A NOTE ON LAPLACE REGRESSION WITH CENSORED DATA

ROGER KOENKER

Abstract. The Laplace likelihood method for estimating linear conditional quantile functions with right censored data proposed by Bottai and Zhang (2010) is demonstrated to be highly non-robust.

1. Introduction

Bottai and Zhang (2010) have recently proposed a method of estimating linear conditional quantile functions for right censored data based on an asymmetric Laplace likelihood formulation. They assert that their procedure is both more accurate and faster than prior proposals by Portnoy (2003) and Peng and Huang (2008). The objective of this note is to demonstrate that neither of these claims are correct.

2. Inconsistency of the Bottai and Zhang Estimator

The flaw in the proposed estimator is immediately apparent from consideration of the simplest median regression setting, where, in effect, it is assumed that the event times, T_i, conditional on covariate vector, x_i, arise from the double exponential (Laplace) density,

$$f(t|x) = \frac{1}{4\sigma} \exp(-|t - x^\top \beta|/2\sigma),$$

with corresponding distribution function,

$$F(t_i|x_i) = \frac{1}{2}(1 + \mathrm{sgn}(t - x^\top \beta)[1 - \exp(-|t - x^\top \beta|/2\sigma)],$$

For right censored data with $Y_i = T_i \wedge C_i$ and censoring indicator, $\delta_i = I(T_i < C_i)$, arising from such a parametric model we have the likelihood,

$$\ell_n(\beta, \sigma) = \sum_{i=1}^{n} \delta_i \log f(y_i|x_i) + (1 - \delta_i) \log F(y_i|x_i)$$

and we can optimize with respect to (β, σ). Optimizing leads to a weighted quantile regression procedure with weights depending upon the estimated scale parameter $\hat{\sigma}$. This weighting is asymptotically the same as that introduced by Portnoy (2003) *provided the Laplacian distributional assumption holds.*

In the uncensored case, so $\delta_i \equiv 1$, it is easily seen that the estimation of β is independent of σ, reducing to conventional median regression, and $\hat{\sigma}$ becomes simply the mean absolute error of the median regression fit. Thus, the Laplacian assumption is harmless. However, in the censored case, the estimation of β and σ become inextricably linked, so when the distributional assumption fails, the weights produced by the Laplace likelihood are no longer correct and the procedure fails to yield a consistent estimator. Note that this is the case even when the errors *are* Laplace

Version: March 23, 2011. This note was partially supported by NSF grant SES-08-50060. All code required to reproduce the results reported here is available in R from the author by request, or from the url `http://www.econ.uiuc.edu/~roger/research/crq/crq.html`.

if one is estimating a quantile other than the median with the asymmetric version of the Laplace likelihood, and in non-iid error conditions the situation is even more problematic since a single scale parameter based on the mean absolute error no longer provides a reliable way to estimate the weighting process. This failing is apparent in the simulation experiments reported in the next section.

3. Some Simulation Evidence

Bottai and Zhang write,

> On the whole, we found that the performance of Portnoy's and Peng and Huang's methods was poorer than that of Wang and Wang (2009). . . In all the simulation scenarios considered, Laplace regression was consistently better than Wang and Wang's method in terms of bias and [coverage probability].

To reevaluate the finite sample performance of the procedure we have tried to reproduce some of the simulation results of Bottai and Zhang (2010). This is somewhat difficult, since the simulation design is incompletely described, however we believe that the setting described below is quite similar to theirs. We focus on the comparison with Portnoy's estimator. We consider nine distinct models. There are three specifications of the covariate effects

$$[M1] \qquad Q_{\log T_i|x_i}(\tau|x_i) = -2 + 6x_{1i} + F_u^{-1}(\tau),$$

$$[M2] \qquad Q_{\log T_i|x_i}(\tau|x_i) = 1 + 2x_{2i} + 3x_{3i} + F_u^{-1}(\tau),$$

