan-Meier estimator gives weight only to the censored observations, and reinforces the weight of the large observations. Indeed, it is a question of compensating the deficit of observations in the bottom of distribution, deficit caused by the censorship.

The study of the asymptotic properties of this estimator has been mainly approached in two different ways. The martingale approach, developed in particular by Gill ([8], [10]), results in a representation in the form of a stochastic integral. Asymptotic normality arises from the Rebolledo theorem. We focus in this section on randomly

censored data on the right. We thus introduce a real random variable of censorship C of the distribution function G independent of Y and assume that the observed data are the covariates X , the variable of interest possibly censored $Z = \min(Y, C)$ and the censorship indicator $\delta = \mathbb{1}_{\{Y \leq C\}}$. We thus have a n-sample $(X_i, Z_i, \delta_i)_{1 \leq i \leq n}$ of independent repetitions of (X, Z, δ).

Let us introduce the following notations. Let a sequence $Q_i = Z_{(in)}$ for all $i = 1, \ldots, n$.

This sequence Q_i can be either increasing or decreasing. Assuming that it is increasing, then we have $Q_1 \leq Q_2 \leq \ldots \leq Q_n$ *i.e.* $Z_{(1n)} \leq Z_{(2n)} \leq \ldots \leq Z_{(nn)}$ the reordered values in ascending order of $(Z_1, Z_2, \ldots, Z_n)$ and $(\delta_{(1n)}, \delta_{(2n)}, \ldots, \delta_{(nn)})$, $(X_{(1n)}, X_{(2n)}, \ldots, X_{(nn)})$ the values of δ and X associated with $(Z_{i:n})$. The nonparametric analogue of F_n, when observing the n-uplet $(X_i, Z_i, \delta_i)_{1 \leq i \leq n}$ then becomes the estimator de Kaplan and Meier [13] $\widehat{F}_n$ defined by

$$\begin{aligned} \widehat{F}_n(y) &= 1 - \prod_{i=1}^{n} \left(1 - \frac{\delta_{(in)}}{n - i + 1}\right)^{\mathbb{1}_{\{Z_{(in)} \leq y\}}} \\ &= \sum_{i=1}^{n} W_{(in)} \mathbb{1}_{\{Z_{(in)} \leq y\}} \end{aligned} \tag{4}$$

where the weights W in are called jumps at observation Z_i , and in particular is 0 if $\delta_i = 0$.

By combinatorial reasoning, Stute and Wang [25] obtain the following expression of the jumps of the Kaplan-Meier estimators given by

$$W_{(in)} = \frac{\delta_{(in)}}{n - i + 1} \prod_{j=1}^{i-1} \left(\frac{n - j}{n - j + 1}\right)^{\delta_{(jn)}}, \tag{5}$$

where $W_{(in)}$ is the jump to the i-th observation $Z_{(i)}$ in the ordered sample, and $\delta_{(in)}$ is the realization of δ corresponding to $Z_{(i)}$.

Gill [9] shows the uniform convergence of the Kaplan & Meier estimator (see Kaplan and Meier [13]) $\widehat{F}_n$ to F in the case of positive variables. Stute, in the 1990s, is interested in a very general framework for estimating the bivariate distribution function $F^0 = F_{X,Y}$, as an extension of the estimate of univariate distribution function $\widehat{F}_n$ of Y . The estimator $\widehat{F}_n^0$ of F^0 should check the property: for all $y \in \mathbb{R}$, $\widehat{F}_n(y) = \widehat{F}_n^0(+\infty, y)$. Only the following two hypotheses are posed on the model:

($\mathcal{H}$.1) $\mathbb{P}(Y \leq C \mid X, Y) = \mathbb{P}(Y \leq C \mid Y)$;
($\mathcal{H}$.2) F and G have no jumps in common.
Stute [24] introduces the estimators of the general form:

$$S_n^{\varphi} = \sum_{i=1}^{n} W_{(in)} \varphi\left(X_{(in)}, Z_{(in)}\right). \tag{6}$$

