DISCUSSION OF: ESTIMATING THE HISTORICAL AND FUTURE PROBABILITIES OF LARGE TERRORIST EVENTS

BY QIURONG CUI, KARL ROHE AND ZHENGJUN ZHANG

Department of Statistics at University of Wisconsin-Madison

First and foremost, we commend the authors for their creative and original investigation. Although this comment will focus on methodological concerns, we note that in general, the advantage of statistical modeling is not necessarily that the solutions are precise, but rather that all assumptions are made explicit. The conclusions in Clauset and Woodard (2013) [1] rely on the strong modeling assumption that the number of deaths in a terrorist incident are (a) independent draws from some unknown probability distribution that is fixed in (b) space and (c) time. Given the politically charged nature of this problem, we are wary of these three assumptions (and thus the conclusions) in the paper.

Many of the inferences in [1] (hereafter referred to as CW) rely on the bootstrap to measure the uncertainty in their statistical estimators. Similarly, other applied papers in the extreme value literature have relied on the bootstrap (e.g. [2]). However, there are many scenarios in which the bootstrap can fail. Both [3] and [4] discuss some of these problems in the context of heavy tailed distributions. In this discussion, we provide a brief simulation that illustrates when the bootstrap succeeds and when it fails in the settings of CW.

This comment investigates the following question under three relevant models:

If the bootstrap is used to create a (nominally) 90% confidence interval, will this interval actually cover the true parameter in 90% of experiments?

The first simulation model is the power law distribution with $\alpha = 2.4$ supported on $[10, \infty)$. The second simulation model comes from [5], a paper that CW cite to justify their method of estimating x_{\min} . To sample a point X_i from this model (which we will refer to as the mixed power law model), sample an observation Y_i uniformly at random from the RAND-MIPT data ([6]) and sample an observation Z_i from power law with $\alpha = 2.4$ (corresponding

CUI ET AL.

to the estimate in CW) and supported on $[10, \infty)$. Then,

$$X_i = \begin{cases} Y_i & \text{if } Y_i < x_{min} = 10\\ Z_i & \text{o.w.} \end{cases}$$

To investigate whether the the bootstrap techniques in CW are sensitive to model misspecification error, the final simulation model is the Generalized Pareto Distribution (GPD) [7]. The GPD distribution is specified by three parameters: location u, scale σ , and shape ξ . The cumulative distribution function of the GDP distribution is

(0.1)
$$F_{u,\sigma,\xi}(x) = \begin{cases} 1 - \left(1 + \frac{\xi(x-u)}{\sigma}\right)^{-1/\xi}, & \xi \neq 0, \\ 1 - \exp\left(-\frac{x-u}{\sigma}\right), & \xi = 0. \end{cases}$$

When $\xi > 0$, the GPD distribution is regularly varying at ∞ with index $-1/\xi$ and thus tail equivalent to the power-law distribution with $\alpha = \xi^{-1} + 1$. Therefore, the GPD distribution can be considered as perturbed power-law distribution. For simplicity, we set $u = 0, \sigma = 1$ in the following simulation study. For comparison purposes, ξ is set to be $1/(\alpha - 1) = 1/1.4$ for $\alpha = 2.4$.

In each of 1000 runs of the experiment, we sample n = 1000 data points from each of the simulation models, we use the code from CW to (1) fit the power law distribution and (2) compute the (nominally) 90% bootstrap confidence intervals for α and p, the probability of observing at least one catastrophic event. For each of the three models, we run the estimation two ways: First with $x_{\min} = 10$ given to the algorithm and, second, where the algorithm estimates x_{\min} . Finally, each bootstrap confidence interval is inspected to see if it contains the true values of α and p (see next paragraph for a discussion on computing the true value of p). In total, this creates six different simulation setups. The simulation results are given in the Table 1 and Figure 1.

To compute the true value of p, we follow the definition in CW

(0.2)
$$p = 1 - E \left[P(X < cat | X > x_{\min})^{n_{\text{tail}}} \right],$$

where $n_{\text{tail}} \sim \text{Binomial}(n = 1000, p_{\text{tail}})$, and *cat* is the size of 9/11 (2749). In all six simulation setups, this value of p is approximately the probability that the maximum of n = 1000 draws is greater than *cat*.¹

In the power law and mixed power law distributions $x_{\min} = 10$ is the lower end of the power-law distribution for the tail model. Using Equation 0.2, p = .3194 under the power law and p = .0244 under the mixed power law.

