The Annals of Applied Statistics
2013, Vol. 7, No. 4, 1891–1894
DOI: 10.1214/13-AOAS614F
Main article DOI: 10.1214/12-AOAS614

DISCUSSION OF "ESTIMATING THE HISTORICAL AND FUTURE PROBABILITIES OF LARGE TERRORIST EVENTS" BY AARON CLAUSET AND RYAN WOODARD

BY QIURONG CUI, KARL ROHE AND ZHENGJUN ZHANG

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 the conclusions in Clauset and Woodard (2013) rely on several strong modeling assumptions. For example, the number of deaths in a terrorist incident is an independent draw from some unknown probability distribution that is fixed in space and time. Readers who are more interested in how Clauset and Woodard (2013) contribute to the discussion on foreign policy and/or national security should 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. Given the politically charged nature of this problem, we are wary of the assumptions (and thus conclusions) in the paper.

Many of the inferences in Clauset and Woodard (2013) (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., Mohtadi and Murshid (2009)]. However, there are many scenarios in which the bootstrap can fail. Both Resnick (2007) and Hall and Weissman (1997) 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 Clauset, Shalizi and Newman (2009), 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 [MIPT, 2009] and sample an observation Z_i from power law with $\alpha = 2.4$ (corresponding to the estimate in CW) and supported on $[10, \infty)$. Then,

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

Received August 2013; revised August 2013.