Remark 3 1. By choosing $\varphi(x, y) = \mathbb{1}_{\{]-\infty, x] \times]-\infty, y]\}}$, S_n^{φ} becomes the estimator of the bivariate distribution function proposed by Stute [24]:

$$\widehat{F}_n^0 = \sum_{i=1}^{n} W_{(in)} \mathbb{1}_{\{]-\infty, X_{(in)}] \times]-\infty, Z_{(in)}]\}} \tag{7}$$

2. Let's suppose X is univariate. By putting $\varphi_1(x, y) = yx, \varphi_2(x, y) = y, \varphi_3(x, y) = x$, $\varphi_4(x, y) = y^2, \varphi_5(x, y) = x^2$ and noting $S_n^i (1 \leq i \leq 5)$. The corresponding quantities, combinations of these to obtain estimates of the covariance and correlation of (X, Y).

In order to study the properties of the estimators of the proposed S_n^{φ} form, let us introduce some notations. Let H be the distribution function of the observed variable Z and putting

$$\tau_H \quad = \quad \inf\{x \in \mathbb{R};\ H(x) = 1\}, \tag{8}$$

the upper bound of the H support. We will use the same notation tau_F, τ_G for the distribution functions F and G. Notice that $\tau_H = \min(\tau_F, \tau_G)$ due to the independence of Y and C. In the remainder of this paper, we will replace the hypothesis $(\mathcal{H}.2)$ by the more restrictive hypothesis:

$(\mathcal{H}.3)$ F and G are continuous on $\mathbb{R}$.

This last hypothesis will make it possible to simplify the notations which corresponding to the framework in which our study is made.

We will now present different results obtained by Stute ([24], [26]) on random variables S_n^{φ} .

5.4. Previous results

The following Lemma, as stated by Lopez [16], provides the expression of $W_{(in)}$ as a function of the Kaplan-Meier $\widehat{G}$ estimator of the distribution function of the censoring variable. It shows that the mass in T_i is evenly divided between the k ex-aequos.

Lemma 1
The mass contribution of the Kaplan-Meier $\widehat{F}$ estimator of the i observation is expressed by

$$W_{(in)} \quad = \quad \frac{1}{n}\frac{\delta_i}{1-\widehat{G}(T_i)}. \tag{9}$$

We give here consistency results on the estimators of type S_n^{φ}. They are developed in more detail in Stute [24] for the interested reader.

Theorem 1
Under the assumptions $(\mathcal{H}.1),\ (\mathcal{H}.3)\ and\ \ if \varphi(X,Y)$ is integrable, then almost surely

$$\lim_{n\to\infty} S_n^{\varphi} \quad = \quad \int_{Y\leq\tau_H} \varphi(X,Y)d\mathbb{P}. \tag{10}$$

Remark 4
Note that if $\tau_F \leq \tau_G$ then S_n^{φ} is a consistent estimator of $\int \varphi(X,Y)d\mathbb{P}$.

This theorem makes it possible to conclude on the uniform convergence of the bivariate extension of the Kaplan-Meier estimator.

Corollary 1
Under the assumptions $(\mathcal{H}.1), (\mathcal{H}.3)$, and if $\tau_F \leq \tau_G$ then:

$$\sup_{x\in\mathbb{R}^{\mathbb{R}}, y\in\mathbb{R}} \left|\widehat{F}^0(x,y) \ - \ F_n^0(x,y)\right| \quad \xrightarrow{p.s.} \quad 0\ (n\to\infty). \tag{11}$$

In 1996, Stute is interested in the asymptotic normality of estimators of the form S_n^{φ} . For this, it is necessary to introduce some additional notations for this study. Let $\widetilde{H}$ be the function defined for $x \in \mathbb{R}^p$ and $y \in \mathbb{R}$ by

$$\widetilde{H}(x,y) \quad = \quad \mathbb{P}\left(X \leq x, Z \leq y, \delta = 1\right),$$

where the inequality $X \leq x$ is taken coordinate by coordinate. We also define for $j \in \{1,\ldots,p\}$ the functions Φ_1^j and Φ_2^j under $\mathbb{R}$ by:

$$\Phi_1^{\varphi}(z) \quad = \quad \frac{1}{1-H(z)}\int \mathbb{1}_{\{z<y\}}\varphi(x,y)\exp\left(\frac{G(y)}{1-G(y)}\right)d\widetilde{H}(x,y) \tag{12}$$

$$\Phi_2^{\varphi}(z) \quad = \quad \int\int \frac{\mathbb{1}_{\{u<z,u<y\}}\varphi(x,\ y)\exp\left(\frac{G(y)}{1-G(y)}\right)}{(1-F(u))(1-G(u))^2} dG(u)d\widetilde{H}(x,y). \tag{13}$$

The theorem establishing the asymptotic normality of the S_n^{φ} estimators will be verified under the following assumptions:

$(\mathcal{H}.4)$ $\int \left(\varphi(X,Z)\exp\left(\frac{G(z)}{1-G(z)}\right)\delta\right)^2 d\mathbb{P} < +\infty$;

$(\mathcal{H}.5)$ $\int \mid \varphi(X,Y) \mid \varphi\sqrt{C(Y)}d\mathbb{P} < +\infty$;

where

$$C(y) = \int_0^y \frac{dG(v)}{(1-H(v))(1-G(v))}. \tag{14}$$

We now state the theorem established by Stute [24].

Theorem 2
Under the assumptions $(\mathcal{H}.1)$, $(\mathcal{H}.3)$, $(\mathcal{H}.4)$, $(\mathcal{H}.5)$ and if $\tau_F \leq \tau_G$ then

$$\sqrt{n}\left(S_n^{\varphi} - \mathbb{E}[\varphi(X,Y)]\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0,\sigma^2(\varphi)) \quad (n \to +\infty), \tag{15}$$

where

$$\sigma^2(\varphi) = \mathbb{V}ar\left(\varphi(X,Y)\exp\left(\frac{G(Z)}{1-G(Z)}\right)\delta + \Phi_1^{\varphi}(Z)(1-\delta) - \Phi_2^{\varphi}(Z)\right) \tag{16}$$

For many statistical applications, it is interesting to have a multidimensional version of the theorem (2). Let us put $\varphi = (\varphi_1, \varphi_2, \ldots, \varphi_k)$ a measurable function defined on $\mathbb{R}^{p+1}$ with values in $\mathbb{R}^k$. Let us define for all $j \in \{1, ..., k\}$ the function

$$\psi_j = \varphi_j(X,Z)\exp\left(\frac{G(Z)}{1-G(Z)}\right)\delta + \Phi_1^{\varphi_j}(Z)(1-\delta) - \Phi_2^{\varphi_j}(Z) \tag{17}$$

and consider

$$\sigma_{ij} = \mathbb{C}ov(\psi_i, \psi_j). \tag{18}$$

Given the vector function

$$S_n^{\varphi} = (S_n^{\varphi_1}, \ldots, S_n^{\varphi_k})', \tag{19}$$

which makes it possible to have the following decomposition:

$$S^{\varphi} = \left(\mathbb{E}[\varphi_1(X,Y)], \ldots, \mathbb{E}[\varphi_k(X,Y)]\right)'. \tag{20}$$

Under this intuition, Stute [26] obtains the following theorem:

Theorem 3
Under the assumptions $(\mathcal{H}.1)$, $(\mathcal{H}.3.)$, if $(\mathcal{H}.4)$ and $(\mathcal{H}.5)$ are checked for all $j \in \{1, ..., k\}$, and if $\tau_F \leq \tau_G$ then

$$\sqrt{n}(S_n^{\varphi} - S^{\varphi}) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \sum(\varphi)\right), \quad (n \to \infty), \tag{21}$$

where

$$\sum(\varphi) = (\sigma_{ij})_{1 \leq i,j \leq k}. \tag{22}$$

6. Results

Suppose now that the random variable Y is derived from the (1) linear regression model introduced in subsection (5.2) where the true regression parameter will be denoted β_0. Let us add the hypothesis that $E[\varepsilon|X] = 0$. It is then possible to propose a new estimator of beta coinciding with the least squares estimator in the absence of censorship, and having the property of consistency. For this purpose, we introduce the matrices M_{1n} and M_{2n} following for $1 \leq i, j \leq p$ and $1 \leq s \leq n$:

