PCM predict output examples

Allen Buxton

September 23, 2007

Abstract

note on the prediction for continuous covariate. cureregr Stata function.

Contents

1 example data, description

small non-cleaved cell (SNCC) non-Hodgkins lymphoma (NHL), Stat Med. 2002 Jan 30;21(2):293-312, page 300.

The outcome measure is event-free survival (EFS), which is the minimum time to disease progression, disease relapse, occurrence of a second malignant neoplasm, or death from any cause. Median follow-up was seven years among non-failures, with 83 per cent of these patients followed at least five years. In children, deaths are almost exclusively preceded by disease recurrence or progression, or are a complication of treatment, so that a cured fraction is recognizable without the complication of high background mortality. The objective of this analysis was to determine whether serum lactate dehydrogenase (LDH) ($\leq$500 IU versus $>$500 IU) and bone marrow status (M1 versus (M2,M3)) at diagnosis are independent predictors of cure in these patients.

Herein, an artificial bootstrap sample from NHL SNCC data with the following variables are available: efsyrs (efs years observation time), exit (exit status), age (years), am3 (age - 3 yrs), am6 (age - 6 yrs), am9 (age - 9 yrs), am12 (age - 12 yrs), am15 (age - 15 yrs), l500 (ldh >500 IU), m23 (marrow (M2,M3)), and marldh, the (m23 by l500) interaction.

2 example PCM fit

```
.         cureregr age l500 m23,sc(m23) sh(m23) ///
>         link(lml) distribution(weibull) class(nonm)nolog

         failure _d:  exit == 1
   analysis time _t:  efsyrs
cf: lml, kn: weibull, model: non_mixture
cf_initial_coef: -.5079 pi: .5478
No. of subjects = 250                             Number of obs   =       250
                                                  LR chi2(5)      =     53.83
Log likelihood = -138.50247                       Prob > chi2     =    0.0000
```

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
cure_frac						
age	.0676627	.0246766	2.74	0.006	.0192974	.116028
l500	1.35413	.2706938	5.00	0.000	.8235795	1.88468
m23	.5573143	.2384032	2.34	0.019	.0900525	1.024576
_cons	-2.216853	.3448536	-6.43	0.000	-2.892754	-1.540953
scale						
m23	-.2831445	.2133382	-1.33	0.184	-.7012796	.1349907
_cons	.605446	.1500396	4.04	0.000	.3113737	.8995182
shape						
m23	.2480482	.1504704	1.65	0.099	-.0468684	.5429648
_cons	.0922628	.0929942	0.99	0.321	-.0900025	.274528

3 predict with a continuous covariate, age

Issuing the `predict, at()` command post estimation can result in a lot of output that may not be of much interest. The illustration that follows shows the first 5 and final six covariate patterns of 145.

```
. predict, at(1/4 4.8)
```

(*Continued on next page*)

```
     | age   l500   m23   _freq |
  1. | 1.7      1     0       1 |
```

```
time: 1    S(t)= .666586  se= .061781  ci: (.530073 - .771698)
time: 2    S(t)= .627493  se= .06648   ci: (.483047 - .741955)
time: 3    S(t)= .623299  se= .067167  ci: (.477588 - .739045)
time: 4    S(t)= .622887  se= .067256  ci: (.477003 - .738791)
time: 4.8  S(t)= .622852  se= .067265  ci: (.47695  - .738771)
```

```
     | age   l500   m23   _freq |
  2. | 1.8      1     0       2 |
```

```
time: 1    S(t)= .664753  se= .061549  ci: (.528964 - .769643)
time: 2    S(t)= .625511  se= .066207  ci: (.481871 - .73969)
time: 3    S(t)= .621301  se= .066893  ci: (.476401 - .736764)
time: 4    S(t)= .620888  se= .066982  ci: (.475815 - .736509)
time: 4.8  S(t)= .620853  se= .06699   ci: (.475762 - .736489)
```

```
     | age   l500   m23   _freq |
  3. |   2      0     0       1 |
```

```
time: 1    S(t)= .898653  se= .029985  ci: (.821143 - .943701)
time: 2    S(t)= .884457  se= .033667  ci: (.798169 - .935315)
time: 3    S(t)= .882896  se= .03413   ci: (.795455 - .93446)
time: 4    S(t)= .882742  se= .034183  ci: (.795165 - .934384)
time: 4.8  S(t)= .882729  se= .034187  ci: (.795139 - .934378)
```

```
     | age   l500   m23   _freq |
  4. |   2      1     0       1 |
```

```
time: 1    S(t)= .661065  se= .061083  ci: (.526714 - .765502)
time: 2    S(t)= .621525  se= .065661  ci: (.479489 - .73513)
time: 3    S(t)= .617286  se= .066344  ci: (.473996 - .732171)
time: 4    S(t)= .61687   se= .066433  ci: (.473406 - .731914)
time: 4.8  S(t)= .616834  se= .066441  ci: (.473353 - .731894)
```

```
     | age   l500   m23   _freq |
  5. | 2.2      1     1       4 |
```

```
time: 1    S(t)= .507776  se= .089906  ci: (.322737 - .666224)
time: 2    S(t)= .430922  se= .09709   ci: (.24112  - .607621)
time: 3    S(t)= .425476  se= .098287  ci: (.234203 - .604667)
time: 4    S(t)= .425252  se= .098375  ci: (.233846 - .604611)
time: 4.8  S(t)= .425246  se= .09838   ci: (.233836 - .604612)
```

