MULTI-STEP ESTIMATION OF REGRESSION COEFFICIENTS IN A LINEAR MODEL WITH CENSORED SURVIVAL DATA

Hira L. Koul
Michigan State University

V. Susarla
SUNY - Binghamton

John Van Ryzin
Columbia University and Brookhaven National Laboratory

1. Introduction

This paper introduces a multi-step procedure for estimating the regression coefficients in a linear model when the dependent variable of interest is a randomly right-censored transform of survival, i.e., log lifetime. The procedure is closely related to that introduced by Buckley and James (1979). Using large sample properties developed by the authors (1981), asymptotic large sample consistency and normality are seen to hold for each iterate of the original estimator. A limited simulation study examines the small sample behavior of the procedure.

The linear regression model considered is:

$$X_i = \alpha + \beta x_i + \varepsilon_i, \quad i = 1,\ldots,n \quad , \tag{1}$$

where $\{x_i\}$ are the known independent (design) variables, $\{\varepsilon_i\}$ are independent,

identically distributed (i.i.d.) error variables with an unknown distribution function F with $E(\varepsilon_1) = 0$, $Var(\varepsilon_1) = \sigma^2$, and (α,β) are the parameters of interest. There exists an extensive literature on inference for α and β based on observing the X_i, but only recently has much work been done on estimating α and β when the X_i's are right-censored. For a discussion of some recent results, see Miller (1976,1981), Buckley and James (1979), and Koul, Susarla and Van Ryzin (1981, 1982). The importance of such a problem in survival data analysis where the X_i's are survival times, or transforms thereof such as log lifetimes, has been pointed out in the above references. Typically right-censored data with follow-up times Y_i can be represented as

$$Z_i = \min\{X_i, Y_i\}$$

and

$$\delta_i = \begin{cases} 1 & \text{if } X_i \leq Y_i \quad \text{(uncensored lifetime)} \\ 0 & \text{if } X_i > Y_i \quad \text{(censored lifetime),} \end{cases}$$

where, here and throughout, $i = 1,\ldots,n$. The paper by Miller (1976) considers the situation where $Y_i = \alpha + \beta x_i + \varepsilon_i'$, where $\{\varepsilon_i'\}$ are i.i.d., independent of $\{\varepsilon_i\}$, and proposes estimators of α and β via the method of Kaplan-Meier (1958) least squares. That is, he suggests minimizing the sum of squares of residuals with weights assigned to the summands in the sum of squares according to the Kaplan-Meier estimator of F based on the residuals. Such a solution leads to an iterative procedure for the estimators of α and β which Miller shows may not converge and could lead to inconsistent estimators of α and β unless the above assumption holds, namely that the Y_i's have means following the regression line. To overcome these possible inconsistency problems, Buckley and James (1979) suggest another method, described below, which also may have the same type of convergence problem as does the Miller method, although less so. In Buckley and James, the $\{Y_i = y_i\}$ is taken as a fixed known sequence. In this paper, we present a modification of the Buckley-James procedure under the assumption that

the $\{Y_i\}$ are i.i.d. random variables. The first-step estimators suggested for the α and β are those presented by Koul, Susarla and Van Ryzin (1981), which were shown to be consistent and asymptotically normal. Using these as first-step estimators, we show that the iterated subsequent-step estimators are also consistent and asymptotically normal.

The next section describes the method of Buckley and James (1979). Section 3 presents our modification of their method, while Section 4 provides the large sample consistency and asymptotic normality of the estimator. Section 5 contains some simulations of our method and that of Buckley-James and compares subsequent iterates using as first-step estimators those of Koul, Susarla and Van Ryzin (1981).

2. The Buckley-James Method

Consider the random variable

$$X_i^* = \delta_i X_i + (1 - \delta_i)\, E[X_i | X_i > y_i] \tag{2}$$

for $i = 1,\ldots,n$. Note that X_i^* is an unbiased estimator of $\alpha + \beta x_i$ for each i. Hence, with $d_i = x_i - \bar{x}$, $n\bar{x} = \sum_i x_i$, where $\sum_i$ represents sum over $i = 1,\ldots,n$, we have, $E[\sum_i d_i(X_i^* - \beta x_i)] = 0$. Therefore, a suitable analogue of the normal equation solution for estimating β is to choose that value of b such that

$$b = (\sum_i d_i X_i^*) / \tau_x^2, \quad \tau_x^2 = \sum_i (x_i - \bar{x})^2 \quad .$$

