

Answers to the referee reports of the submission entitled “OneStep - Le Cam’s one-step estimation procedure”

We would like to thank the referees for their valuable comments that improve the paper. We consider in the revised version of our manuscript all the points raised by the referees. We give below an answer to each point:

Reviewer 1:

1. *The paper focuses [...] this is not a huge issue but the beauty of the one-step approach is that it increases the efficiency of simple but possibly highly inefficient estimator. For many distributions, simple estimators can be computed using the order statistics. Some examples are :*

- (a) *The paper uses initial estimators based on the empirical characteristic function for the location and scale parameters of the Cauchy distribution; this does seem unnecessarily complicated. For example, simple estimators of the location and scale parameters are, respectively, the sample median and one half the sample interquartile range.*

The new sequence of initial guess estimators proposed by the referee is simple and fast to be computed. Consecutively, we add it as the default method in the package for the Cauchy distribution. The corresponding paragraph in the paper and the simulations are also corrected accordingly. Our former procedure (Koutrouvelis, 1982) is now an example in the documentation in order to use the argument `init` (see below).

- (b) *For the Weibull distribution, simple initial estimators can be obtained from a Weibull plot, whereby $\ln(X(i))$ is plotted against $\ln(-\ln(1 - i/(n + 1)))$ and the two parameters can be estimated via the estimated slope and intercept.*

The new sequence of initial guess estimators proposed by the referee is simple and fast to be computed. Consecutively, we add it as the default method in the package for the Weibull distribution. The corresponding paragraph in the paper and the simulations are also corrected accordingly.

2. *The observation that an initial estimator need only be $O(n^{\delta/2})$ -consistent for $1/2 < \delta \leq 1$ is interesting. I believe that this goes back to work by Peter Bickel back in the 1970s but I can’t find an exact reference (if indeed one exists). You mention that this results allows us to use estimators based on subsamples; in certain situations, it also allows us to use mode-type estimators of location and scale, which typically are $O(n^{1/3})$ -consistent.*

Unfortunately, we did not find the reference. Nevertheless, we change in the revised version of the paper the sentence page 7 “More recently, it has been shown [...] that for a $n^{\delta/2}$ -consistent [...]” by “Indeed, it can be shown been (see [...]) that for a $n^{\delta/2}$ -consistent [...]”. The fact that mode-type estimators can be used in the “improved” Le Cam one-step procedure is very interesting. We mention it as an example in the paper for the use of the `init` parameter (see below) and add it in the documentation for the χ^2 distribution.

3. *Package: I tried it out and everything seems to work well, at least using the defaults. I did find the documentation to be confusing - it seems you need to be familiar with `fitdistplus` in addition to `OneStep` to fully understand the documentation. For example, it wasn’t clear if you could specify your own initial estimates in the function `onestep`. More examples in the documentation would be useful.*

The `init` parameter has been included in the package and in the revised version of the paper. It allows one to use its own sequence of initial guess estimators. We add several examples in the

documentation: for the gamma distribution (Ye and Chen, 2017), for the Cauchy distribution (Koutrouvelis, 1982) and for the Chi2 distribution (Grenander, 1965).

Reviewer 2:

The package is of importance, written well, and generally motivated well.

1. *A flaw is that "computational cost" is described in the various tables with no units attached. Please say what these numbers mean - are they milliseconds? Seconds? How was this timing performed? Give the relevant R function used for benchmarking, and discuss the specifics of the computer on which timing was measured (the type of CPU, at least). The author's own function `benchonestep` can at least be mentioned.*

We add to the revised version that the timing performance (given in seconds) are done on the Monte-Carlo simulations with the `proc.time` command ("`elapsed`" time) on a laptop with an Intel Core i7 2.7 GHz processor with 8GB RAM. We mention in the text the `benchonestep` and `benchonestep.replicate` functions.

2. *The authors indicate in the abstract that a major (the major?) advantage of their method is that it is "computed faster than the maximum likelihood estimator for large datasets." They also say on page 1 that the method is "appropriate for very large datasets." Can you illustrate this with an example on a very large dataset? Simulation of such a case would better illustrate the method's advantages.*

As mentioned by the referee, we add to the revised version a paragraph in the first example (Gamma distribution) that the one-step procedure can be computed faster than the a simulation maximum likelihood estimator for large datasets. We add also a simulation with large datasets of size $n = 10^r$ for $r = 3, \dots, 9$. It shows on a simple example the performance of the one-step estimation in terms of computation time.

3. *In all the tables, remove the "1" that labels the first column. I recommend replacing the "1" with the correct unit of time, and also adding to each table a statistic quantifying the error that the figures are used to illustrate. This will allow an easy comparison of performance, though the figures are also nice, and should be kept in.*

The "1" has been replaced by "Computation time (s)" in the Tables.

For the error, we propose in the revised version to use the Cramer-von Mises (CvM) statistics T . The asymptotic distribution of T is tabulated for instance in the `gofest` package. We add in the Tables the computed CvM statistics for both coordinates ϑ_1 and ϑ_2 .

4. *Explain somewhere (in the text once, or in each caption) what the difference between the blue and red lines is for the method of moments in Figures 1-3.*

Red and blue lines have been described (theoretical or empirical asymptotic distributions) in all captions.

5. *For the final "generic" example, can you also quantify performance in terms of error?*

We add in the generic example a Table with the timing performance and the CvM statistics in order to quantify also the variance efficiency (error).

6. *Minor edits: All the minor edits have been corrected.*

Sincerely yours,

Alexandre Brouste