User-Defined Maximum Likelihood Estimation via EasyReg International

Herman J. Bierens

March 14, 2007

Abstract

In this note I will explain how to build the EasyReg code for the conditional treatment model in Bierens and Carvalho (2007).

1 Introduction

During the period of $1980 - 1985$, the National Institute of Justice (NIJ) sponsored a controlled experiment to evaluate the impact of reemployment programs for recent released prisoners. Three well established programs were chosen, in Boston, San Diego and Chicago, to participate in the Employment Services for Ex-Offenders Program, henceforth ESEO. A total of $2,045$ prisoners who voluntarily accepted to participate were randomly assigned to either a treatment group ($G = 1$) or a control group ($G = 0$). Those in the first group received, besides the normal services (orientation, screening, evaluation, support services, job development seminar, and job search coaching), special services which consisted of an assignment to a follow-up specialist who provided support during the job search and the 180 days following job placement. The control group received only normal services. The inclusion of special services was a major response to the increasing belief that some past employment programs had failed because ex-inmates lost contact with their original programs.

2 The data

In this guided tour on user-defined maximum likelihood (ML) estimation, I will use the following subset ESEO variables, corresponding to the final model in Bierens and Carvalho (2007). The number of observations is 503.

2.1 Covariates

- $G = 1$ if selected in the treatment group, $= 0$ if selected in the control group
- AGE: Age in years;
- $SANDIEGO$: Indicator for San Diego;
- $CHICAGO$: Indicator for Chicago;
- $BOSTON = 1 - SANDIEGO - CHICAGO$

2.2 Dependent variables

- T_s ($s = search$) is the duration of the job search, i.e., the time between the date of release and the date of placement in the first job after release;
- T_c ($c = crime$) is the recidivism time, i.e., the time between the date of release and the date of the first arrest after release.

Both durations are interval-censored: $T_s \in (a, b]$, $T_c \in (p, q]$, $b \leq p$, where

$$(a, b], (p, q] \in \{(0, 1], (1, 6], (6, 12], (12, \infty]\}.$$

In the data set that comes with this guided tour these variables have already been transformed to the following six dummy variables

$$\begin{gathered}
I(T_s \in (0, 1]) \times I(T_c \in (1, 6]) \\
I(T_s \in (0, 1]) \times I(T_c \in (6, 12]) \\
I(T_s \in (0, 1]) \times I(T_c > 12) \\
I(T_s \in (1, 6]) \times I(T_c \in (6, 12]) \\
I(T_s \in (1, 6]) \times I(T_c > 12) \\
I(T_s \in (6, 12]) \times I(T_c > 12)
\end{gathered}$$

where $I(.)$ is the indicator function: $I(true) = 1$, $I(false) = 0$.

3 The model

3.1 Job search

We find that only the location dummy variables $CHICAGO$ and $SANDIEGO$ matter for job search. Moreover, we cannot reject the null hypothesis that the baseline hazard is of the Weibull type. Thus, the survival function now takes the form

$$S_s(t|X_s) = P\left[T_s > t|X\right] = \exp\left(-\exp\left(\beta_s' X_s\right) t^{\alpha_s}\right) \tag{1}$$

where $X_s =$ (CHICAGO, SANDIEGO, 1)$'$

3.2 Recidivism

The conditional survival function of T_c given T_s, X_c, and G is specified as

$$\begin{aligned} S_c(t|T_s, X_c, G) &= P\left[T_c > t|T_s, X_c, G\right] = I\left(t \le T_s\right)\exp\left(-\alpha_c.t\right) \\ &+ I\left(t > T_s\right)\exp\left(-\alpha_c.T_s\right).\exp\left(-\alpha_c.\exp\left(G.\beta_c' X_c\right).\left(t - T_s\right)\right), \end{aligned} \tag{2}$$

where $X_c =$ (BOSTON, AGE, 1)$'$

3.3 Log-likelihood

It follows from (**??**) and (**??**) that for $0 \le a < b \le p < q$,

$$\begin{aligned} & P\left[T_c \in (p,q], T_s \in (a,b]|X, G\right] \\ &= \int_a^b S_c(t|t_s, X_c, G) d\left(-S_s(t_s|X_s)\right) \\ &= -\int_a^b \exp\left(-\alpha_c.\left(1 - \exp\left(G.\beta_c' X_c\right)\right) t_s\right) d\left(\exp\left(-\exp\left(\beta_s' X_s\right) t_s^{\alpha_s}\right)\right) \\ &\times \left(\exp\left(-\alpha_c.\exp\left(G.\beta_c' X_c\right).p\right) - \exp\left(-\alpha_c.\exp\left(G.\beta_c' X_c\right).q\right)\right) \\ &= \int_{S_s(b|X_s)}^{S_s(a|X_s)} \exp\left[-\alpha_c.\exp\left(-\alpha_s^{-1}\beta_s' X_s\right)\right. \\ &\times \left.\left(1 - \exp\left(G.\beta_c' X_c\right)\right)\left(\ln\left(1/u\right)\right)^{1/\alpha_s}\right] du \\ &\times \left(\exp\left(-\alpha_c.\exp\left(G.\beta_c' X_c\right).p\right) - \exp\left(-\alpha_c.\exp\left(G.\beta_c' X_c\right).q\right)\right). \end{aligned}$$

