Response to the reviewers' comments on the article

Statistical Inference for inter-arrival times of extreme events in bursty time series

Katharina Hees Peter Straka Smarak Nayak

11 März 2020

We would like to thank the reviewers very much for their careful reviews and the helpful comments that gave us the opportunity to improve the manuscript. Below we explain point by point how we addressed the comments during our revision of the manuscript.

Referee 1

Major points and questions

"It should be noted, however, that this process has light-tailed Mittag-Leffler marginals, while the authors of the paper under review encounter a different, heavy-tailed Mittag-Leffler law."

- We have added the following sentence near the end of Section 3: "We caution the reader that, somewhat confusingly, there is another distribution called the"light-tailed" Mittag-Leffler distribution. This is in fact the limiting distribution of the renewal process $N(t)$ above (see @limitCTRW)."
- We have also added the following reference: Meerschaert, Mark M, and Hans-Peter Scheffler. 2004. "Limit Theorems forContinuous-Time Random Walks with Infinite Mean Waiting Times."J. Appl.Probab.41 (3): 623–38. https://doi.org/10.1239/jap/1091543414.

"(p.5, l. 9-13) Why is it true that the limiting distribution exists? It is well known…"

- In case of a continuous distribution one can choose the normalization in such a way that a non-degenerate limit distributions exists, but this is not neccessarily true for a linear normalization. We just forgot to add the assumption, that the distribution of the magnitudes is out of the domain of attraction of an extreme value distribution. We added this accordingly.

"Note: It is important to observe that due to the fact that the waiting times Wj 's have infinite means, the renewal counting process N(t) cannot be made stationary (through a random shift) and hence this model does not fall into the classic context of Hsing-Leadbetter and Davis-Hsing."

- We have added a remark at the end of Section 3, saying: "When $0 < \beta < 1$, the renewal process $N(t)$ is *not stationary*, and hence the results by @Hsing88 on the exceedances of stationary sequences do not apply."

"Given the first comment above, can you elucidate the connection between the two different kinds of Mittag-Leffler distributions – the light-tailed distributional limit for the counting process N(t) and the heavy-tailed limit model for the T(l)'s?"

- We have addressed this, see first point.

"The use of k is very confusing … These two sentences are so unclear, so poorly written… Be specific, choose your words carefully and introduce labels/titles in the plots."

- We have re-written the entire section titled "Simulation study".

“(Section 5 and 6) In these 2 sections it is implied that p = k/n . . . . While a complete rigor in the discussion on the statistical methodology is perhaps not necessary, itis important for the reader to clearly see where “asymptotic approximations” are applied."

- We have re-written Sections 5 & 6, and made sure that p is used as the theoretical exceedance probability, whereas the fraction of exceedances k/n is denoted as $\hat{p}$. We have also clarified throughout where we assuming approximations.

“(Section 5) The methodology in this section (pages 9 and 10) is not very clearly described. Why give 2 algorithms if you are using just the second one, mostly? . . . The scaling/rescaling of the scale coefficient by the choice of l and k is very confusing. Please revise Section 5 thoroughly.”

- We have re-written Section 5, as per the referee’s suggestions, and we can now do without the “practical adjustments” section.

“(page 11, line 12) Poor scientific language: “. . . the hill estimator gets useless . . . Also a few lines after that, they claim that the scale cannot be estimated. This s not entirely true. The classic work of Peter Hall { entitled On Some Simple Estimates of an Exponent of Regular Variation and published in the Journal of the Royal Statistical Society. Series B (Methodological), Vol. 44, No. 1 (1982), pp. 37{42 { does include natural asymptotically normal estimators for both the tail exponent and the scale. This said, there is a tremendous amount of work on bias correction for the Hill estimator (and I assume the scale) that the authors can implement, should they desire. I realize that this may be going beyond the scope of their work, but proper pointers to the literature should be given and the notion of second order regular variation (required to establish asymptotic distributions) should be mentioned.”