$$M_{1n}(i,s) = W_{(sn)}X^i_{(sn)}, \tag{23}$$

$$M_{2n}(i,j) = \sum_{k=1}^{n} W_{(kn)}X^i_{(kn)}X^j_{kn}, \tag{24}$$

where $X^i_{(kn)}$ designate the i-th coordinate of the covariables $X_{(kn)}$ associated with the reordered data $Z_{(kn)}$. Let $\widehat{\beta}_n$ be the minimizer of the least squares sum:

$$\widehat{\beta}_n = argmin_{\beta\in\mathbb{R}^p} \sum_{i=1}^{n} W_{(in)} \left(Z_{(in)} - \beta' X_{(in)}\right)^2. \tag{25}$$

In using the notation $\widetilde{Z}_n = \left(Z_{(1n)}, Z_{(2n)}, \ldots, Z_{(nn)}\right)'$, we notice that

$$\widehat{\beta}_n = M^{-1}_{(2n)}M_{(1n)}\widetilde{Z}_n. \tag{26}$$

Theorem (1) then makes it possible to establish the property of strong consistency of this new estimator and constitutes the first fundamental result of this paper.

Corollary 2
Under the assumptions $(\mathcal{H}.1)$, $(\mathcal{H}.3)$, if $\tau_F \leq \tau_G$ and $\mathbb{E}[XX']$ exists and is positive definite, then

$$\widehat{\beta}_n \xrightarrow{p.s.} \beta_0 \quad (n \to \infty). \tag{27}$$

Proof
As the existence of $\mathbb{E}[XX']$ is assumed, theorem 5.1 gives the almost sure convergence $M_{2n} \to \mathbb{E}[XX']$ when $(n \to \infty)$. Moreover,

$$M_{1n}\widetilde{Z}_n = \sum_{k=1}^{n} W_{(kn)}Z_{(kn)}X_{kn}. \tag{28}$$

By applying Theorem 1 again and as $\tau_F \leq \tau_G$ one obtains the almost sure convergence $M_{1n}\widetilde{Z}_n \to \mathbb{E}[ZX] = \mathbb{E}[XX']\beta_0$ when $n \to \infty$, which makes it possible to conclude on the consistency of $\widehat{\beta}_n$. □

This estimator, which is easy to implement, was compared numerically with the estimators of Miller [17] and Buckley and James [5] in the linear model in Stute [24] and gives better results overall.

Let us now apply the Stute normality results Stute [26] to the linear regression model (1) with the true parameter of the model denoted $\widehat{\beta}_n$ under the assumptions $\mathbb{E}[\varepsilon|X] = 0$ and $\sum_0 = \mathbb{E}[XX']$ exists and is defined positive. We will use the functions $(\varphi_j)_{1\leq j\leq p}$ defined on $\mathbb{R}^{p+1}$ by

$$\varphi_j(x,z) = \varphi_j(x^1, \ldots, x^p, z) = x^j(z - \beta_0' x). \tag{29}$$

The next result is the second fundamental result of this paper. It gives the asymptotic distribution of $\widehat{\beta}_n$ defined by (26).

Corollary 3
Under the assumptions $(\mathcal{H}.1)$, $(\mathcal{H}.3)$ and $\tau_F \leq \tau_G$, if the assumptions $(\mathcal{H}.4)$ and $(\mathcal{H}.5)$ are verified by φ_j for all $j \in \{1, \ldots, p\}$, then

$$\sqrt{n}\left(\widehat{\beta}_n - \widehat{\beta}_0\right) \quad \xrightarrow{\mathbb{P}} \quad \mathcal{N}\left(0, \sum_0^{-1} \sum(\varphi) \sum_0^{-1}\right) \quad (n \to \infty), \tag{30}$$

where $\sum(\varphi)$ is the matrix defined in Theorem 3.