(*output omitted*)

```
     | age   l500   m23   _freq |
140. |  17      0     0       2 |
```

```
time: 1    S(t)= .744648  se= .062609  ci: (.597136 - .844846)
time: 2    S(t)= .712641  se= .069198  ci: (.55204  - .82434)
time: 3    S(t)= .709175  se= .070087  ci: (.546716 - .822356)
time: 4    S(t)= .708835  se= .070194  ci: (.546141 - .822187)
time: 4.8  S(t)= .708805  se= .070204  ci: (.546089 - .822174)
```

(*Continued on next page*)

```
      +-------------------------------+
      |  age   l500   m23   _freq     |
      |-------------------------------|
141.  |   17      1     0        1    |
      +-------------------------------+

time: 1    S(t)= .319167  se= .096782  ci: (.146354 - .50728)
time: 2    S(t)= .269223  se= .094944  ci: (.108379 - .460752)
time: 3    S(t)= .264186  se= .094913  ci: (.104436 - .456447)
time: 4    S(t)= .263695  se= .094933  ci: (.104022 - .456077)
time: 4.8  S(t)= .263653  se= .094936  ci: (.103985 - .456048)

      +-------------------------------+
      |  age   l500   m23   _freq     |
      |-------------------------------|
142.  |  17.2     0     0        1    |
      +-------------------------------+

time: 1    S(t)= .741663  se= .06384   ci: (.59122  - .843711)
time: 2    S(t)= .709359  se= .070521  ci: (.545719 - .823083)
time: 3    S(t)= .705862  se= .07142   ci: (.540356 - .821086)
time: 4    S(t)= .705519  se= .071528  ci: (.539778 - .820915)
time: 4.8  S(t)= .70549   se= .071538  ci: (.539725 - .820902)

      +-------------------------------+
      |  age   l500   m23   _freq     |
      |-------------------------------|
143.  |  17.6     1     1        1    |
      +-------------------------------+

time: 1    S(t)= .146423  se= .068138  ci: (.045568 - .302665)
time: 2    S(t)= .091951  se= .053937  ci: (.020994 - .228974)
time: 3    S(t)= .088695  se= .053337  ci: (.019436 - .225532)
time: 4    S(t)= .088563  se= .053334  ci: (.019358 - .225463)
time: 4.8  S(t)= .088559  se= .053335  ci: (.019355 - .225463)

      +-------------------------------+
      |  age   l500   m23   _freq     |
      |-------------------------------|
144.  |  17.9     1     1        1    |
      +-------------------------------+

time: 1    S(t)= .140767  se= .068359  ci: (.041342 - .299206)
time: 2    S(t)= .08756   se= .053552  ci: (.018608 - .225658)
time: 3    S(t)= .084397  se= .052892  ci: (.017194 - .222193)
time: 4    S(t)= .084268  se= .052886  ci: (.017123 - .22212)
time: 4.8  S(t)= .084265  se= .052887  ci: (.017121 - .22212)

      +-------------------------------+
      |  age   l500   m23   _freq     |
      |-------------------------------|
145.  |  18.7     0     0        1    |
      +-------------------------------+

time: 1    S(t)= .718359  se= .07392   ci: (.544107 - .83545)
time: 2    S(t)= .68381   se= .081302  ci: (.495749 - .813937)
time: 3    S(t)= .68008   se= .082269  ci: (.490122 - .811842)
time: 4    S(t)= .679713  se= .082384  ci: (.489519 - .811662)
time: 4.8  S(t)= .679682  se= .082394  ci: (.489464 - .811648)
```

