Additive and Multiplicative Laplace Models for Survival Percentiles

Andrea Bellavia

Unit of Biostatistics, Unit of Nutritional Epidemiology
Institute of Environmental Medicine
Karolinska Institutet, Stockholm
http://www.imm.ki.se/biostatistics
andrea.bellavia@ki.se

Nordic and Baltic Stata Users Group meeting
September 4th, 2015

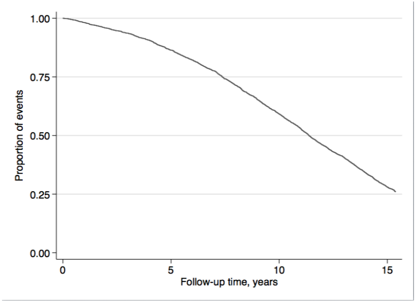
Acknowledgements

- Nicola Orsini
- Matteo Bottai
- Paolo Frumento
- Andrea Discacciati
- Alicja Wolk

Survival Percentiles

- In time-to-event analysis we define the *p*th survival percentile as the time *t* by which p% of the study population has experienced the event of interest, and (1-p)% have not
- Example - The minimal value of T is 0, when everyone is alive. The time by which 50% of the participants have died is called 50th survival percentile, or median survival
- In the same way we can define all survival percentiles
- Survival percentiles are depicted in the survival curve

Survival Percentiles (2)



- Evaluating survival percentiles changes the prospective
- In the classical approach the time is fixed and the probability (risk) is evaluated. Here the probability is fixed and the time by which that proportion of cases is achieved is evaluated

Modelling Survival Percentiles

- At the univariable level percentiles can be estimated with the non-parametric Kaplan-Meier method
- In Stata: `sts graph`, `stqkm`.
- `stqkm` provides differences in survival percentiles with CI. It can be installed by typing:

```
net install stqkm, ///
from(http://www.imm.ki.se/biostatistics/stata)
```

Adjusted survival percentiles

- Common situation in epidemiology
- We have two main approaches:
- 1) Estimate a multivariable parametric (AFT, flexible parametric survival), or semi-parametric (COX) model. Back-calculate the survival function. Derive adjusted survival percentiles
- Computational and mathematical complexity, plus relying on the original model assumption, limit this application

Adjusted survival percentiles

- 2) Quantile regression for censored data
- Recent developments (Powell, Portnoy, Peng-Huang)
- Bottai & Zhang introduced Laplace regression in 2010
- Recent developments have largely extended the potentiality and advantages of this method

Laplace regression

- When the time variables T_i may be censored we observe the covariates x_i, $y_i = \min(t_i, c_i)$, and $d_i = I(t_i \leq c_i)$
- The aim is to estimate the τ^{th} conditional quantile of T_i
- Given a quantile τ, a response variable Y, and a set of covariates x, a Laplace regression model establishes a linear relationship between a given percentile of T and a set of covariates

$$t_i(\tau) = x_i' \beta(\tau) + \sigma_i(\tau) u_i$$

- u_i follows the Asymmetric Laplace distribution
- Estimation is conducted via gradient search algorithm (Bottai et al. 2014), and standard errors are preferrabily estimated via bootstrap

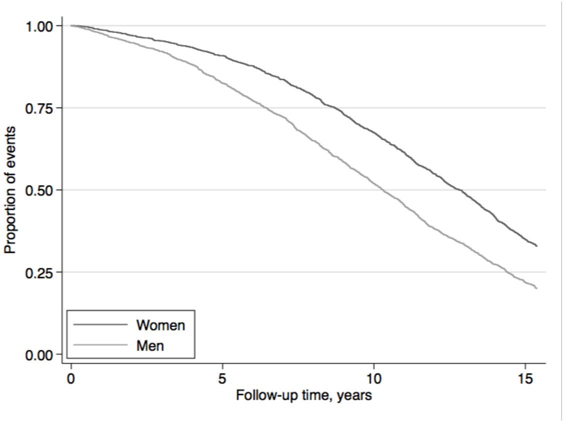
Additive Laplace regression

- In its basic form, a Laplace regression model establishes a linear association between a predictor E and the pth survival percentile of the time variable T

$$T(p|E = e) = \beta_{p0} + \beta_{p1} \cdot e$$

- βs estimated from an additive Laplace regression are interpreted in terms of survival percentile differences (PD), absolute differences in time by which the chosen proportion of cases is achieved.

