



**AgEcon** SEARCH  
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

*The World's Largest Open Access Agricultural & Applied Economics Digital Library*

**This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.**

**Help ensure our sustainability.**

Give to AgEcon Search

AgEcon Search

<http://ageconsearch.umn.edu>

[aesearch@umn.edu](mailto:aesearch@umn.edu)

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

# THE STATA JOURNAL

## Editors

H. JOSEPH NEWTON  
Department of Statistics  
Texas A&M University  
College Station, Texas  
editors@stata-journal.com

NICHOLAS J. COX  
Department of Geography  
Durham University  
Durham, UK  
editors@stata-journal.com

## Associate Editors

CHRISTOPHER F. BAUM, Boston College  
NATHANIEL BECK, New York University  
RINO BELLOCCO, Karolinska Institutet, Sweden, and  
University of Milano-Bicocca, Italy  
MAARTEN L. BUIS, WZB, Germany  
A. COLIN CAMERON, University of California–Davis  
MARIO A. CLEVES, University of Arkansas for  
Medical Sciences  
WILLIAM D. DUPONT, Vanderbilt University  
PHILIP ENDER, University of California–Los Angeles  
DAVID EPSTEIN, Columbia University  
ALLAN GREGORY, Queen's University  
JAMES HARDIN, University of South Carolina  
BEN JANN, University of Bern, Switzerland  
STEPHEN JENKINS, London School of Economics and  
Political Science  
ULRICH KOHLER, University of Potsdam, Germany

FRAUKE KREUTER, Univ. of Maryland–College Park  
PETER A. LACHENBRUCH, Oregon State University  
JENS LAURITSEN, Odense University Hospital  
STANLEY LEMESHOW, Ohio State University  
J. SCOTT LONG, Indiana University  
ROGER NEWSON, Imperial College, London  
AUSTIN NICHOLS, Urban Institute, Washington DC  
MARCELLO PAGANO, Harvard School of Public Health  
SOPHIA RABE-HESKETH, Univ. of California–Berkeley  
J. PATRICK ROYSTON, MRC Clinical Trials Unit,  
London  
PHILIP RYAN, University of Adelaide  
MARK E. SCHAFFER, Heriot-Watt Univ., Edinburgh  
JEROEN WEESIE, Utrecht University  
IAN WHITE, MRC Biostatistics Unit, Cambridge  
NICHOLAS J. G. WINTER, University of Virginia  
JEFFREY WOOLDRIDGE, Michigan State University

## Stata Press Editorial Manager

LISA GILMORE

## Stata Press Copy Editors

DAVID CULWELL and DEIRDRE SKAGGS

The *Stata Journal* publishes reviewed papers together with shorter notes or comments, regular columns, book reviews, and other material of interest to Stata users. Examples of the types of papers include 1) expository papers that link the use of Stata commands or programs to associated principles, such as those that will serve as tutorials for users first encountering a new field of statistics or a major new technique; 2) papers that go “beyond the Stata manual” in explaining key features or uses of Stata that are of interest to intermediate or advanced users of Stata; 3) papers that discuss new commands or Stata programs of interest either to a wide spectrum of users (e.g., in data management or graphics) or to some large segment of Stata users (e.g., in survey statistics, survival analysis, panel analysis, or limited dependent variable modeling); 4) papers analyzing the statistical properties of new or existing estimators and tests in Stata; 5) papers that could be of interest or usefulness to researchers, especially in fields that are of practical importance but are not often included in texts or other journals, such as the use of Stata in managing datasets, especially large datasets, with advice from hard-won experience; and 6) papers of interest to those who teach, including Stata with topics such as extended examples of techniques and interpretation of results, simulations of statistical concepts, and overviews of subject areas.

The *Stata Journal* is indexed and abstracted by *CompuMath Citation Index*, *Current Contents/Social and Behavioral Sciences*, *RePEc: Research Papers in Economics*, *Science Citation Index Expanded* (also known as *SciSearch*, *Scopus*, and *Social Sciences Citation Index*).