However, the factor $E[X_i | X_i > y_i]$ is unknown in X_i^* and therefore b cannot be calculated directly from the data. Buckley and James (1979) suggest estimating this expectation by using the product limit estimator $\hat{F}_{0,b}(\cdot)$ based on the censored residuals $(\delta_i, Z_i - bx_i)$. Then

$$(3) \qquad E[X_i | X_i > y_i] = bx_i + E[X_i - bx_i | X_i - bx_i > y_i - bx_i]$$

can be estimated by

$$(4) \qquad \bar{X}_i(b) = bx_i + \sum_{k \in A_i} (y_k - bx_k)\, v_k(b) / \hat{F}(y_i - bx_i) \quad ,$$

where $v_k(b)$ is the jump under $\hat{F}_{0,b}$ at all uncensored $X_k - bx_k$ and A_i is the set of all k such that $X_k - bx_k > y_i - bx_i$. Note that the y_i's are fixed. Substituting (4) as an estimator of (3) into b gives the equation

$$(5) \qquad b = \sum_i [\delta_i d_i Z_i + (1 - \delta_i)\, \bar{X}_i(b)] / \tau_x^2 \quad .$$

Buckley and James suggest solving for b in (5) iteratively using $b_0 = \sum_i (\delta_i d_i Z_i) / \tau_x^2$ as the initial estimator of β. They also provide some simulation results for their estimator and a heuristic discussion of its large sample properties.

3. A Modification of the Buckley-James Method

In our modification of Buckley and James (1979), we assume $\{Y_i\}$ to be i.i.d. random variables with some continuous distribution G and that $\{Y_i\}$ is independent of $\{\varepsilon_i\}$. Consider again the relation of (1) written as

$$(6) \qquad X_i^* = \delta_i X_i + (1 - \delta_i)\, E[X_i | \delta_i = 0] \quad .$$

It is easy to see that $E[X_i^*] = \alpha + \beta x_i$ for all i, and the X_i^* follow the regression line of interest. Thus, the Buckley-James observation holds when the $\{Y_i\}$ are random provided the $\{X_i\}$ and $\{Y_i\}$ are independent. However, the $\{Y_i\}$ could be dependent within themselves. Thus, natural estimators of β and α

based on the X_i^* in (6) are

$$(7) \qquad \tilde{\beta} = \left(\sum_i d_i X_i^* \right) / \tau_x^2 \quad , \quad \tilde{\alpha} = n^{-1} \sum_i X_i^* - \tilde{\beta}\, \bar{x} \quad .$$

The X_i^* in (6) themselves cannot be used directly since the $E[X_i | \delta_i = 0]$ are unknown. To overcome this difficulty we will estimate $E[X_i | \delta_i = 0]$ based on a preliminary estimate of β, say, β^*.

Note that the definition of conditional expectation implies

$$(8) \qquad E[X_i | \delta_i = 0] = \frac{\int s G(s)\, dF(s - \alpha - \beta x_i)}{\int G(s)\, dF(s - \alpha - \beta x_i)} = \frac{P_{i,1}}{P_{i,0}} \quad ,$$

where $P_{i,j}$, $j = 0,1$ are defined as indicated. Therefore, if a first-step estimator (α^*, β^*) is available, one can estimate $1 - F$ by $1 - F^*$ given by the Kaplan-Meier (1958) estimator based on the censored residuals $(\delta_i, Z_i - \alpha^* - \beta^* x_i)$ and $1 - G$ can be estimated by the Kaplan-Meier estimator $1 - G^*$ using $(1 - \delta_i, Z_i)$; that is, we treat the lifetimes as censoring the follow-up times. Thus, $P_{i,j}$ for $j = 0,1$ can be estimated by

$$(9) \qquad P_{i,j}^* = \int s^j\, G^*(s)\, dF^*(s - \alpha^* - \beta^* x_i) \quad ,$$

for $i = 1,\ldots,n$ and $j = 0,1$. Substituting (9) into (7) and (8) we have a second-step estimator $\hat{\beta}$ of β given by

$$(10) \qquad \hat{\beta} = \sum_i d_i \{ \delta_i Z_i + (1 - \delta_i)\, P_{i,1}^* / P_{i,0}^* \} / \tau_x^2 \quad .$$