Example



50th PD=2.2 - Median survival is 2 years longer for women

In Stata

The program can be installed from:

```
net install laplace, ///
from(http://www.imm.ki.se/biostatistics/stata)
```

Example

```
sysuse cancer, clear
xi:  laplace studytime i.drug, q(.5) fail(died)
```

Equivariance to Monotonic Transformation

- Thanks to a peculiar property of the quantiles, the definition of a multiplicative model for survival percentiles is straightforward
- This property is defined as equivariance to monotonic transformation (EMT): let h be a non-decreasing function, then for any Y

$$Q_{h(Y)}(\tau) = h(Q_Y(\tau))$$

- In words, for any random variable T the quantiles of the transformed random variable $h(T)$ are the transformed quantiles of the original T

Multiplicative Laplace regression

- To define a multiplicative model for survival percentiles the most intuitive approach is to specify a model that is linear on the logarithm of time
- The property of EMT assures that this can be achieved by simply operating a logarithmic transformation on the original time variable and by fitting a linear model on the logarithm of time

$$\log[T(p|E=e)] = \beta^*_{p0} + \beta^*_{p1} \cdot e$$

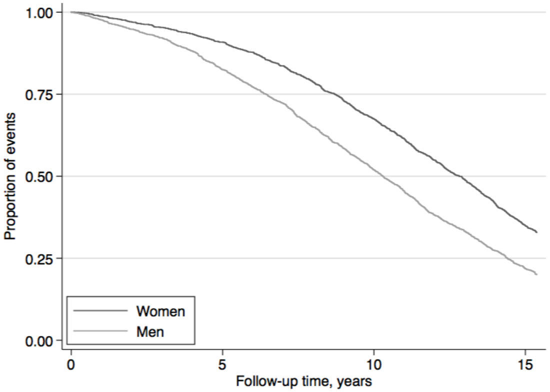
Coefficients' interpretation

- βs estimated from this model do not have a simple interpretation
- However, it is possible to operate an exponential transformation to go back on the original time scale

$$T(p|E = e) = \exp(\beta^*_{p0} + \beta^*_{p1} \cdot e) = \exp(\beta^*_{p0}) \cdot \exp(\beta^*_{p1} \cdot e)$$

- $\exp(\beta^*_{p1})$ can be interpreted as percentile ratio (PR) associated with the exposure, and shows how much faster/slower exposed participants attain the fixed proportion of p% of cases

Example



50th PR=1.22 - Median survival is achieved 22% faster by men

In Stata

The estimation of a multiplicative model for the *p*th survival percentile can be achieved in Stata by including the option `link` in the `laplace` command

Example

```
sysuse cancer, clear
xi:  laplace studytime i.drug, q(.5) fail(died) ///
link(log)
```

Example

- Mortality data from 15.000 subjects
- 8400 participants (58%) died in 15 years of follow-up
- We focus on the impact of smoking on median survival adjusting for baseline age

Additive Model

$T(50) = \beta_0 + \beta_1 \cdot smoking + \beta_2 \cdot age$

```
laplace _t smoking age , q(50) fail(_d)
```

Multiplicative model

$\log[T(50)] = \beta_0^* + \beta_1^* \cdot smoking + \beta_2^* \cdot age$

```
laplace _t smoking age , q(50) fail(_d) link(log)
```

Example - Results

Percentile Difference

The 50th PD (difference in median survival) is estimated by β_1
50th PD= -2.6 years, 95% CI: -3.0, -2.3

Percentile Ratio

The 50th PR (median ratio) is estimated by $\exp(\beta_1^*)$
50th PR=0.79, 95% CI: 0.76, 0.81

Median survival was attained 21% slower in nonsmokers than in smokers. This acceleration resulted in a median survival difference of 2.6 years