For more information on the *Stata Journal*, including information for authors, see the webpage

<http://www.stata-journal.com>

**Subscriptions** are available from StataCorp, 4905 Lakeway Drive, College Station, Texas 77845, telephone 979-696-4600 or 800-STATA-PC, fax 979-696-4601, or online at

<http://www.stata.com/bookstore/sj.html>

**Subscription rates** listed below include both a printed and an electronic copy unless otherwise mentioned.

U.S. and Canada		Elsewhere	
<b>Printed &amp; electronic</b>		<b>Printed &amp; electronic</b>	
1-year subscription	\$ 98	1-year subscription	\$138
2-year subscription	\$165	2-year subscription	\$245
3-year subscription	\$225	3-year subscription	\$345
1-year student subscription	\$ 75	1-year student subscription	\$ 99
1-year university library subscription	\$125	1-year university library subscription	\$165
2-year university library subscription	\$215	2-year university library subscription	\$295
3-year university library subscription	\$315	3-year university library subscription	\$435
1-year institutional subscription	\$245	1-year institutional subscription	\$285
2-year institutional subscription	\$445	2-year institutional subscription	\$525
3-year institutional subscription	\$645	3-year institutional subscription	\$765
<b>Electronic only</b>		<b>Electronic only</b>	
1-year subscription	\$ 75	1-year subscription	\$ 75
2-year subscription	\$125	2-year subscription	\$125
3-year subscription	\$165	3-year subscription	\$165
1-year student subscription	\$ 45	1-year student subscription	\$ 45

Back issues of the *Stata Journal* may be ordered online at

<http://www.stata.com/bookstore/sjj.html>

Individual articles three or more years old may be accessed online without charge. More recent articles may be ordered online.

<http://www.stata-journal.com/archives.html>

The *Stata Journal* is published quarterly by the Stata Press, College Station, Texas, USA.

Address changes should be sent to the *Stata Journal*, StataCorp, 4905 Lakeway Drive, College Station, TX 77845, USA, or emailed to [sj@stata.com](mailto:sj@stata.com).



Copyright © 2013 by StataCorp LP

**Copyright Statement:** The *Stata Journal* and the contents of the supporting files (programs, datasets, and help files) are copyright © by StataCorp LP. The contents of the supporting files (programs, datasets, and help files) may be copied or reproduced by any means whatsoever, in whole or in part, as long as any copy or reproduction includes attribution to both (1) the author and (2) the *Stata Journal*.

The articles appearing in the *Stata Journal* may be copied or reproduced as printed copies, in whole or in part, as long as any copy or reproduction includes attribution to both (1) the author and (2) the *Stata Journal*.

Written permission must be obtained from StataCorp if you wish to make electronic copies of the insertions. This precludes placing electronic copies of the *Stata Journal*, in whole or in part, on publicly accessible websites, file servers, or other locations where the copy may be accessed by anyone other than the subscriber.

Users of any of the software, ideas, data, or other materials published in the *Stata Journal* or the supporting files understand that such use is made without warranty of any kind, by either the *Stata Journal*, the author, or StataCorp. In particular, there is no warranty of fitness of purpose or merchantability, nor for special, incidental, or consequential damages such as loss of profits. The purpose of the *Stata Journal* is to promote free communication among Stata users.

The *Stata Journal* (ISSN 1536-867X) is a publication of Stata Press. Stata, **STATA**, Stata Press, Mata, **MATA**, and NetCourse are registered trademarks of StataCorp LP.

# A command for Laplace regression

Matteo Bottai  
Unit of Biostatistics  
Institute of Environmental Medicine  
Karolinska Institutet  
Stockholm, Sweden  
matteo.bottai@ki.se

Nicola Orsini  
Unit of Biostatistics and Unit of Nutritional Epidemiology  
Institute of Environmental Medicine  
Karolinska Institutet  
Stockholm, Sweden  
nicola.orsini@ki.se

**Abstract.** We present the new `laplace` command for estimating Laplace regression, which models quantiles of a possibly censored outcome variable given covariates. We illustrate `laplace` with an example from a clinical trial on survival in patients with metastatic renal carcinoma. We also report the results of a small simulation study.

**Keywords:** st0294, laplace, quantile regression, censored outcome, survival analysis, Kaplan–Meier

## 1 Introduction

Estimating percentiles for a time-to-event variable of interest conditionally on covariates may offer a useful complement to current approaches to survival analysis. For example, comparing survival across treatments or exposure levels in observational studies at various percentiles (for example, at the 50th or 10th percentiles) provides important insights. At the univariate level, this can be accomplished with the Kaplan–Meier estimator.