4 options to extensive predict output

4.1 option 1, summarize

```
.         summarize age , detail
                        age (years)
-------------------------------------------------------------
      Percentiles      Smallest
 1%           1.8           1.7
 5%           2.8           1.8
10%          3.15           1.8       Obs                 250
25%           5.6             2       Sum of Wgt.         250
50%          9.05                     Mean              9.006
                        Largest       Std. Dev.      4.174151
75%          12.2          17.2
90%            15          17.6       Variance       17.42354
95%          16.2          17.9       Skewness       .1954961
99%          17.6          18.7       Kurtosis       2.139681
.         foreach i of numlist 1 5 10 25 50 75 90 95 99 {
  2.                local age`i' = `r(p`i')'
  3.                }
```

get a more meaningful predit summary

```
.         foreach i of numlist 5 50 95 {
  2.                qui replace age=`age`i''
  3.                di `""'
  4.                di `"{res}===== age percentile = `i' ======{txt}"'
  5.                predict , at(1/4 4.8)
  6.                di `"{res}==end age percentile = `i' ======{txt}"'
  7.                }
===== age percentile = 5 =====

     | age   l500   m23   _freq |
  1. | 2.8      0     0      90 |

time: 1    S(t)= .893328  se= .030183  ci: (.816361 - .939213)
time: 2    S(t)= .878438  se= .033861  ci: (.792817 - .930198)
time: 3    S(t)= .876801  se= .034331  ci: (.790025 - .929284)
time: 4    S(t)= .876639  se= .034385  ci: (.789725 - .929203)
time: 4.8  S(t)= .876626  se= .034389  ci: (.789698 - .929197)

     | age   l500   m23   _freq |
  2. | 2.8      0     1      14 |

time: 1    S(t)= .833421  se= .057507  ci: (.68198  - .916909)
time: 2    S(t)= .797446  se= .068972  ci: (.61961  - .898501)
time: 3    S(t)= .794724  se= .0701    ci: (.614101 - .897389)
time: 4    S(t)= .794611  se= .070159  ci: (.613828 - .897358)
time: 4.8  S(t)= .794608  se= .070163  ci: (.61382  - .897357)

     | age   l500   m23   _freq |
  3. | 2.8      1     0      87 |

time: 1    S(t)= .64602   se= .059223  ci: (.517278 - .748555)
time: 2    S(t)= .6053    se= .063487  ci: (.469505 - .716522)
time: 3    S(t)= .600942  se= .06416   ci: (.463915 - .713437)
time: 4    S(t)= .600515  se= .06425   ci: (.463312 - .713172)
time: 4.8  S(t)= .600478  se= .064259  ci: (.463257 - .713151)
```

(*Continued on next page*)

```
     | age   l500   m23   _freq |
  4. | 2.8      1     1      59 |
```

```
time: 1    S(t)= .493716  se= .087098  ci: (.316017 - .648927)
time: 2    S(t)= .41615   se= .093579  ci: (.234717 - .588423)
time: 3    S(t)= .410675  se= .094759  ci: (.227794 - .585438)
time: 4    S(t)= .41045   se= .094848  ci: (.227435 - .585386)
time: 4.8  S(t)= .410444  se= .094853  ci: (.227424 - .585386)
==end age percentile = 5 =====
===== age percentile = 50 =====
```

```
     |  age   l500   m23   _freq |
  1. | 9.05      0     0      90 |
```

```
time: 1    S(t)= .84183   se= .034228  ci: (.7607   - .897282)
time: 2    S(t)= .820507  se= .038214  ci: (.730647 - .88275)
time: 3    S(t)= .818174  se= .038785  ci: (.72699  - .881342)
time: 4    S(t)= .817944  se= .038857  ci: (.726589 - .881224)
time: 4.8  S(t)= .817925  se= .038862  ci: (.726552 - .881215)
```

```
     |  age   l500   m23   _freq |
  2. | 9.05      0     1      14 |
```

```
time: 1    S(t)= .757199  se= .064533  ci: (.60225  - .858515)
time: 2    S(t)= .70788   se= .075719  ci: (.53056  - .828355)
time: 3    S(t)= .704195  se= .076966  ci: (.524151 - .826632)
time: 4    S(t)= .704043  se= .077039  ci: (.523826 - .82659)
time: 4.8  S(t)= .704039  se= .077043  ci: (.523816 - .826589)
```

```
     |  age   l500   m23   _freq |
  3. | 9.05      1     0      87 |
```

```
time: 1    S(t)= .513292  se= .054577  ci: (.401903 - .613897)
time: 2    S(t)= .464736  se= .057391  ci: (.349619 - .571925)
time: 3    S(t)= .459639  se= .058067  ci: (.343387 - .568206)
time: 4    S(t)= .45914   se= .058172  ci: (.342699 - .567911)
time: 4.8  S(t)= .459098  se= .058182  ci: (.342636 - .56789)
```

```
     |  age   l500   m23   _freq |
  4. | 9.05      1     1      59 |
```

```
time: 1    S(t)= .340512  se= .06174   ci: (.223509 - .460888)
time: 2    S(t)= .262323  se= .061785  ci: (.151157 - .38761)
time: 3    S(t)= .257073  se= .062775  ci: (.144836 - .384806)
time: 4    S(t)= .256857  se= .062869  ci: (.144492 - .384804)
time: 4.8  S(t)= .256852  se= .062874  ci: (.144481 - .384807)
==end age percentile = 50 =====
```