The second-step estimator of α resulting from this procedure would be

$$\hat{\alpha} = n^{-1} \sum_i \{\delta_i Z_i + (1 - \delta_i) P^*_{i,1} / P^*_{i,0}\} - \hat{\beta}\, \bar{x} \quad . \tag{11}$$

Given these second-step estimators, one could then repeat this process to obtain a third-step estimator. Multi-step estimators could be derived by continuing this procedure. In the next section, we show that starting with the consistent and asymptotically normal estimators of (α,β), the second-step estimator (and hence subsequent-step) estimators are also consistent and asymptotically normal. Section 5 investigates by simulation the change in the estimators over the early steps and compares them with the Buckley-James estimators.

4. Some (Intuitive) Large Sample Properties of $\hat{\beta}$

Consider the second-step estimator $\hat{\beta}$ for β given by (10). If $\{P^*_{i,j}\}$ are consistent estimators of $\{P_{i,j}\}$, then $\hat{\beta}$ can be expected to be a consistent estimator of β. Therefore, we consider the behavior of $P^*_{i,j}$ under i.i.d. censoring with distribution G.

Under i.i.d. censoring and certain other regularity conditions β^* can be chosen to be a consistent estimator of β (see Koul, Susarla and Van Ryzin (1981)). In fact, β^* can be taken such that $\tau_x(\beta^* - \beta) \xrightarrow{d} U \sim N(0,\sigma^2)$, where $\xrightarrow{d}$ stands for convergence in distribution as $n \to \infty$ and $N(0,\sigma^2)$ stands for a normal distribution with mean zero and variance σ^2.

We can express the difference $\hat{P}_{i,0} - P_{i,0}$ as

$$\hat{P}_{i,0} - P_{i,0} = \int \{G^*(s + \alpha^* + \beta^* x_i) - G(s + \alpha^* + \beta^* x_i)\}\, dF^*(s) \tag{12}$$

$$+ \int G(s + \alpha^* + \beta^* x_i)\, d(F^* - F)\,(s) \quad .$$

Note that the first term on the right-hand side of (12) can be bounded by $\sup_s |G^*(s) - G(s)|$ which converges to zero with probability one under quite general conditions due to the recent results of Földes and Rejtö (1981). The same situation exists for the second term on the right-hand side of (12) which is bounded by $\sup_s |F^*(s) - F(s)|$. Since these convergent bounds are independent of i, equation (12) yields $\lim_n \sup_i |P^*_{i,0} - P_{i,0}| = 0$ with probability one. Similarly, it can be expected that $\lim_n \sup_i |P^*_{i,1} - P_{i,1}| = 0$ with probability one, under quite general conditions. For example, using results similar to those of Susarla and Van Ryzin (1980) for estimating the mean, it would suffice to have the $\{x_i\}$ be bounded and the G having a heavier tail than F. Hence, from these results, we expect $\hat{\beta}$ of (10) to be a strongly consistent estimator of β.

Note that there is no need to have an estimator of α to implement (10) to find $\hat{\beta}$. This is similar to the situation noted by Buckley and James (1979) in their estimate of β. Thus, in the remainder of the arguments we take $\alpha = \alpha^* = 0$ for simplicity of notation.

To study the asymptotic normality of the second-step $\hat{\beta}$, we consider the random variable $\tau_x(\hat{\beta}-\beta)$. In all statements which follow, by $o_p(1)$ we mean a term which converges to zero in probability as $n \to \infty$. Consider now

$$\tau_x(\hat{\beta}-\beta) = \tau_x^{-1} \sum_i \left\{ d_i(1-\delta_i)\left(\frac{\int (s+\beta^* x_i)\, G^*(s+\beta^* x_i)\, dF^*(s)}{\int G^*(s+\beta^* x_i)\, dF^*(s)} - E[X_i \mid 1-\delta_i] \right) \right\} \tag{13}$$

$$+ \tau_x^{-1} \sum_i d_i(\delta_i Z_i - E[\delta_i Z_i]) = I + II \quad .$$

Since II is easily seen to converge in distribution as a sum of independent random variables provided $\tau_x^2 \to \infty$ as $n \to \infty$, we concentrate our efforts on obtaining an approximation (in probability) of I. For the following details, let

