Laplace Regression
A Practical Introduction

Andrea Bellavia

Institute of Environmental Medicine
Karolinska Institutet
Laplace Regression

Instead of looking at the risk of experiencing an event we focus on the time to event. Given a survival percentile we look at absolute survival differences.
Example - Data Introduction

- Study population: 71,706 men and women from COSM and SMC ’97
- Exposure: Fruit and vegetables consumption, servings/day
- Outcome: time to death
- Potential confounders: age, gender, smoking, alcohol, physical activity, bmi, energy intake, education

<table>
<thead>
<tr>
<th>Categories of FV consumption</th>
<th>Cases</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;=1</td>
<td>1,093</td>
<td>3,663</td>
</tr>
<tr>
<td>1-3</td>
<td>3,826</td>
<td>20,283</td>
</tr>
<tr>
<td>3-5</td>
<td>3,463</td>
<td>23,703</td>
</tr>
<tr>
<td>5-7</td>
<td>1,827</td>
<td>13,926</td>
</tr>
<tr>
<td>&gt;7</td>
<td>1,230</td>
<td>10,131</td>
</tr>
<tr>
<td>Total</td>
<td>11,439</td>
<td>71,706</td>
</tr>
</tbody>
</table>
Cox Regression

- Presenting both Cox and Laplace is always recommended, especially if the event is not certain.
- `stset` the data. Using follow-up as primary time scale facilitates the comparison with Laplace. But check if age is the same.

```stata
stset doxm, fail(dead) origin(doe) scale(30.4166666667)
gen timeacm=_t
```

```stata
xi: stcox i.fvc2 i.age7
```

| _t | Haz. Ratio | Std. Err. | z    | P>|z|  | [95% Conf. Interval] |
|----|------------|-----------|------|------|----------------------|
| _Ifvc2_1 | 1.673878 | .0582794 | 14.80 | 0.000 | 1.563463 - 1.792092 |
| _Ifvc2_2 | 1.19692  | .0281034 | 7.66  | 0.000 | 1.143087 - 1.253289 |
| _Ifvc2_4 | .88547   | .0256063 | -4.21 | 0.000 | .8366785 - .9371069 |
| _Ifvc2_5 | .8416861 | .0279458 | -5.19 | 0.000 | .7886575 - .8982803 |

```stata
xi: stcox i.fvc2 i.age7 i.packyc i.alcoholc5 pa97 i.bmic4 energy i.educat i.data
```

| _t | Haz. Ratio | Std. Err. | z    | P>|z|  | [95% Conf. Interval] |
|----|------------|-----------|------|------|----------------------|
| _Ifvc2_1 | 1.230924 | .0669806 | 3.82  | 0.000 | 1.106403 - 1.36946 |
| _Ifvc2_2 | 1.078128 | .0328063 | 2.47  | 0.013 | 1.015708 - 1.144383 |
| _Ifvc2_4 | .9262542 | .0331951 | -2.14 | 0.033 | .8634254 - .9936548 |
| _Ifvc2_5 | .9019392 | .0381979 | -2.44 | 0.015 | .8300957 - .9800006 |
From Cox to Laplace

HR=1.23. Compared to the median category of FV consumption (3-5 servings/day), those who consumed less than 1 serving per day had 23% higher “mortality rate”.

sts graph, by(fvc2) ylabel(0.7(0.1)1)

We now want to provide a multivariable adjusted estimate of the horizontal line (differences in the time by which the first 10% of the population has died, according to levels of FV consumption).
Laplace – running the model

\[ xi: \text{laplace} \text{timeacm} \text{i.fvc2} \text{i.age7}, \text{failure(dead)} q(10) \]

Time variable (months, years..)

Covariates

Main Exposure

Failure Variable

Focus on 10th percentile.

n.b. You need to have installed the package \text{laplace} on Stata:

\text{net install laplace, from(http://www.imm.ki.se/biostatistics/stata) replace}
Laplace – Interpreting the Results

We focused on the 10th percentile of survival, the time by which the first 10% of the population has died. After adjustment for potential confounders we observed that those who ate less than 1 serving per day of FV lived an average of 17 months shorter than those in the median category of consumption (10th PD = -17; 95% CI: -25, -9).
Laplace – Going Further

• We can go further in the analysis, since these results suggest a non-linear association

• A more flexible approach (splines) could much better describe the association

This is what we did here:

*Fruit and vegetable consumption and all-cause mortality: a dose-response analysis*¹-³
Andrea Bellavia, Susanna C Larsson, Matteo Bottai, Alicja Wolk, and Nicola Orsini

• When using Laplace regression, extension of the model is extremely easy.

• Time is directly modeled as a function of the covariates, and no transformations are therefore required (log, exp..)

We generate splines (right restricted):

```plaintext
generate fvp=-fv
gen fvpn1=max(0, 8-fv)^3
gen fvpn2=max(0, 5-fv)^3
gen fvpn3=max(0, 3-fv)^3
```
Laplace with Splines

The output you get from this command can not be directly interpreted. A post estimation command is required:

```
qui: levelsof fv
xb1c fvp fv pn*, cov(fv) ref(5) at(0,0.5,1,2,3,4,5,8)
```

**Stata output**

<table>
<thead>
<tr>
<th>fvp</th>
<th>xb</th>
<th>(95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-37.39</td>
<td>(-58.44--16.34)</td>
</tr>
<tr>
<td>0.5</td>
<td>-25.55</td>
<td>(-36.00--15.11)</td>
</tr>
<tr>
<td>1</td>
<td>-16.90</td>
<td>(-23.03--10.77)</td>
</tr>
<tr>
<td>2</td>
<td>-6.89</td>
<td>(-11.44--2.33)</td>
</tr>
<tr>
<td>3</td>
<td>-2.85</td>
<td>(-6.29--0.59)</td>
</tr>
<tr>
<td>4</td>
<td>-0.95</td>
<td>(-3.24--1.33)</td>
</tr>
<tr>
<td>5</td>
<td>0.00</td>
<td>(0.00--0.00)</td>
</tr>
</tbody>
</table>

**Comment in the paper:** “Compared with an FV consumption of 5 servings/d, lower levels of consumption were progressively associated with shorter survival up to 3 y for those who never consumed FV daily (PD: -37 mo; 95% CI: -58, -16 mo).”
Graphing the Dose-Response (1)

- With a minor change to the post-estimation command we get the entire curve

```
qui: levelsof fv
xblc fvp fvnp*, cov(fv) ref(5) at(`r(levels)') line
```

![Graph of Dose-Response](image)
Graphing the Dose-Response (2)

- ... that can be edited for publication (commands not shown).
References


• http://www.imm.ki.se/biostatistics/laplace/