(*Continued on next page*)

```
===== age percentile = 95 =====
```

	age	l500	m23	_freq
1.	16.2	0	0	90

```
time: 1    S(t)= .756307  se= .057953  ci: (.619903 - .849472)
time: 2    S(t)= .725476  se= .064176  ci: (.576461 - .829466)
time: 3    S(t)= .722133  se= .065028  ci: (.5713   - .827535)
time: 4    S(t)= .721805  se= .065131  ci: (.570742 - .827371)
time: 4.8  S(t)= .721777  se= .065141  ci: (.570691 - .827359)
```

	age	l500	m23	_freq
2.	16.2	0	1	14

```
time: 1    S(t)= .636873  se= .088851  ci: (.437319 - .781826)
time: 2    S(t)= .570957  se= .099567  ci: (.356541 - .737448)
time: 3    S(t)= .566143  se= .100815  ci: (.349696 - .734885)
time: 4    S(t)= .565945  se= .100893  ci: (.349352 - .734818)
time: 4.8  S(t)= .56594   se= .100897  ci: (.349342 - .734817)
```

	age	l500	m23	_freq
3.	16.2	1	0	87

```
time: 1    S(t)= .338962  se= .091663  ci: (.171051 - .515387)
time: 2    S(t)= .288496  se= .090789  ci: (.129811 - .469144)
time: 3    S(t)= .283381  se= .090885  ci: (.125447 - .464888)
time: 4    S(t)= .282882  se= .09092   ci: (.124986 - .464524)
time: 4.8  S(t)= .282839  se= .090924  ci: (.124945 - .464496)
```

	age	l500	m23	_freq
4.	16.2	1	1	59

```
time: 1    S(t)= .174189  se= .066417  ci: (.068553 - .319966)
time: 2    S(t)= .114086  se= .0551    ci: (.034826 - .245712)
time: 3    S(t)= .110405  se= .054793  ci: (.032503 - .242396)
time: 4    S(t)= .110255  se= .054808  ci: (.032384 - .242342)
time: 4.8  S(t)= .110251  se= .054811  ci: (.03238  - .242342)
==end age percentile = 95 =====
```

4.2 option 2, use egen

```
.         foreach i of numlist 33 68 {
  2.                 egen temp0=pctile(age_copy),p(`i')
  3.                 qui replace age=temp0
  4.                 drop temp0
  5.                 di `""'
  6.                 di `"{res}===== age percentile = `i' =====`txt'"'
  7.                 predict , at(1/4 4.8)
  8.                 di `"{res}==end age percentile = `i' =====`txt'"'
  9.                 }
===== age percentile = 33 =====

     | age   l500   m23   _freq |
     |------------------------- |
  1. | 6.9      0     0      90 |

time: 1    S(t)= .861684  se= .03196   ci: (.784592 - .912696)
time: 2    S(t)= .84278   se= .035725  ci: (.757286 - .90011)
time: 3    S(t)= .840708  se= .036248  ci: (.753984 - .898873)
time: 4    S(t)= .840504  se= .036311  ci: (.753624 - .898767)
time: 4.8  S(t)= .840487  se= .036316  ci: (.75359  - .89876)

     | age   l500   m23   _freq |
     |------------------------- |
  2. | 6.9      0     1      14 |

time: 1    S(t)= .786255  se= .06143   ci: (.634707 - .88055)
time: 2    S(t)= .741777  se= .072732  ci: (.566433 - .854723)
time: 3    S(t)= .738437  se= .07394   ci: (.560334 - .853224)
time: 4    S(t)= .738299  se= .074008  ci: (.560026 - .853185)
time: 4.8  S(t)= .738296  se= .074012  ci: (.560017 - .853185)

     | age   l500   m23   _freq |
     |------------------------- |
  3. | 6.9      1     0      87 |

time: 1    S(t)= .561795  se= .052685  ci: (.452443 - .657554)
time: 2    S(t)= .515541  se= .055796  ci: (.401498 - .618151)
time: 3    S(t)= .510648  se= .056476  ci: (.395392 - .614594)
time: 4    S(t)= .510169  se= .056578  ci: (.394719 - .614308)
time: 4.8  S(t)= .510128  se= .056589  ci: (.394657 - .614287)

     | age   l500   m23   _freq |
     |------------------------- |
  4. | 6.9      1     1      59 |

time: 1    S(t)= .393982  se= .068365  ci: (.261337 - .523863)
time: 2    S(t)= .314426  se= .070528  ci: (.184182 - .453277)
time: 3    S(t)= .308978  se= .071624  ci: (.177423 - .450359)
time: 4    S(t)= .308754  se= .07172   ci: (.177059 - .450342)
time: 4.8  S(t)= .308749  se= .071725  ci: (.177048 - .450344)
==end age percentile = 33 =====
```