$F_i(s) = F(s - \beta x_i)$, $a_i = \int sG(s)\, dF_i(s)$, and $b_i = \int (1 - F_i(s))\, dG(s)$ and for $i = 1,\ldots,n$, $F_i^*(s) = F^*(s - \beta^* x_i)$. Note that $F_i^*(s)$ is the Kaplan-Meier estimator based only on $\{\delta_i, z_i - \beta^* x_i\}$. Now write I as:

$$
\begin{aligned}
I = \quad & \tau_x^{-1} \sum_i \left\{ d_i(1-\delta_i) \left(\frac{\int sG^*\, dF_i^*}{\int G^*\, dF_i^*} - \frac{\int sG\, dF_i}{\int G^*\, dF_i^*} \right) \right\} \\
& + \tau_x^{-1} \sum_i \left\{ d_i(1-\delta_i) \left(\frac{\int sG\, dF_i}{\int G^*\, dF_i^*} - \frac{\int sG\, dF_i}{\int G\, dF_i} \right) \right\} \qquad (14) \\
& + \tau_x^{-1} \sum_i \left\{ d_i(1-\delta_i) \left(\frac{\int sG\, dF_i}{\int G\, dF_i} - E[X_i \mid 1-\delta_i] \right) \right\} \\
= \quad & I_1 + I_2 + I_3 \quad .
\end{aligned}
$$

Since I_3 is a sum of independent random variables, it can be shown to be asymptotically normal. Thus, consider I_1 and I_2.

We first deal with I_2 which can be rewritten as

$$
\begin{aligned}
I_2 &= \tau_x^{-1} \sum_i d_i(1-\delta_i) \left\{ \int sG\, dF_i \right\} \frac{\left[\int G\, dF_i - \int G^*\, dF_i^* \right]}{\left(\int G\, dF_i \right) \left(\int G^*\, dF_i^* \right)} \\
&= \tau_x^{-1} \sum_i d_i(1-\delta_i) \frac{\int sG\, dF_i}{\left(\int G\, dF_i \right)^2} \left\{ \int G\, dF_i - \int G^*\, dF_i^* \right\} + o_p(1) \qquad (15) \\
&= \tau_x^{-1} \sum_i \frac{a_i}{b_i^2} d_i(1-\delta_i) \left\{ \int (G - G^*)\, dF_i^* + \int G\, d(F_i - F_i^*) \right\} + o_p(1) \\
&= \tau_x^{-1} \sum_i \frac{a_i}{b_i^2} d_i(1-\delta_i) \left\{ \int (G - G^*)\, dF_i + \int (F_i - F_i^*)\, dG \right\} + o_p(1) \quad .
\end{aligned}
$$

Since the term in braces in the last line of (15) is approximately centered, replace $(1-\delta_i)$ by its expectation b_i to obtain

$$(16) \qquad I_2 = \tau_x^{-1} \sum_i d_i \frac{a_i}{b_i} \{ \int (G - G^*)\, dF_i + \int (F_i - F_i^*)\, dG \} + o_p(1) \quad .$$

The first term of (16) can be treated by an approximation as in Koul, Susarla and Van Ryzin (1981). This involves first approximating the term by a U-statistic, and then reducing it to a sum of independent centered random variables, denoted by $A_1 + \cdots + A_n$.

To approximate the term $\sum_i d_i(a_i/b_i) \int (F_i - F_i^*)dG$ in (16), we have to study the process $\{|F_i^*(s) - F_i(s)| \; -\infty < s < \infty\}$, for $i = 1,\ldots,n$, more carefully. Recall here that $F_i(s) = F(s-\beta x_i)$. A Taylor expansion yields

$$F_i^*(s) - F_i(s) = \exp\{\ln F_i^*(s)\} - \exp\{\ln F_i(s)\}$$

$$(17) \qquad = \sum_j \frac{\delta_j [Z_j - \beta x_j < s - \beta x_i + (\beta^* - \beta)(x_j - x_i)]}{1 + \sum_k [Z_k - \beta x_k > Z_j - \beta x_j + (\beta^* - \beta)(x_j - x_k)]} - \ln F_i(s) + o_p(1) \; ,$$

where $o_p(1)$ is independent of i and $[S]$ is the indicator function of the set S. The random variable in (17) can further be approximated by