Laplace regression can be used to estimate the effect of risk factors and important predictors on survival percentiles while adjusting for other covariates. The user-written `clad` command (Jolliffe, Krushelnytsky, and Semykina 2000) estimates conditional quantiles only when censoring times are fixed and known for all observations (Powell 1986), and its applicability is limited.

In this article, we present the `laplace` command for estimating Laplace regression (Bottai and Zhang 2010). In section 3, we describe the syntax and options. In section 3, we illustrate `laplace` with data from a randomized clinical trial. In section 4, we sketch the methods and formulas. In section 5, we present the results of a small simulation study.

## 2 The laplace command

### 2.1 Syntax

```
laplace depvar [indepvars] [if] [in] [, quantiles(numlist) failure(varname)
      sigma(varlist) reps(#) seed(#) tolerance(#) maxiter(#) level(#) ]
```

by, statsby, and xi are allowed with laplace; see [U] 11.1.10 Prefix commands.

See [R] **qreg postestimation** for features available after estimation.

### 2.2 Options

**quantiles**(*numlist*) specifies the quantiles as numbers between 0 and 1; numbers larger than 1 are interpreted as percentages. The default is **quantiles**(0.5), which corresponds to the median.

**failure**(*varname*) specifies the failure event; the value 0 indicates censored observations. If **failure**() is not specified, all observations are assumed to be uncensored.

**sigma**(*varlist*) specifies the variables to be included in the scale parameter model. The default is constant only.

**reps**(#) specifies the number of bootstrap replications to be performed for estimating the variance–covariance matrix and standard errors of the regression coefficients.

**seed**(#) sets the initial value of the random-number seed used by the bootstrap. If **seed**() is specified, the bootstrapped estimates are reproducible (see [R] **set seed**).

**tolerance**(#) specifies the tolerance for the optimization algorithm. When the absolute change in the log likelihood from one iteration to the next is less than or equal to #, the **tolerance**() convergence criterion is met. The default is **tolerance**(1e-10).

**maxiter**(#) specifies the maximum number of iterations. When the number of iterations equals **maxiter**(), the optimizer stops, displays an x, and presents the current results. The default is **maxiter**(2000).

**level**(#) specifies the confidence level, as a percentage, for confidence intervals. The default is **level**(95) or as set by **set level**.

## 2.3 Saved results

`laplace` saves the following in `e()`:

### Scalars

<code>e(N)</code>	number of observations	<code>e(n.q)</code>	number of estimated quantiles
<code>e(N_fail)</code>	number of failures	<code>e(reps)</code>	number of bootstrap replications

### Macros

<code>e(cmd)</code>	<code>laplace</code>	<code>e(qlist)</code>	requested quantiles
<code>e(cmdline)</code>	command as typed	<code>e(vcetype)</code>	title used to label Std. Err.
<code>e(depvar)</code>	name of dependent variable	<code>e(properties)</code>	<code>b V</code>
<code>e(eqnames)</code>	names of equations	<code>e(predict)</code>	program used to implement <code>predict</code>

### Matrices

<code>e(b)</code>	coefficient vector	<code>e(V)</code>	variance–covariance matrix of the estimators
-------------------	--------------------	-------------------	--

### Functions

<code>e(sample)</code>	marks estimation sample
------------------------	-------------------------

## 3 Example: Survival in metastatic renal carcinoma

We illustrate the use of `laplace` with data from a clinical trial on 347 patients with metastatic renal carcinoma. The patients were randomly assigned to either interferon- $\alpha$  (IFN) or oral medroxyprogesterone (MPA) (Medical Research Council Renal Cancer Collaborators 1999). A total of 322 patients died during follow-up. The outcome of primary research interest is overall survival.

```
. use kidney_ca_1
(kidney cancer data)
. quietly stset months, failure(cens)
```

The numeric variable `months` represents the time to event or censoring, and the binary variable `cens` indicates the failure status (0 = censored, 1 = death).