$$[M3] \qquad Q_{\log T_i|x_i}(\tau|x_i) = -2 + 6x_{1i} + x_{1i}F_u^{-1}(\tau),$$

and there are three choices of the quantile function, F_u^{-1}: Standard Gaussian, Gumbel with scale $\pi/\sqrt{3}$, and t_3 with scale $\sqrt{3}$, designated F1, F2, and F3 respectively in Table 1. The covariates are generated iid-ly from standard uniform, standard Gaussian and Bernoulli (p = 0.5) distributions respectively. The precise nature of the censoring used in the Bottai-Zhang simulations is somewhat unclear, but they comment that the censoring variable, C_i is uniformly distributed on $[0, a]$ with a chosen for each model "to make the censoring rate approximately equal to 20 and 40%" [sic]. In the simulations reported here, this same uniform censoring is employed with $a = 35$ for models M1 and M3, and $a = 150$ for model M2. This yields between 1/4 to 1/3 censoring in each of the nine models investigated. In all cases we estimate four coefficients for the three quartiles and compare performance of the Bottai-Zhang estimator with that of Portnoy's `crq` estimator as implemented in the `quantreg` R package of Koenker (2010). Prior comparisons reported in Koenker (2008) indicate that the Peng and Huang (2008) estimator performs very similarly to that of Portnoy (2003). For the implementation of the Bottai-Zhang estimator we have used the code provided in the electronic supplement to their published paper.

Before considering all nine models we conducted a preliminary exercise with the simplest of the nine, model M1 with Gaussian error. In Figures 1 and 2 report bias and (scaled) root mean squared error for the Bottai-Zhang estimator and the Portnoy estimator, for four sample sizes: $n \in \{100, 500, 1000, 10,000\}$. There are four parameters for each of the fitted models, and we evaluate performance for the three quartile estimates. Figure 1 reports the absolute value of the bias scaled by 100, for the two estimators. Figure 2 reports root mean squared errors scaled by the square root of sample size n. In all cases 1000 replications were done.

As can be seen in Figure 1, the bias of the Portnoy estimate is quite small and stable over the range of sample sizes considered, however the Bottai-Zhang estimator has quite substantial bias, somewhat attentuated by larger sample sizes, but remaining quite substantial even at the $n = 10,000$ sample size. The impact of this bias effect is also apparent in Figure 2 where root mean squared errors particularly for the non-zero coefficients β_1 and β_2 for the first and third quartiles are increasing in sample sample size. Note that because the root mean squared errors are scaled by the square root of sample size they should be quite stable with respect to n, and we see that this is true for the Portnoy crq estimator, but it is not the case for the Bottai-Zhang Laplace likelihood estimator.