Hence, denoting

$$c = \alpha_c.\exp\left(-\alpha_s^{-1}\beta_s' X_s\right)\left(1-\exp\left(G.\beta_c' X_c\right)\right) \quad (3)$$
$$d = 1/\alpha_s \quad (4)$$
$$r = -\alpha_c.\exp\left(G.\beta_c' X_c\right) \quad (5)$$

we have

$$P\left[T_s \in (a,b], T_c \in (p,q] | X, G\right]$$
$$= \left(\exp\left(r.p\right) - \exp\left(r.q\right)\right) \int_{S_s(b|X_s)}^{S_s(a|X_s)} \exp\left(-c.\left(\ln\left(1/x\right)\right)^d\right) dx.$$

The latter integral is available in EasyReg as transformation

Integral(exp[-c.(ln(1/x))^d]dx|a<x<b)(a,b,c,d)
$= \int_a^b \exp\left[-c.\left(\ln\left(1/x\right)\right)^d\right] dx,\ 0 \le a < b \le 1,\ d > 0.$

Thus,

$$P_1 = P\left[T_s \in (0,1], T_c \in (1,6] | X, G\right]$$
$$= \left(\exp\left(r\right) - \exp\left(6r\right)\right) \int_{S_s(1|X_s)}^{1} \exp\left(-c.\left(\ln\left(1/x\right)\right)^d\right) dx.$$
$$P_2 = P\left[T_s \in (0,1], T_c \in (6,12] | X, G\right]$$
$$= \left(\exp\left(6r\right) - \exp\left(12r\right)\right) \int_{S_s(1|X_s)}^{1} \exp\left(-c.\left(\ln\left(1/x\right)\right)^d\right) dx.$$
$$P_3 = P\left[T_s \in (0,1], T_c > 12 | X, G\right]$$
$$= \exp\left(12r\right) \int_{S_s(1|X_s)}^{1} \exp\left(-c.\left(\ln\left(1/x\right)\right)^d\right) dx.$$
$$P_4 = P\left[T_s \in (1,6], T_c \in (6,12] | X, G\right]$$
$$= \left(\exp\left(6r\right) - \exp\left(12r\right)\right) \int_{S_s(6|X_s)}^{S_s(1|X_s)} \exp\left(-c.\left(\ln\left(1/x\right)\right)^d\right) dx.$$
$$P_5 = P\left[T_s \in (1,6], T_c > 12 | X, G\right]$$
$$= \exp\left(12r\right) \int_{S_s(6|X_s)}^{S_s(1|X_s)} \exp\left(-c.\left(\ln\left(1/x\right)\right)^d\right) dx.$$
$$P_6 = P\left[T_s \in (6,12], T_c > 12 | X, G\right]$$
$$= \exp\left(12r\right) \int_{S_s(12|X_s)}^{S_s(6|X_s)} \exp\left(-c.\left(\ln\left(1/x\right)\right)^d\right) dx$$
$$P_7 = 1 - \sum_{i=1}^{6} P_i$$

Denoting

$$\begin{aligned}
Y_1 &= I(T_s \in (0,1]) \times I(T_c \in (1,6]) \\
Y_2 &= I(T_s \in (0,1]) \times I(T_c \in (6,12]) \\
Y_3 &= I(T_s \in (0,1]) \times I(T_c > 12) \\
Y_4 &= I(T_s \in (1,6]) \times I(T_c \in (6,12]) \\
Y_5 &= I(T_s \in (1,6]) \times I(T_c > 12) \\
Y_6 &= I(T_s \in (6,12]) \times I(T_c > 12) \\
Y_7 &= 1 - \sum_{i=1}^{6} Y_i
\end{aligned}$$

the log-likelihood is now the sum of terms of the form $\ln\left(\sum_{i=1}^{7} Y_i.P_i\right)$.