### 3.1 Median survival

We estimate a Laplace regression model where the response variable is time to death or censoring (`months`) and the binary indicator for treatment (`trt`) is the only covariate. We specify the event status with the option `failure()`. The default percentile is the median (`q50`).

```
. laplace months trt, failure(cens)
```

Laplace regression

	No. of subjects =	347
	No. of failures =	322

---

	months	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]
q50	trt	3.130258	1.195938	2.62	0.009	.7862628 5.474254
	_cons	6.80548	.7188408	9.47	0.000	5.396578 8.214382

The estimated median survival in the MPA group is 6.8 months (95% confidence interval: [5.4, 8.2]). The difference (`trt`) in median survival between the treatment groups is 3.1 months (95% confidence interval: [0.8, 5.5]). Median survival among patients on IFN can be obtained with the postestimation command `lincom`.

```
. lincom _cons + trt
```

(1) [q50]trt + [q50]\_cons = 0

---

	months	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
(1)		9.935738	.9557906	10.40	0.000	8.062423 11.80905

Percentiles of survival time by treatment group can also be obtained from the Kaplan–Meier estimate of the survivor function by using the command `stci`.

```
. stci, by(trt)
```

failure \_d: cens  
analysis time \_t: months

trt	no. of subjects	50%	Std. Err.	[95% Conf. Interval]
MPA	175	6.80548	.8902896	4.86575 8.15342
IFN	172	9.830137	.8982793	7.7589 11.7041
total	347	7.956164	.5699226	6.90411 9.1726

The estimated median in the IFN group (9.8 months) differs slightly from the `laplace` estimate (9.9 months) shown above. The Kaplan–Meier curve in the IFN group is flat at the 50th percentile between 9.83 and 9.96 months of follow-up. The command `stci` shows the lower limit of this interval while `laplace` shows a middle value.

### 3.2 Multiple survival percentiles

When it is relevant to estimate multiple percentiles of the distribution of survival time, these can be specified with the option `quantiles()`.

```
. laplace months trt, failure(cens) quantiles(25 50 75) rep(100) seed(123)
Laplace regression                               No. of subjects =      347
                                                No. of failures =      322
```

months	Coef.	Bootstrap Std. Err.	z	P> z	[95% Conf. Interval]		
q25	trt	1.509151	.8289345	1.82	0.069	-.1155312	3.133832
	_cons	2.49863	.399623	6.25	0.000	1.715384	3.281877
q50	trt	3.130258	1.209658	2.59	0.010	.7593719	5.501145
	_cons	6.80548	.9100921	7.48	0.000	5.021732	8.589227
q75	trt	3.663238	3.482536	1.05	0.293	-3.162407	10.48888
	_cons	15.87945	1.714295	9.26	0.000	12.5195	19.23941

The treatment effect is larger at higher percentiles of survival time. The difference between the two treatment groups at the 25th, 50th, and 75th percentiles is 1.5, 3.1, and 3.7 months, respectively. When bootstrap is requested, one can test for differences in treatment effects across survival percentiles with the postestimation command `test`.

```
. test [q25]trt = [q50]trt
( 1) [q25]trt - [q50]trt = 0
      chi2( 1) =      2.59
      Prob > chi2 =      0.1076
```

We fail to reject the hypothesis that the treatment effects at the 25th and 50th survival percentiles are equal ( $p$ -value > 0.05).

Figure 1 shows the predicted percentiles from the 1st to the 99th in each treatment group. The difference of 3 months in median survival between groups is represented by the horizontal distance between the points A and B. Approximately 30% and 40% of the patients on MPA and IFN, respectively, are estimated to live longer than 12 months. The absolute difference of about 10% in the probability of surviving 12 months is represented by the vertical distance between the points C and D.