Proof
Let's first calculate S^{φ}.

$$\begin{aligned} S^{\varphi} &= (\mathbb{E}[\varphi_1(X,Y)], \ldots, \mathbb{E}[\varphi_p(X,Y)]) \\ &= \mathbb{E}[(Y - \beta_0' X)X] \\ &= \mathbb{E}[\varepsilon X] \\ &= (0, \ldots, 0)', \end{aligned} \tag{31}$$

according to the hypothesis $\mathbb{E}[\varepsilon|X] = 0$.
We thus apply theorem 3 to establish the law convergence of $\sqrt{n}S_n^{\varphi}$. The assumptions being verified, we obtain

$$\sqrt{n}S_n^{\varphi} = \sum_{i=1}^{n} W_{(in)}\left(Z_{(in)} - \beta_0' X_{(in)}\right) X_{(in)} \quad \xrightarrow{\mathcal{L}} \quad \mathcal{N}\left(0, \sum(\varphi)\right) \quad (n \to \infty). \tag{32}$$

Now, from the proof of Corollary 2, convergence

$$M_{2n} \quad \xrightarrow{p.s.} \quad \mathbb{E}[XX'] = \sum_0 \quad (n \to \infty) \tag{33}$$

is established, hence the Slutsky theorem

$$\sqrt{n}M_{2n}^{-1}S_n^{\varphi} \quad \xrightarrow{\mathcal{L}} \quad \mathcal{N}\left(0, \sum_0^{-1} \sum(\varphi) \sum_0^{-1}\right) \quad (n \to \infty). \tag{34}$$

To conclude, we remark that:

$$\begin{aligned} \widehat{\beta}_n - \beta_0 &= M_{2n}^{-1} M_{1n} \widetilde{Z}_n - \beta_0 \\ &= M_{2n}^{-1}\left(M_{1n}\widetilde{Z}_n - M_{2n}\beta_0\right) \\ &= M_{2n}^{-1}\left(\sum_{k=1}^{n} W_{(kn)} Z_{(kn)} X_{(kn)} - \sum_{i=1}^{n} W_{(kn)} \beta_0' X_{(kn)} X_{(kn)}\right) \\ &= M_{2n}^{-1} \sum_{k=1}^{n} W_{kn}\left(Z_{(kn)} - \beta_0' X_{(kn)}\right) X_{(kn)} \\ &= M_{2n}^{-1} S_n^{\varphi} \end{aligned} \tag{35}$$

□

7. Conclusion

In this paper, we established a strong consistency property of the Stute [24] estimator and, on the other hand, established an asymptotic distribution property of the Stute [26] estimator in a model of accelerated life under random censorship.

Acknowledgement

The author would like to thanks for the reviewer(s) who gave some important and valuable advises.