$$\frac{1}{n} \sum_j \delta_j \frac{[Z_j - \beta x_j < s - \beta x_i + (\beta^* - \beta)(x_j - x_i)]}{n^{-1} \sum_k [Z_k - \beta x_k > Z_j - \beta x_j] + n^{-1} \sum_k (x_j - x_k)\, f_k(Z_j - \beta x_j)(\beta^* - \beta)}$$

$$- \ln F_i(s) + o_p(1) \quad ,$$

where f_k is the density of $Z_k - \beta x_k$. The factor

$|\beta^*-\beta|\,|x_j| < \tau_x\,|\beta^*-\beta|\,(\tau_x^{-1}\max_i|x_i|) \xrightarrow{p} 0$ whenever $\max_i|x_i|\,\tau_x^{-1} \to 0$ and $\beta^* \xrightarrow{p} \beta$ as $n \to \infty$. Therefore,

$$(19)\quad F_i^*(s) - F_i(s) = \frac{1}{n}\sum_j \frac{\delta_j[Z_j - \beta x_j < s - \beta x_i]}{n^{-1}\sum_k[Z_k - \beta x_k > z_j - \beta x_j]} - \ln F_i(s) + o_p(1) \quad .$$

This term is similar to the terms dealt with in Koul, Susarla and Van Ryzin (1981) which can be approximated by a sum of independent centered random variables, denoted by $B_1 + \cdots + B_n$. Combining this result with that earlier for the first-term of (16), we have

$$(20)\qquad I_2 = A_1 + \cdots + A_n + B_1 + \cdots + B_n + o_p(1) \quad .$$

To treat the term I_1 in (14), rewrite I_1 as

$$(21)\qquad I_1 = \tau_x^{-1}\sum_i d_i(1-\delta_i)(b_i^*)^{-1}\{\int s(G^* dF_i^* - G dF_i)\} \quad ,$$

and noting that $\lim_n b_i^* = \lim \int G^* dF_i^* = \int G dF_i = b_i$ and that $\lim_n \sup_i \sup_s |F_i^*(s) - F_i(s)| = 0$ with probability one, we can approximate I_1 as

$$(22)\qquad I_1 = \tau_x^{-1}\sum_i d_i(1-\delta_i)\, b_i^{-1}\{\int s(G^* - G)dF_i + \int s(F_i^* - F_i)dG\} + o_p(1) \quad .$$

By writing $G^*(s) - G(s) = \exp\{\ln G^*(s)\} - \exp\{\ln G(s)\}$ and similarly for $F_i^* - F_i$ and doing a Taylor expansion similar to that in (17), I_1 can be approximated as a sum of independent centered random variables, denoted by $C_1 + \cdots + C_n$. This result combined with (13), (14) and (20) yield $\tau_x(\hat{\beta} - \beta) = II + (A_1 + \cdots + A_n) + (B_1 + \cdots + B_n) + (C_1 + \cdots + C_n) + I_3$, which we see is the sum of a triangular array of independent random variables, and hence by the Lindeberg-Feller central limit theorem will be asymptotically normal. Note that the conditions that $\tau_x^2 \to \infty$ as $n \to \infty$, $\max_i |x_i| / \tau_x \to 0$ and that G have heavier tails than F are required for this to hold. Thus, asymptotic normality of $\hat{\beta}$ is expected to hold provided the first-step estimator β^* is consistent. Such a consistent first-step estimator is given by Koul, Susarla and Van Ryzin (1981). That subsequent finite-step estimators of β are consistent and asymptotically normal follows inductively.

5. Some Simulation Results

Table 1 presents the results of six simulations of the multi-step procedure of this paper and that of Buckley and James (1979). In all cases, the table entries are based on 500 simulations of the situation described. All simulations are for the two sample problem with $n = 50$ observations where $x_i = 0$ for $i = 1,\ldots,25$ and $x_i = 1$ for $i = 26,\ldots,50$. The simulation sample standard errors for the simulation averages of $\hat{\alpha}$ and $\hat{\beta}$ for the 500 repetitions were in all cases $\leq .017$ for $\hat{\alpha}$ and $\leq .031$ for $\hat{\beta}$ and thus are not individually given to save space. The first-step estimates for both methods were taken as the estimator of α and β as defined in Koul, Susarla and Van Ryzin (1981) with $M_n = n$, and are denoted in the tables as method M_1. The second, third and fourth-step estimators for the estimators introduced in this paper are denoted by M_2, M_3, and M_4, respectively, while those of Buckley and James are referred to as M_2^*, M_3^* and M_4^*. Furthermore, in each case of Table 1 the error distribution for the simulation was taken as $\varepsilon_i \sim N(0,1)$. The follow-up distribution G(y) for the first five cases are exponential with mean μ and are denoted by $E(\mu)$ in Table 1 while the six case has a right-sided logistic distribution given by $1 - G(y) =$