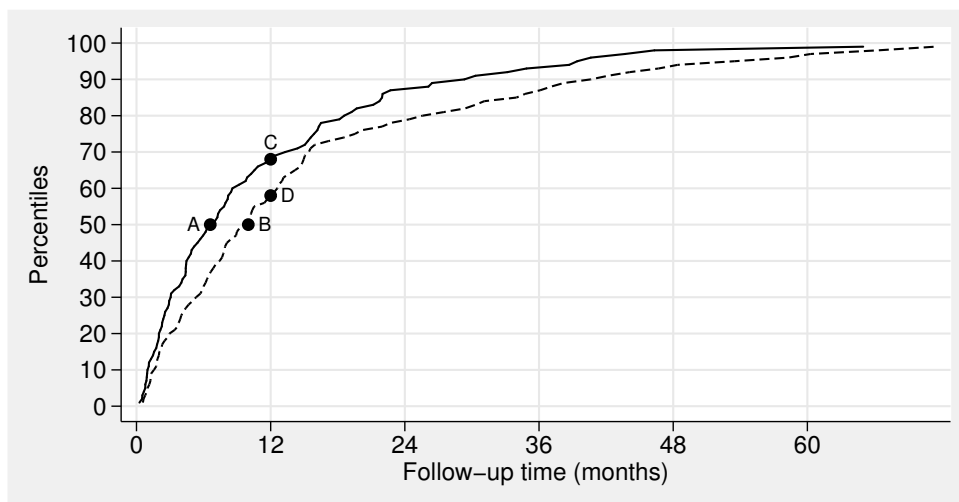


Figure 1. Survival percentiles in the MPA (solid line) and IFN (dashed line) groups estimated with Laplace regression. The horizontal distance between the points A and B (3.1 months) indicates the difference in median survival between groups. The vertical distance between C and D (about 10%) indicates the difference in the proportion of patients estimated to survive 12 months.

### 3.3 Interactions between covariates

Royston, Sauerbrei, and Ritchie (2004) analyzed the same data and described how a continuous prognostic factor, white cell count (`wcc`), affects the treatment effect as measured by a relative hazard. We now perform a similar analysis by using Laplace regression for the median survival. We include as covariates the treatment indicator (`trt`), three equally sized classes of white cell counts (`cwcc`) by means of two indicator variables, and their interactions.

```
. xi: laplace months i.trt*i.cwcc, failure(cens)
i.trt          _Itrt_0-1          (naturally coded; _Itrt_0 omitted)
i.cwcc         _Icwcc_0-2        (naturally coded; _Icwcc_0 omitted)
i.trt*i.cwcc   _ItrtXcwc_#_#    (coded as above)
Laplace regression                               No. of subjects =      347
                                                No. of failures  =      322
```

months	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
q50						
_Itrt_1	8.01462	2.270786	3.53	0.000	3.563962	12.46528
_Icwcc_1	2.262442	2.068403	1.09	0.274	-1.791554	6.316438
_Icwcc_2	-2.496523	1.645959	-1.52	0.129	-5.722544	.7294982
_ItrtXcwc_1_1	-5.737988	3.241483	-1.77	0.077	-12.09118	.6152021
_ItrtXcwc_1_2	-7.751629	2.645534	-2.93	0.003	-12.93678	-2.566478
_cons	6.90203	1.658547	4.16	0.000	3.651337	10.15272

The predicted median survival can be obtained with standard postestimation commands such as `predict` or `adjust`.

```
. adjust, by(trt cwcc) format(%2.0f) noheader
```

treatment	White Cell Counts		
	Low	Medium	High
MPA	7	9	4
IFN	15	11	5

Key: Linear Prediction

The between-treatment-group difference in median survival varies from 8 months in the low white cell count category to 1 month in the high white cell count category. We test for interaction between treatment and white cell counts with the postestimation command `testparm`.

```
. testparm _ItrtX*
( 1) [q50]_ItrtXcwc_1_1 = 0
( 2) [q50]_ItrtXcwc_1_2 = 0
      chi2( 2) =      8.59
      Prob > chi2 =    0.0137
```

We reject the null hypothesis of equal treatment effect across categories of white cell counts ( $p = 0.0137$ ). The treatment effect seems to be largest in patients with low white cell counts.

### 3.4 Laplace regression with uncensored data

Suppose all the values for the variable `months` were uncensored times at death. The `laplace` command can be used with uncensored observation by omitting the `failure()` option. In this case, `laplace` is simply an alternative to the standard quantile regression commands `qreg` and `sqreg`.

```
. qui laplace months trt
. adjust, by(trt) format(%3.2f) noheader
```

treatment	xb
MPA	6.77
IFN	9.89

Key: xb = Linear Prediction

```
. qui qreg months trt
. adjust, by(trt) format(%3.2f) noheader
```

treatment	xb
MPA	6.77
IFN	9.96

Key: xb = Linear Prediction

The number of observations in the MPA group is odd (175 patients), and the sample median survival is 6.77 months. The number of observations in the IFN group is even (172 patients), and the median is not uniquely defined. The two nearest values are 9.83 and 9.96 months. The command `qreg` picks the larger of the two, while `laplace` picks a value in between.