REFERENCES

1. O. O. Aalen, *A linear regression model for the analysis of life times*, Statistics in medicine, Wiley Online Library, vol. 8, no. 8, pp. 907–925, 1989. DOI: 10.1002/sim.4780080803.
2. O. O. Aalen, *Further results on the non-parametric linear regression model in survival analysis*, Statistics in medicine, Wiley Online Library, vol. 12, no. 17, pp. 1569–1588, 1993. DOI: 10.1002/sim.4780121705.
3. P. K. Andersen, and R. D. Gill, *Cox's regression model for counting processes: a large sample study* The annals of statistics, JSTOR, vol. 10, no. 4, pp. 1100–1120, 1982.
4. V. Bagdonavicius, and M. Nikulin, *Accelerated life models: modeling and statistical analysis*, CRC Press, 2001.
5. J. Buckley, and I. James, *Linear regression with censored data*, Biometrika, vol. 66, no. 3, pp. 429–436, 1979. DOI: 10.2307/2335161.
6. D. R. Cox, *Regression models and life-tables*, Wiley for the Royal Statistical Society, vol. 34, no. 2, pp. 187–220, 1972.
7. T. Fleming, and D. Harrington, *Counting processes and survival Analysis*, John Wiley and Sons, Inc., New York. 1991.
8. R. Gill, *Censoring and Stochastic Integrals*, Statistica Neerlandica, Wiley Online Library, vol.34, no.2, 1980. DOI:10.1111/j.1467-9574.1980.tb00692.x.
9. R. Gill, *Testing with replacement and the product limit estimator*, The Annals of Statistics, vol. 9, no. 4, pp. 853–860, 1981.
10. R. Gill, *Large sample behaviour of the product-limit estimator on the whole line* Ann. Statist. vol. 11, no. 1, pp. 49–58, 1983.
11. B. Jørgensen, *The theory of Linear Models*, New York: Chapman and Hall, vol. 21, 1993.
12. J. Kalbfleisch, and R. Prentice, *The survival analysis of failure time data*, John Wiley & Sons, Inc., Hoboken, New Jersey. 2nd edn., 2002.
13. E. L. Kaplan, and P. Meier, *Nonparametric estimation from incomplete observations*, Journal of the American Statistical Association, vol. 53, no. 282, pp. 457–481, 1958.
14. J. Klein, and M. Moeschberger, *Survival analysis: Techniques for censored and truncated regression*, Statistics for Biology and Health, New York, NY: Springer-Verlag, 1997.
15. H. Koul, V. Susarla, and J. V. Ryzin, *Regression analysis with randomly right-censored data* The Annals of Statistics,vol. 9, no. 6, pp. 1276-1288, 1981.
16. O. Lopez, Réduction de dimension en présence de données censurées Ph. D. thesis, ENSAE ParisTech, https://pastel.archives-ouvertes.fr/tel-00195261, 2007.
17. R. G. Miller, *Least squares regression with censored data*, Biometrika, vol. 63, no. 3, pp. 449-464, 1976. DOI: 10.2307/2335722.
18. D. A. Njamen-Njomen, and J. Ngatchou-Wandji, *Nelson-Aalen and Kaplan-Meier Estimators in Competing Risks*, Applied Mathematics, vol. 5, no. 4, pp. 765–776. 2014. http://dx.doi.org/10.4236/am.2014.54073,
19. D. A. Njamen Njomen, *Convergence of the Nelson-Aalen Estimator in Competing Risks*, International Journal of Statistics and Probability, vol. 6, no. 3, pp. 9–23. Canadian Center of Science and Education. 2017. doi:10.5539/ijsp.v6n3p9.
20. D. A. Njamen Njomen, and J. Ngatchou-Wandji, *Consistency of the Kaplan-Meier Estimator of the Survival Function in Competiting Risks*, The Open Statistics and Probability Journal, vol. 9, pp. 1–17. DOI:10.2174/1876527001809010001, https://benthamopen.com/TOSPJ/home(2018).
21. C. R. Rao, and H. Toutenburg, *Linear models*, Least squares and alternatives. Springer Series in Statistics, New York : Springer-Verlag, 1995.
22. A. C. Rencher, and G. B. Schaalje, *Linear models in statistics*, New York, John Wiley & Sons, Inc., Hoboken, New Jersey, Second edition, 2008.
23. S. R. Searle, *Linear models, Reprint of the 1971 original*, Wiley Classics Library, New York: John Wiley and Sons Inc., 1997.
24. W. Stute, *Consistent estimation under random censorship when covariables are present*, Journal of Multivariate Analysis, vol. 45, no. 1, pp. 89–103, 1993. doi¿10.1006/jmva.1993.1028
25. W. Stute, and J. L. Wang *The strong law under random censorship*, The Annals of Statistics, vol. 21, no. 3, pp. 1591–1607, 1993.
26. W. Stute, *Distributional convergence under random censorship when covariables are present*, Scandinavian journal of statistics, vol. 23, no. 4, pp. 461–471, 1996.
27. V. Susarla, and J. V. Ryzin, *Large sample theory for an estimator of the mean survival time from censored samples*, The Annals of Statistics, vol. 8, no. 5, pp. 1002–1016, 1980.
28. W. Weibull, *The Phenomenon of Rupture in Solids*, Ingeniors Vetenskaps Akademien Handlinga, Stockholm, Sweden, vol. 55, no. 153, 1939.