(*Continued on next page*)

```
===== age percentile = 68 =====
```

	age	l500	m23	_freq
1.	10.9	0	0	90

```
time: 1    S(t)= .822722  se= .037596  ci: (.734331 - .883987)
time: 2    S(t)= .799144  se= .041933  ci: (.70138  - .867857)
time: 3    S(t)= .79657   se= .042562  ci: (.697368 - .866307)
time: 4    S(t)= .796316  se= .04264   ci: (.696928 - .866177)
time: 4.8  S(t)= .796295  se= .042647  ci: (.696887 - .866168)
```

	age	l500	m23	_freq
2.	10.9	0	1	14

```
time: 1    S(t)= .72963   se= .06837   ci: (.568656 - .838598)
time: 2    S(t)= .676008  se= .079468  ci: (.493985 - .80462)
time: 3    S(t)= .672021  se= .080742  ci: (.48734  - .802693)
time: 4    S(t)= .671857  se= .080818  ci: (.487002 - .802646)
time: 4.8  S(t)= .671853  se= .080822  ci: (.486992 - .802646)
```

	age	l500	m23	_freq
3.	10.9	1	0	87

```
time: 1    S(t)= .469616  se= .060486  ci: (.348001 - .582032)
time: 2    S(t)= .419594  se= .062973  ci: (.295653 - .538508)
time: 3    S(t)= .414382  se= .063594  ci: (.289533 - .534647)
time: 4    S(t)= .413872  se= .063693  ci: (.288863 - .534338)
time: 4.8  S(t)= .413829  se= .063703  ci: (.288801 - .534315)
```

	age	l500	m23	_freq
4.	10.9	1	1	59

```
time: 1    S(t)= .294947  se= .059548  ci: (.18483  - .41355)
time: 2    S(t)= .219452  se= .057455  ci: (.119165 - .339162)
time: 3    S(t)= .214481  se= .058239  ci: (.113574 - .336357)
time: 4    S(t)= .214277  se= .058321  ci: (.11327  - .336356)
time: 4.8  S(t)= .214272  se= .058326  ci: (.113261 - .336359)
==end age percentile = 68 =====
```

4.3 option, other constants

Any constant can be temporarily substituted as for *age* in the examples above prior to issuing the `predict` command. In some cases, the average, standard score, or some function of the continuous variable may be of interest.

Of course, as with any regression, distinct groups and associated factors can be constructed to handle continuous variables.

5 do file

```
use "example_stata.dta", clear
stset efsyrs, failure(exit==1)

which cureregr

/* make a copy of age since predict needs the same variable names as the */
/* model is fit with */
gen age_copy=age
cureregr age l500 m23,sc(m23) sh(m23) ///
link(lml) distribution(weibull) class(nonm)nolog

/* with option 'detail' summarize returns the following percentiles: */
/* 1 5 10 25 50 75 90 95 99 */
summarize age , detail
foreach i of numlist 1 5 10 25 50 75 90 95 99 {
local age`i' = `r(p`i')'
}
foreach i of numlist 5 50 95 {
qui replace age=`age`i''
di `""'
di `"{res}===== age percentile = `i' =====`{txt}"'
predict , at(1(0.5)4)
di `"{res}==end age percentile = `i' =====`{txt}"'
}

/* alternative method for above percentiles or for percentiles not above, */
/* such as the 33rd and 68th percentiles, use egen */
foreach i of numlist 33 68 {
egen temp0=pctile(age_copy),p(`i')
qui replace age=temp0
drop temp0
di `""'
di `"{res}===== age percentile = `i' =====`{txt}"'
predict , at(1(0.5)4)
di `"{res}==end age percentile = `i' =====`{txt}"'
}
replace age=age_copy
drop age_copy
```