## 4 Methods and formulas

In this section, we follow the description provided by Bottai and Zhang (2010). Suppose we have a sample of size  $n$ . Let  $t_i$ ,  $i = 1, \dots, n$ , be a continuous outcome variable,  $c_i$  be a continuous censoring variable, and  $x_i = \{x_{1,i}, \dots, x_{r,i}\}'$  and  $z_i = \{z_{1,i}, \dots, z_{s,i}\}'$  be two vectors of covariates. The sets of covariates contained in  $x_i$  and  $z_i$  may partially or entirely overlap. We assume that  $c_i$  is independent of  $t_i$  conditionally on the covariates. Suppose we observe  $(y_i, d_i, x'_i, z'_i)$ , with  $y_i = \min(t_i, c_i)$  and  $d_i = I(t_i \leq c_i)$ , where  $I(A)$  denotes the indicator function of the event  $A$ . We assume that

$$t_i = x'_i \beta_p + \exp(z'_i \sigma_p) \varepsilon_i \quad (1)$$

where  $\beta_p = \{\beta_{p,1}, \dots, \beta_{p,r}\}'$  and  $\sigma_p = \{\sigma_{p,1}, \dots, \sigma_{p,s}\}'$  indicate the unknown parameter vectors, and  $\varepsilon_i$  are independent and identically distributed error terms that follow a standard Laplace distribution,  $f(\varepsilon_i) = p(1-p) \exp\{[I(\varepsilon_i \leq 0) - p]\varepsilon_i\}$ . For any given  $p \in (0, 1)$ , the  $p$ -quantile of the conditional distribution of  $t_i$  given  $x_i$  and  $z_i$  is  $x'_i \beta_p$  because  $P(t_i \leq x'_i \beta_p | x_i, z_i) = p$ .

The command `laplace` estimates the  $(r+s)$ -dimensional parameter vector  $\{\beta'_p, \sigma'_p\}$  by maximizing the Laplace likelihood function described by Bottai and Zhang (2010). It uses an iterative maximization algorithm based on the gradient of the log likelihood that generates a finite sequence of parameter values along which the likelihood increases. Briefly, from a current parameter value, the algorithm searches the positive semiline in the direction of the gradient for a new parameter value where the likelihood is larger.

The algorithm stops when the change in the likelihood is less than the specified tolerance. Convergence is guaranteed by the continuity and concavity of the likelihood.

The asymptotic variance of the estimator  $\widehat{\beta}_p$  for the parameter  $\beta_p$  is derived by considering the estimating condition reported by Bottai and Zhang (2010, eq. 4),  $S(\widehat{\beta}_p) = 0$ , where

$$S(\widehat{\beta}_p) = \frac{1}{\exp(z'_i \widehat{\sigma}_p)} \sum_{i=1}^n x_i \left\{ p - I(y_i \leq x'_i \beta_p) - I(y_i \leq x'_i \beta_p) (1 - d_i) \frac{p - 1}{1 - \widehat{F}(y_i | x_i)} \right\}$$

with  $\widehat{F}(y_i | x_i) = p \exp\{(1 - p)(y_i - x'_i \widehat{\beta}_p) / \exp(z'_i \widehat{\sigma}_p)\}$ . Following the standard asymptotic theory for method of moments estimators,  $\widehat{\beta}_p$  approximately follows a normal distribution with mean  $\beta_p^*$  and variance  $\widehat{V}$ , where  $\beta_p^*$  indicates the expected value of  $\beta_p$ ,  $\widehat{V} = H(\widehat{\beta}_p)^{-1} S(\widehat{\beta}_p)' S(\widehat{\beta}_p) H(\widehat{\beta}_p)^{-1}$ , and  $H(\widehat{\beta}_p) = \partial S(\beta_p) / \partial \beta'_p |_{\beta_p = \widehat{\beta}_p}$ . The derivative in  $H(\widehat{\beta}_p)$  is evaluated numerically. Alternatively, the standard errors can be obtained with bootstrap by specifying the `reps()` option.

## 5 Simulation

In this section, we present the setup and results of a small simulation study to assess the finite sample performance of the Laplace regression estimator under different data-generating mechanisms. We contrast the performance of Laplace with that of the Kaplan–Meier estimator, a standard, nonparametric, uniformly consistent, and asymptotically normal estimator of the survival function. To generate the survival estimates, we used the `sts` command.