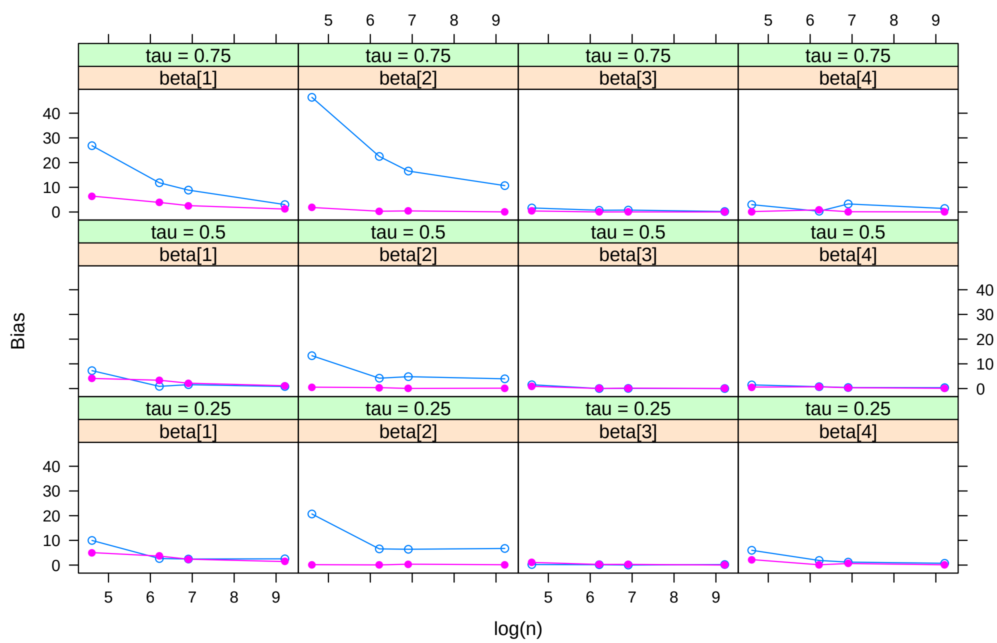

FIGURE 1. Absolute Values of Bias (scaled by 100) for the Portnoy (magenta filled circles) and Bottai-Zhang (blue open circles) censored quantile regression estimators.

In Table 2 we report comparative performance of the Bottai-Zhang (Laplace likelihood) estimator relative to the Portnoy `crq` estimator for all nine models. In all cases $n = 200$ and 1000 replications were done. Entries in the table represent ratio of root mean square errors for the two methods, with entries greater than one indicating superior performance for the Portnoy estimator. As is evident from the table the Portnoy method dominates the Bottai-Zhang estimator in all of the cases investigated.

4. COMPUTATIONAL SPEED

Finally, we come to the Bottai-Zhang claim that their estimator, as implemented in R with the Nelder-Mead option of the R function `optim`, is *faster* than the Portnoy estimator. This is truly puzzling, since our R timings of a typical instance of the simulation estimation of the three quartiles is roughly 40 times faster for the Portnoy method than the Bottai-Zhang method when the sample size is 500, 20 times faster

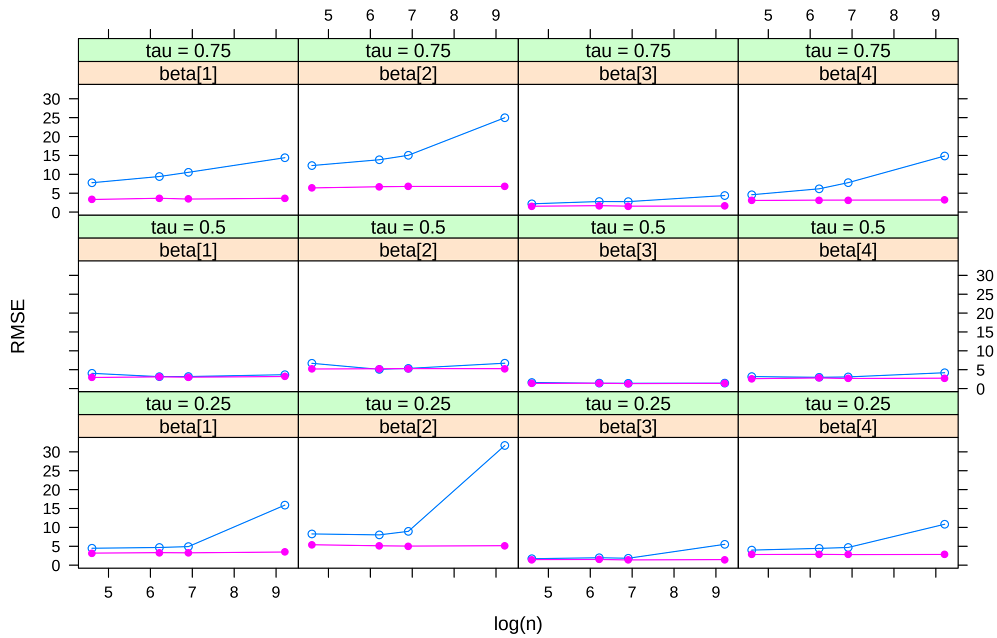

FIGURE 2. Root mean squared error (scaled by $\sqrt{n}$) for the Portnoy (magenta filled circles) and Bottai-Zhang (blue open circles) censored quantile regression estimators.

when $n = 1000$, and even at $n = 10,000$ is slightly faster despite the fact that it has estimated the model at more than 100 τ's rather than only three. These timings were made without any bootstrapping.