4 The EasyReg code

The initial model variables are

Z(1) =	I(T_s in (0,1]) x I(T_c in (1,6])
Z(2) =	I(T_s in (0,1]) x I(T_c in (6,12])
Z(3) =	I(T_s in (0,1]) x I(T_c > 12)
Z(4) =	I(T_s in (1,6]) x I(T_c in (6,12])
Z(5) =	I(T_s in (1,6]) x I(T_c > 12)
Z(6) =	I(T_s in (6,12]) x I(T_c > 12)
Z(7) =	G
Z(8) =	AGE
Z(9) =	CHICAGO
Z(10) =	SANDIEGO
Z(11) =	BOSTON
Z(12) =	1

EasyReg automatically adds the constant 1 to the list of initial variables.

First, lets create the parameters, using the "Linear combination" transformation:

New Z	EasyReg parameters	Linear combination of:	Actual value
Z(13) =	b(1)	Z(12)	α_s
Z(14) =	b(2)Z(9)+b(3)Z(10)+b(4)	Z(9),Z(10),Z(12)	$\beta_s' X_s$
Z(15) =	b(5)	Z(12)	α_c
Z(16) =	b(6)Z(11)+b(7)Z(8)+b(8)	Z(11),Z(8),Z(12)	$\beta_c' X_c$

Next, lets make the variables (**??**), (**??**) and (**??**), using the following transformations

New Z	Transformation	Variable(s)	Actual value
Z(17)	Ratio	Z(12),Z(13)	$1/\alpha_s = d$
Z(18)	Multiply	Z(7),Z(16)	$G.\beta_c' X_c$
Z(19)	EXP(z)	Z(18)	$\exp(G.\beta_c' X_c)$
Z(20)	Multiply	Z(17),Z(14)	$\alpha_s^{-1}\beta_s' X_s$
Z(21)	Negative: -z	Z(20)	$-\alpha_s^{-1}\beta_s' X_s$
Z(22)	EXP(z)	Z(21)	$\exp(-\alpha_s^{-1}\beta_s' X_s)$
Z(23)	Subtract	Z(12),Z(19)	$1 - \exp(G.\beta_s' X_s)$
Z(24)	Multiply	Z(15),Z(19)	$\alpha_c.\exp(G.\beta_c' X_c)$
Z(25)	Negative: -z	Z(24)	$-\alpha_c.\exp(G.\beta_c' X_c) = r$
Z(26)	Multiply	Z(15),Z(22),Z(23)	c

In order to make $\exp(p.r)$ for $p = 1, 6, 12$ and $S_s(a|X_s)$ for $a = 1, 6, 12$ we need to create these constants first:

New Z	Transformation	c	Variable	Actual value
Z(27)	c*z with c a constant	1	Z(12)	1
Z(28)	c*z with c a constant	6	Z(12)	6
Z(29)	c*z with c a constant	12	Z(12)	12

We are now able to make $\exp(p.r)$ for $p = 1, 6, 12$ and their differences

New Z	Transformation	Variable(s)	Actual value
Z(30)	Multiply	Z(28),Z(25)	$6.r$
Z(31)	Multiply	Z(29),Z(25)	$12.r$
Z(32)	EXP(z)	Z(25)	$\exp(r)$
Z(33)	EXP(z)	Z(30)	$\exp(6.r)$
Z(34)	EXP(z)	Z(31)	$\exp(12.r)$
Z(35)	Subtract	Z(32),Z(33)	$\exp(r) - \exp(6.r)$
Z(36)	Subtract	Z(33),Z(34)	$\exp(6.r) - \exp(12.r)$

Next, we make $S_s(a|X_s)$ for $a = 1, 6, 12$. Recall that

$$\begin{aligned} S_s(a|X_s) &= \exp\left(-\exp\left(\beta_s' X_s\right) a^{\alpha_s}\right) \\ &= \exp\left(-\exp\left(\beta_s' X_s\right) \exp\left(\alpha_s \ln\left(a\right)\right)\right) \\ &= \exp\left(-\exp\left(\beta_s' X_s + \alpha_s \ln a\right)\right) \end{aligned}$$

so that

$$\begin{aligned} S_s(1|X_s) &= \exp\left(-\exp\left(\beta_s' X_s\right)\right) \\ S_s(6|X_s) &= \exp\left(-\exp\left(\beta_s' X_s + \alpha_s \ln\left(6\right)\right)\right) \\ S_s(12|X_s) &= \exp\left(-\exp\left(\beta_s' X_s + \alpha_s \ln\left(12\right)\right)\right) \end{aligned}$$