We generated 500 samples from (1) in each of the six different simulation scenarios that arose from the combination of two sample sizes and three data-generating mechanisms. In each scenario, we estimated five percentiles ( $p = 0.10, 0.30, 0.50, 0.70, 0.90$ ) with Laplace regression and the Kaplan–Meier estimator. The two sample sizes were  $n = 100$  and  $n = 1,000$ . The three different data-generating mechanisms were obtained by changing the values of  $z_i$ ,  $\sigma_p$ , and the censoring variable  $c_i$ . In all simulation scenarios,  $x_i = (1, x_{1,i})'$ , with  $x_{1,i} \sim \text{Bernoulli}(0.5)$ ,  $\beta_p = (5, 3)'$ , and  $\varepsilon_i$  was a standard normal centered at the quantile being estimated.

In scenario number 1,  $z_i = 1$ ,  $\sigma_p = 1$ , and the censoring variable was set equal to a constant  $c_i = 1,000$  for all individuals. In this scenario, no observations were censored, and Laplace regression was equivalent to ordinary quantile regression. In scenario number 2,  $z_i = 1$ ,  $\sigma_p = 1$ , and the censoring variable was generated from the same distribution as the outcome variable  $t_i$ . This ensured an expected censoring rate of 50% in both covariate patterns ( $x_{1,i} = 0, 1$ ). In scenario number 3,  $z_i = (1, x_{1,i})'$  and  $\sigma_p = (0.5, 0.5)'$ . The censoring variable  $c_i$  was generated from the same distribution as the outcome variable  $t_i$ . In this scenario, the standard deviation of  $t_i$  was equal to 0.5 when  $x_{1,i} = 0$  and equal to 1 when  $x_{1,i} = 1$ .

The following table shows the observed relative mean squared error multiplied by 1,000 for the predicted quantile in the group  $x_{1,i} = 1$  in each combination of sample size (**obs**), data-generating scenario (**scenario**), and percentile (**percentile**) for Laplace (top entry) and Kaplan–Meier (bottom entry).

```
. table percentile scenario obs, contents(mean msel mean msekm) format(%4.3f)
> stubwidth(12)
```

percentile	obs and scenario					
	100			1000		
	1	2	3	1	2	3
10	1.187	1.395	1.268	0.129	0.136	0.126
	1.233	1.496	1.320	0.132	0.140	0.132
30	0.597	0.685	0.680	0.064	0.073	0.067
	0.606	0.792	0.831	0.064	0.078	0.075
50	0.496	0.570	0.653	0.053	0.065	0.073
	0.505	0.860	0.941	0.053	0.074	0.075
70	0.513	0.639	0.711	0.050	0.131	0.144
	0.518	1.329	1.050	0.050	0.113	0.094
90	0.728	1.661	1.930	0.063	0.876	0.955
	0.731	1.835	1.701	0.063	0.478	0.450

The relative mean squared error was smaller for Laplace than for Kaplan–Meier at lower quantiles and with the smaller sample size.

Figure 2 shows the relative mean squared error of Laplace ( $x$  axis) and Kaplan–Meier ( $y$  axis) estimators of the quantile in group  $x_{1,i} = 1$  over all simulation scenarios.

The Laplace estimator had fewer extreme values than Kaplan–Meier. The overall concordance correlation coefficient (command `concord`) was 72.2%. After the 10% largest differences were excluded, the coefficient was 99.1%.

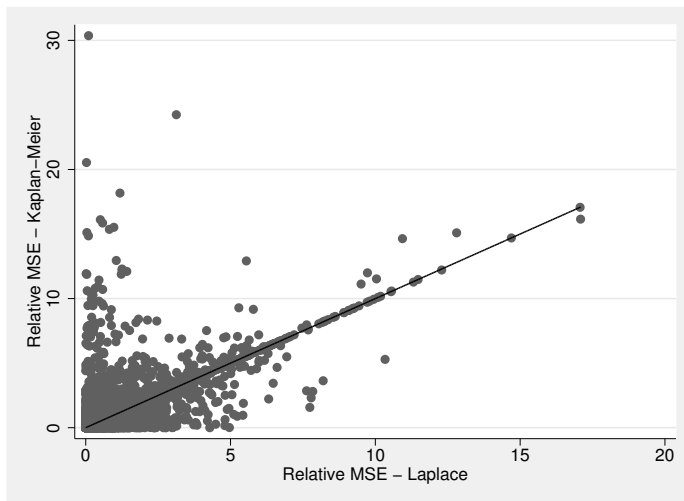


Figure 2. Relative mean squared error of Laplace ( $x$  axis) and Kaplan–Meier ( $y$  axis) estimators of the percentiles in group  $x_{1,i} = 1$  over all simulation scenarios. The solid 45-degree line indicates the equal relative mean squared error of the two estimators.