¹In fact, for n = 1000, they are equal in the first four decimal places.

DISCUSSION

Under the GPD distribution an exact x_{\min} is not available for computing p. We therefore applied an analogy of Kolmogorov-Smirnov criteria to minimize the maximum distance of the GPD distribution above x_{\min} and power-law distribution, i.e., we find (x_{\min}, α) that minimize

$$f(x_{\min}, \alpha) = \max_{x: x > x_{\min}} \left| \left(\frac{\sigma + \xi(x - u)}{\sigma + \xi(x_{\min} - u)} \right)^{-1/\xi} - \left(\frac{x}{x_{\min}} \right)^{1-\alpha} \right|.$$

Although the limiting behavior of x_{\min} is left an undeveloped problem in CW, the solution of the above optimization problem could be a possible option heuristically. Numerical optimization (grid search in matlab) yields $\alpha = 2.31$ and $x_{\min} = 13.44$. With these values, the value of p in Equation 0.2 is .0242.



Fig 1: One thousand confidence intervals for p under the mixed power law distribution with $x_{\min} = 10$ (on left) and GPD with x_{\min} estimated (on right). These confidence intervals have poor coverage properties for the true value of p (represented as a red line). Each confidence interval is computed from a sample of n = 1000 data points and 1000 bootstrap samples. The order of the intervals is shuffled so that \hat{p} (blue line) is increasing.

Table 1 reports the results for the bootstrap confidence intervals for both the probability p of a catastrophic event and α , the power law parameter. In brief, out of the six different setups, the confidence intervals failed for two of them. Under both (1) the mixed power law distribution (from [5]) with x_{\min} given and (2) the GPD with x_{\min} estimated, the nominally 90% bootstrap confidence intervals cover the true values of p with probabilities .64 and .71 respectively. One reason the bootstrap fails under the mixed power law distribution is possibly due to an observation in CW, that a fixed choice of x_{\min} underestimates the uncertainty in \hat{p} due to the tails unknown structure. The reason the bootstrap fails under the GPD is that the algorithm tends to underestimate x_{\min} and α , i.e. it is inclined for heavier tails. This is

CUI ET AL.

consistent with the other discussants who suggest that x_{\min} is potentially too small.

Although the bootstrap gives a straightforward path to computing confidence intervals, these simulations suggest that their coverage performance is sensitive to the data generating model and whether or not x_{\min} is estimated or known.

TABLE	1
-------	---

For each of the three simulation models, x_{\min} is either given as 10 or estimated (est.) by the algorithm. The table presents (a) the estimated bias in coverage probabilities for the (nominally) 90 % bootstrap confidence intervals and (b) the median width of the

 $confidence \ intervals$

Model	x_{\min}	Coverage bias (%)		width of CI	
		α	p	α	p
$PL, \alpha = 2.4$	10	-1.0	9	.1	.2
PL-Mix, $\alpha = 2.4$	10	-28.9	-25.3	.7	.03
$\text{GPD}, \xi = 1/1.4$	10	6	2.4	.6	.1
Power Law, $\alpha = 2.4$	est.	+1.5	+.9	.2	.3
PL-Mix, $\alpha = 2.4$	est.	+.5	+.4	.8	.1
$GPD, \xi = 1/1.4$	est.	-11.6	-18.4	.5	.2

References.

- [1] CLAUSET, A. and WOODARD, R. (2013). Estimating the historical and future probabilities of large terrorist events. The Annals of Applied Statistics.
- [2] MOHTADI, H and MURSHID, AP (2009). Risk of catastrophic terrorism: an extreme value approach. Journal of Applied Econometrics. 537-559.
- [3] RESNICK, SI (2007). Heavy tail Phenomena; Probabilistic and Statistical Modeling. Springer.
- [4] HALL, P. and WEISSMAN, I. (1997). On the estimation of extreme tail probabilities. The Annals of Statistics. 1311-1326. MR1447753
- [5] CLAUSET, A. and SHALIZI, CR and NEWMAN, M EJ, (2009). Power-law distributions in empirical data. SIAM review. 661-703.
- [6] MIPT. (2008). Terrorism Knowledge Base. (acessed 19 August 2013).
- [7] COLES, S. (2001). An Introduction to Statistical Modeling of Extreme Values. Springer. MR1932132

DEPARTMENT OF STATISTICS UNIVERSITY OF WISCONSIN, MADISON, WI 53706 USA E-MAIL: karlrohe@stat.wisc.edu

4