$2e^{-y}/(1+e^{-y})$ on $[0,\infty]$ and is denoted by LOG(½).

Upon examining Table 1, it is clear that the method of this paper performed better than that of Buckley-James over the initial three iterates after the first-step. Thus, based on these limited simulations, we feel the multi-step procedure introduced in this paper holds considerable promise. Simulations for differing sample sizes and for regression situations other than the two samples are under investigation and will be presented elsewhere.

TABLE 1. Simulation of 4-step estimators based on 500 replicates.

Simulated Case	Method of Estimation	Average of Estimates (α,β)	Average Mean Square Error of Estimates (α,β)
	M_1	(-.200,.017)	(.082,.077)
$(\alpha,\beta) = (0,0)$	M_2	(-.144,.015)	(.074,.087)
Follow-up dist. =E(1)	M_3	(-.137,.015)	(.076,.096)
Average censoring:	M_4	(-.135,.015)	(.077,.100)
Sample 1 = .44	M_2^*	(-.190,.013)	(.085,.082)
Sample 2 = .44	M_3^*	(-.186,.012)	(.087,.088)
	M_4^*	(-.185,.012)	(.088,.092)
	M_1	(-.227,-.797)	(.087,.115)
$(\alpha,\beta) = (0,-1)$	M_2	(-.130,-.906)	(.067,.096)
Follow-up dist. = E(2)	M_3	(-.105,-.939)	(.066,.098)
Average censoring:	M_4	(-.096,-.949)	(.066,.099)
Sample 1 = .44	M_2^*	(-.200,-.875)	(.084,.094)
Sample 2 = .22	M_3^*	(-.193,-.892)	(.084,.094)
	M_4^*	(-.191,-.987)	(.084,.094)

Table 1 (continued)

Simulated Case	Method of Estimation	Average of Estimates (α,β)	Average Mean Square Error of Estimates (α,β)
	M_1	(.960,.760)	(.083,1.872)
$(\alpha,\beta) = (1,2)$	M_2	(1.021,1.485)	(.053,.400)
Follow-up dist. = E(20)	M_3	(.976,1.743)	(.048,.176)
Average censoring:	M_4	(.967,1.855)	(.048,.129)
Sample 1 = .18	M_2^*	(.961,1.338)	(.049,.590)
Sample 2 = .62	M_3^*	(.927,1.560)	(.051,.306)
	M_4^*	(.908,1.655)	(.053,.214)
	M_1	(.895,.234)	(.119,3.584)
$(\alpha,\beta) = (1,2)$	M_2	(1.055,1.022)	(.076,1.217)
Follow-up dist. = E(10)	M_3	(.961,.1.387)	(.058,.576)
Average censoring:	M_4	(.942,1.597)	(.058,.326)
Sample 1 = .30	M_2^*	(.911,.815)	(.068,1.692)
Sample 2 = .79	M_3^*	(.871,1.147)	(.071,.958)
	M_4^*	(.839,1.343)	(.078,.620)
	M_1	(-.008,1.708)	(.047,1.895)
$(\alpha,\beta) = (0,3)$	M_2	(.012,2.506)	(.046,.366)
Follow-up dist. =E(20)	M_3	(-.008,2.776)	(.044,.155)
Average censoring:	M_4	(-.012,2.884)	(.043,.118)
Sample 1 = .07	M_2^*	(-.012,2.276)	(.044,.662)
Sample 2 = .62	M_3^*	(-.037,2.487)	(.044,.372)
	M_4^*	(-.052,2.574)	(.045,.277)

Table 1 (continued)