The following two tables show the performance of the estimator of the asymptotic standard error for the regression coefficients  $\hat{\beta}_{p,0}$  (first table) and  $\hat{\beta}_{p,1}$  (second table). In each cell of each table, the top entry is the average estimated asymptotic standard error, and the bottom entry is the corresponding observed standard deviation across the simulated samples.

```
. table percentile scenario obs, contents(mean s0 mean ms0) format(%4.3f)
> stubwidth(12)
```

percentile	obs and scenario					
	100			1000		
	1	2	3	1	2	3
10	0.237	0.228	0.131	0.076	0.077	0.039
	0.235	0.251	0.123	0.073	0.082	0.039
30	0.185	0.200	0.098	0.059	0.062	0.031
	0.182	0.193	0.097	0.058	0.067	0.032
50	0.176	0.194	0.097	0.056	0.060	0.030
	0.169	0.185	0.093	0.053	0.064	0.032
70	0.188	0.198	0.098	0.059	0.064	0.032
	0.185	0.207	0.103	0.057	0.071	0.035
90	0.225	0.227	0.114	0.077	0.076	0.038
	0.231	0.255	0.141	0.072	0.087	0.046

```
. table percentile scenario obs, contents(mean s1 mean ms1) format(%4.3f)
> stubwidth(12)
```

percentile	obs and scenario					
	100			1000		
	1	2	3	1	2	3
10	0.349	0.353	0.276	0.109	0.110	0.087
	0.330	0.351	0.263	0.104	0.113	0.088
30	0.277	0.292	0.232	0.084	0.089	0.070
	0.265	0.269	0.216	0.079	0.092	0.066
50	0.255	0.279	0.219	0.080	0.086	0.068
	0.250	0.257	0.226	0.077	0.086	0.073
70	0.272	0.293	0.227	0.084	0.090	0.070
	0.265	0.277	0.236	0.081	0.094	0.076
90	0.337	0.339	0.246	0.109	0.108	0.085
	0.325	0.320	0.284	0.104	0.109	0.098

The estimated standard errors were similar to the observed standard deviation across all cells for both regression coefficients.

## 6 Acknowledgment

Nicola Orsini was partly supported by a Young Scholar Award from the Karolinska Institutet's Strategic Program in Epidemiology.

## 7 References

- Bottai, M., and J. Zhang. 2010. Laplace regression with censored data. *Biometrical Journal* 52: 487–503.
- Jolliffe, D., B. Krushelnytsky, and A. Semykina. 2000. sg153: Censored least absolute deviations estimator: CLAD. *Stata Technical Bulletin* 58: 13–16. Reprinted in *Stata Technical Bulletin Reprints*, vol. 10, pp. 240–244. College Station, TX: Stata Press.
- Medical Research Council Renal Cancer Collaborators. 1999. Interferon- $\alpha$  and survival in metastatic renal carcinoma: Early results of a randomised controlled trial. *Lancet* 353: 14–17.
- Powell, J. L. 1986. Censored regression quantiles. *Journal of Econometrics* 32: 143–155.
- Royston, P., W. Sauerbrei, and A. Ritchie. 2004. Is treatment with interferon-alpha effective in all patients with metastatic renal carcinoma? A new approach to the investigation of interactions. *British Journal of Cancer* 90: 794–799.

**About the author**

Matteo Bottai is a professor of biostatistics in the Unit of Biostatistics at the Institute of Environmental Medicine at Karolinska Institutet in Stockholm, Sweden.

Nicola Orsini is an associate professor of medical statistics and an assistant professor of epidemiology in the Unit of Biostatistics and the Unit of Nutritional Epidemiology at the Institute of Environmental Medicine at Karolinska Institutet in Stockholm, Sweden.