New Z	Transformation	Variable(s)	Actual value
Z(37)	LOG(z)	Z(28)	ln(6)
Z(38)	LOG(z)	Z(29)	ln(12)
Z(39)	Multiply	Z(13),Z(37)	$\alpha_s \ln(6)$
Z(40)	Multiply	Z(13),Z(38)	$\alpha_s \ln(12)$
Z(41)	Add up	Z(14),Z(39)	$\beta_s' X_s + \alpha_s \ln(6)$
Z(42)	Add up	Z(14),Z(40)	$\beta_s' X_s + \alpha_s \ln(12)$
Z(43)	EXP(z)	Z(14)	$\exp(\beta_s' X_s)$
Z(44)	EXP(z)	Z(41)	$\exp(\beta_s' X_s + \alpha_s \ln(6))$
Z(45)	EXP(z)	Z(42)	$\exp(\beta_s' X_s + \alpha_s \ln(12))$
Z(46)	Negative: -z	Z(43)	$-\exp(\beta_s' X_s)$
Z(47)	Negative: -z	Z(44)	$-\exp(\beta_s' X_s + \alpha_s \ln(6))$
Z(48)	Negative: -z	Z(45)	$-\exp(\beta_s' X_s + \alpha_s \ln(12))$
Z(49)	EXP(z)	Z(46)	$S_s(1|X_s)$
Z(50)	EXP(z)	Z(47)	$S_s(6|X_s)$
Z(51)	EXP(z)	Z(48)	$S_s(12|X_s)$

We are now able to make the integrals $\int_{S_s(b|X_s)}^{S_s(a|X_s)} \exp\left(-c.\left(\ln\left(1/x\right)\right)^d\right) dx$:

Integral(exp[-c.(ln(1/x))^d]dx|a<x<b)(a,b,c,d)

New Z	Variables a,b,c,d	Actual value
Z(52)	Z(49),Z(12),Z(26),Z(17)	$\int_{S_s(1|X_s)}^{1} \exp\left(-c.\left(\ln\left(1/x\right)\right)^d\right) dx$
Z(53)	Z(50),Z(49),Z(26),Z(17)	$\int_{S_s(6|X_s)}^{S_s(1|X_s)} \exp\left(-c.\left(\ln\left(1/x\right)\right)^d\right) dx$
Z(54)	Z(51),Z(50),Z(26),Z(17)	$\int_{S_s(12|X_s)}^{S_s(6|X_s)} \exp\left(-c.\left(\ln\left(1/x\right)\right)^d\right) dx$

We now have all the building material to make the probabilities $P_i, i = 1, ..., 7$.

New Z	Transformation	Variable(s)	Actual value
Z(55)	Multiply	Z(35),Z(52)	P_1
Z(56)	Multiply	Z(36),Z(52)	P_2
Z(57)	Multiply	Z(34),Z(52)	P_3
Z(58)	Multiply	Z(36),Z(53)	P_4
Z(59)	Multiply	Z(34),Z(53)	P_5
Z(60)	Multiply	Z(34),Z(54)	P_6
Z(61)	Add up	Z(55),Z(56),...,Z(60)	$\sum_{i=1}^{6} P_1$
Z(62)	Subtract	Z(12),Z(61)	P_7

The log-likelihood can now be constructed as follows:

New Z	Transformation	Variable(s)	Actual value
Z(63)	Add up	Z(1),Z(2),...,Z(6)	$\sum_{i=1}^{6} Y_i$
Z(64)	Subtract	Z(12),Z(63)	Y_7
Z(65)	Multiply	Z(1),Z(55)	$Y_1.P_1$
Z(66)	Multiply	Z(2),Z(56)	$Y_2.P_2$
Z(67)	Multiply	Z(3),Z(57)	$Y_3.P_3$
Z(68)	Multiply	Z(4),Z(58)	$Y_4.P_4$
Z(69)	Multiply	Z(5),Z(59)	$Y_5.P_5$
Z(70)	Multiply	Z(6),Z(60)	$Y_6.P_6$
Z(71)	Multiply	Z(64),Z(62)	$Y_7.P_7$
Z(72)	Add up	Z(65),Z(66),...,Z(71)	$\sum_{i=1}^{7} Y_i P_i$
Z(73)	c*z with c=1E-10	Z(12)	0.0000000001
Z(74)	Maximum	Z(72),Z(73)	$\max\left(1E\text{-}10, \sum_{i=1}^{7} Y_i P_i\right)$
Z(75)	LOG(z)	Z(74)	log-likelihood

The steps Z(73) and Z(74) are done for numerical reasons.

5 Reference

Bierens, H. J. and J. R. Carvalho (2007), "Job Search, Conditional Treatment and Recidivism: The Employment Services for Ex-Offenders Program Reconsidered", working paper, PennState University, downloadable from http://econ.la.psu.edu/~hbierens/PAPERS.HTM