5. Conclusions

We cannot offer any explanation for what went wrong – suffice it to say that both of the central claims of Bottai and Zhang (2010) fail to stand up to closer examination. As for the "heuristic interpretation" offered for the Laplace likelihood approach, it is worth mentioning again that in the uncensored case their method reduces to conventional quantile regression, so the scale parameter estimation is harmless, but it should also be emphasized that the proposed Nelder-Mead algorithm is a poor alternative to standard linear programming methods. In the censored case, the scale parameter of the Laplace likelihood is not so innocuous, it is unable to adapt to the myriad possibilities that give rise to the linear conditional quantile model, and consequently its performance is inevitably poor.

References

BOTTAI, M., AND J. ZHANG (2010): "Laplace regression with censored data," *Biometrical Journal*, 52, 487–503.

KOENKER, R. (2008): "Censored Quantile Regression Redux," *Journal of Statistical Software*, 27, 1–14.

KOENKER, R. (2010): `quantreg`*: An R Package for Quantile Regression*Version 4.57, `http://CRAN.R-project.org/package=quantreg`.

PENG, L., AND Y. HUANG (2008): "Survival Analysis with Quantile Regression Models," *Journal of American Statistical Association*, 103, 637–649.

Coefficient	M1			M2			M3		
	F1	F2	F3	F1	F2	F3	F1	F2	F3
$\tau = \mathbf{0.25}$									
$\hat{\beta}_0$	1.19	1.15	3.41	1.41	1.34	6.29	1.21	1.19	6.17
$\hat{\beta}_1$	1.30	1.17	1.94	1.37	1.37	3.75	1.41	1.24	3.26
$\hat{\beta}_2$	1.12	1.02	2.78	1.25	1.09	4.87	1.26	1.07	4.93
$\hat{\beta}_3$	1.28	1.09	3.46	1.50	1.19	5.70	1.53	1.16	6.48
$\tau = \mathbf{0.5}$									
$\hat{\beta}_0$	1.12	1.31	4.40	1.25	2.06	7.34	1.37	1.79	5.90
$\hat{\beta}_1$	1.07	1.31	2.08	1.17	2.01	2.51	1.27	1.90	3.01
$\hat{\beta}_2$	1.05	1.10	3.64	1.15	1.44	3.13	1.25	1.30	4.40
$\hat{\beta}_3$	1.08	1.25	2.72	1.22	1.60	3.25	1.35	1.64	4.06
$\tau = \mathbf{0.75}$									
$\hat{\beta}_0$	2.54	3.82	16.51	2.60	3.35	19.08	3.75	6.10	21.78
$\hat{\beta}_1$	2.00	3.43	4.82	1.83	3.38	5.84	2.71	4.61	6.42
$\hat{\beta}_2$	2.26	2.51	9.14	1.42	1.80	14.33	3.73	2.91	15.98
$\hat{\beta}_3$	1.85	3.70	9.92	1.65	2.18	9.06	2.74	5.57	12.71

Table 1. Ratio of Root Mean Squared Errors: Table entries report the ratio of the RMSE for the Bottai-Zhang (Laplace likelihood) estimator to the corresponding Portnoy crq estimator, so entries larger than one indicate superior performance for the Portnoy method. All entries are based on 1000 replications.

Portnoy, S. (2003): "Censored Quantile Regression," *Journal of American Statistical Association*, 98, 1001–1012.

Wang, H. J., and L. Wang (2009): "Locally Weighted Censored Quantile Regression," *Journal of American Statistical Association*, 104, 1117–1128.