Simulated Case	Method of Estimation	Average of Estimates (α,β)	Average Mean Square Error of Estimates (α,β)
	M_1	(-.248,-.790)	(.094,.113)
$(\alpha,\beta) = (0,-1)$	M_2	(-.144,-.907)	(.066,.091)
Follow-up dist. = LOG(½)	M_3	(-.115,-.945)	(.064,.093)
Average censoring:	M_4	(-.104,-.959)	(.064,.095)
Sample 1 = .50	M_2^*	(-.234,-.884)	(.094,.087)
Sample 2 = .26	M_3^*	(-.227,-.908)	(.093,.087)
	M_4^*	(-.222,-.915)	(.093,.088)

6. Concluding Remarks

This paper presents a multi-step estimator for the α and β in linear model (1) when the independent variable X_i is randomly right-censored. These estimators are modifications of the Buckley-James estimators. The large sample properties shown to hold here might be extendable to the Buckley-James case of fixed censoring or to our modification of the Buckley-James case when the Y_i are not i.i.d. (see Susarla and Van Ryzin (1979)) if in our method we replace $(1-\hat{G})$ in our formulas by the estimator one would get using the Kaplan-Meier for the non i.i.d. case as an estimator of $\lim_n\{n^{-1}\sum_i(1-G_i(t)\}$, assuming this exists. This seems worth further investigation.

We remark that everything mentioned in this paper easily extends to the multiple regression model where

$$X_i = \underset{\sim}{C}_i\,\underset{\sim}{\beta} + \varepsilon_i,\ i=1,\ldots,n \quad ,$$

with $\underset{\sim}{C}_i$ being the i^{th} row of the $n \times p$ design matrix C, $\underset{\sim}{\beta}$ is the $p \times 1$ vector of regression coefficients, and G_i, Y_i, Z_i and δ_i are as above. Then, the

second-step estimator of $\underset{\sim}{\beta}$ would be given by

$$\hat{\underset{\sim}{\beta}} = (C'C)^{-1} C' \hat{\underset{\sim}{X}}^* \quad , \quad \hat{\underset{\sim}{X}}^* = (\hat{X}_1^*, \ldots, \hat{X}_n^*)' \quad ;$$

with $\hat{X}_i^* = \delta_i Z_i + (1 - \delta_i) P_{i,1}^* / P_{i,0}^*$ where $P_{i,j}^*$ is given by (9) with the first-step uncensored residuals $Z_i = \underset{\sim}{C}'\underset{\sim}{\beta}^*$ with $\delta_i = 1$ used to estimate $1 - F^*$ by the Kaplan-Meier method, $1 - G^*$ is estimated as before, and $\underset{\sim}{\beta}^*$ is the first-step estimate of $\underset{\sim}{\beta}$ given by (5.2) in Koul, Susarla and Van Ryzin (1981) with $M_n = n$.

Clearly, further simulations (or theoretical) studies of the convergence properties (speed, etc.) of our multi-step procedure seem warranted and are anticipated.

ACKNOWLEDGEMENTS

The research of Professors Susarla and Van Ryzin was supported by NIH grant No. 1-R01-GM28405 at Columbia University. The authors wish to thank Ms. Sonja Johansen for the programming of the simulation results of Section 5.

REFERENCES

Buckley, J. and James, I. (1979). Linear regression with censored data. Biometrika 66, 429-436.

Földes, A. and Rejtö, L. (1980). Strong uniform consistency for nonparametric survival curve estimators from randomly censored data. Annals of Statistics 9, 122-129.

Kaplan, E.L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. Journal of the American Statistical Association 53, 457-481.

Koul, H., Susarla, V. and Van Ryzin, J. (1981). Regression analysis with randomly right-censored data. Annals of Statistics 9, 1276-1288.

Koul, H., Susarla, V. and Van Ryzin, J. (1982). Least squares regression analysis with censored survival data. To appear in Topics in Applied Statistics (T.W. Dwivedi, Ed.). Marcel Dekker, New York.

Miller, R.G., Jr. (1976). Least squares regression with censored data. Biometrika 63, 449-464.

Miller, R.G., Jr. (1981). Survival Analysis. Wiley, New York.

Susarla, V. and Van Ryzin, J. (1979). Large sample theory for survival curve estimators under variable censoring. In Optimization Methods in Statistics, 475-508. Academic Press, New York

Susarla, V. and Van Ryzin, J. (1979). Large sample theory for the mean survival time from censored data. Annals of Statistics 8, 1002-1016.