Interaction in time-to-event analysis

- Statistical interaction can be evaluated on the additive or the multiplicative scale, and presentation of both scales is recommended
- In survival analysis, because of the popularity of Cox regression, the multiplicative scale alone is usually presented
- We defined the concept of interaction in the context of survival percentiles and presented how to evaluate additive and multiplicative interaction (*Epidemiology*, 2016, accepted for publication)

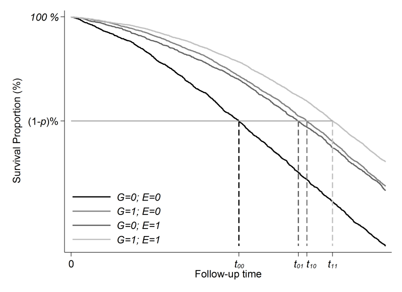
Interaction in the context of survival percentiles

Additive interaction

$$I_{add} = (t_{11} - t_{00}) - [(t_{10} - t_{00}) + (t_{01} - t_{00})]$$

Multiplicative interaction

$$I_{mul} = (t_{11} \cdot t_{00})/(t_{10} \cdot t_{01})$$



Model-based additive and multiplicative interaction

- Inclusion of a product term between two predictors G and E in an additive and multiplicative Laplace model will serve as a test for additive and multiplicative interaction, respectively

Additive interaction

$$T(p|G = g, E = e) = \beta_{p0} + \beta_{p1} \cdot g + \beta_{p2} \cdot e + \beta_{p3} \cdot g \cdot e$$

Multiplicative interaction

$$\log[T(p|G = g, E = e)] = \beta^*_{p0} + \beta^*_{p1} \cdot g + \beta^*_{p2} \cdot e + \beta^*_{p3} \cdot g \cdot e$$

- β_{p3} and $\exp(\beta^*_{p3})$ will test for the presence of additive and multiplicative interaction between G and E

Example

We evaluate the interaction between smoking and education in predicting median survival

Additive Model

$$T(50) = \beta_0 + \beta_1 \cdot smoke + \beta_2 \cdot educat + \beta_3 \cdot smoke \cdot educat$$

```
laplace _t smoke educat inter , q(50) fail(_d)
```

Multiplicative model

$$\log[T(50)] = \beta_0^* + \beta_1^* \cdot smoke + \beta_2^* \cdot educat + \beta_3^* \cdot smoke \cdot educat$$

```
laplace _t smoke educat inter, q(50) fail(_d) ///
link(log)
```

Example - Results

Additive Interaction

Interaction on the additive scale is estimated by β_3

I_{add}= 2.1 years, 95% CI: 1.2, 2.9

Multiplicative Interaction

Interaction on the multiplicative scale is estimated by $\exp(\beta_3^*)$

I_{mul}=1.08, 95% CI: 1.00, 1.17

Summary

- Survival percentiles are defined as the time points by which specific proportion of events are achieved
- Statistical models for survival percentiles, such as Laplace regression, offer all modelling advantages such as multivariable adjustment and interaction assessment
- Thanks to properties of the quantiles, the Laplace models can be defined in both the additive and multiplicative scales
- The additive and multiplicative Laplace models estimate survival percentile differences (absolute measures) and percentile ratios (relative measures), respectively
- An additional importante advantage is that inclusion of product terms in the additive and multiplicative models will serve as tests for additive and multiplicative interactions in the metric of time

References on Laplace regression

- Bottai M, Zhang J. *Laplace regression with censored data.* Biometrical Journal. 2010
- Orsini N et al. *Evaluating percentiles of survival.* Epidemiology. 2012
- Bottai M, Orsini N. *A command for Laplace regression.* Stata Journal. 2013
- Bottai M, Orsini N, Geraci M *A gradient search maximization algorithm for the asymmetric Laplace likelihood*. Journal of Statistical Computation and Simulation. 2014
- Bellavia A. et al. *Using Laplace Regression to Model and Predict Percentiles of Age at Death When Age Is the Primary Time Scale* Am Journal of Epi. 2015
- Bellavia A., Bottai M., Orsini N. *Evaluating additive interaction using survival percentiles*. Epidemiology, 2016 - In press
- Bellavia A., Bottai M., Orsini N. *Survival percentile ratios*. Submitted