- KATHI TO DO

“Last but not least, given that the authors are dealing with more or less standard semi- parametric estimation of heavy-tail exponents and scales, they could formulate limit the- orems on the asymptotic distribution of their statistics and derive Wold-type confidence intervals for the estimated parameters. . . . ”

- In the last paragraph of Section 5, we have included pointers to the literature for confidence intervals for the shape and scale parameters of the Mittag-Leffler distribution, and how these lead to “confidence bands” in our plots, based on the CTRE package.

Typos and suggestions

1. fixed
2. fixed
3. fixed
4. fixed
5. fixed
6. fixed
7. fixed
8. fixed
9. fixed
10. fixed
11. fixed
12. fixed
13. fixed
14. fixed
15. We reformulated the whole part and added titles.
16. fixed
17. no longer applies
18. fixed
19. fixed
20. fixed
21. fixed

22. fixed
23. fixed
24. fixed
25. fixed
26. The section has been re-written
27. fixed
28. fixed
29. fixed
30. fixed
31. fixed
32. fixed
33. We have commented as follows: “The stability plot for the tail stabilizes nicely around 0.85 (dashed line), while the scale parameter stabilizes less obviously near 3×10^7 (dashed line). The growth of the scale parameter for lower threshold appears to be closer to linear in p, rather than proportional to $p^{1/0.85}$ as suggested by the Mittag-Leffler fits. The reason for this is likely that the overall goodness of fit as compared to an exponential distribution is improved due to the peaked shape of the Mittag-Leffler distribution near 0, rather than its tail behaviour at ∞. The reported fit should hence come with the caveat that a Mittag-Leffler distribution models exceedance times well only up to certain time-scales. More research is needed into the modelling of scale transitions, where inter-exceedance times appear to have different power laws across different time scales.”
34. We have changed “fitted” to “estimated”.
35. We have added confidence limits.
36. We have added titles to the two figures, and to other figures where feasible.
37. fixed
38. fixed

Referee 2

“I understand the authors find inspiration in Peaks-Over-Threshold (POT) method of extreme value theory, although, I do not see a real connection be- tween two models. Observe that the peaks over threshold l are ignored, so random variable $\tau(l)$ could be substituted by any other independent geo- metric (or exponential) random variable with mean $1/p(l)$.”

- TO DO

“Moreover, I think, this would simplify some arguments, including the proof of the theorem.”

- TO DO

“The CTRE model considered here can be viewed as a special family of renewal processes (with power-law waiting times). Their relatively simple structure makes them amenable to theoretical analysis but for the same rea- son, I am not convinced that the CTRE model represents a really good alternative to cluster models as a model for bursty data. The authors’ claim that”visually similar statistical behaviour" can be seen in some real life data set seems a bit too vague to me."

- To DO

“Although auto-correlation plots given in Figure 7, after log-transforming data (this does not always ensure infiniteness of moments note) seem to indicate limited dependence in the sequence $logT(l, n)$, it is well known that the choice of such a transformation can in influence appearance of a corresponding graph.”

- To DO

“The authors finally apply two methods (log-moments and MLE) to es- timate parameters of the limiting Mittag-Leffler distribution, and obtain in the process an estimator of the key tail parameter β. These methods seems to outperform the standard Hill estimator in a simulation study although the latter utilizes a much larger data set typically. This phenomena asks for at least some explanation and a further study, because if it holds more generally, the authors could use their approach as a superior alternative to the Hill estimator for any power law distribution with $\beta < 1$."

- TO DO

"The manuscript seems nicely organized. However, writing is not always precise, the paper contains some typos and not entirely rigorous mathemati- cal arguments.""

- TO DO

"I also do not see a clear connection with the POT method suggested in the abstract, mainly because exceedances really play insignifficant role in the analysis.""

- TO DO

"P.3. Definition MPR uses the left and the right end point of the distribution without precise explanation. The role of the sequence (J_k) in the rest of the text should be also discussed."

- We added a definition for the left and the right point. NOCH ETWAS ZU DEN J_k ERGAENZEN.

"P.3. in the CTRE definition, in $(T(l,n), W(l,n))_n$, one might expect $X(l,n)$."

- You are completely right. This was a typo. We changed this accordingly.

"P.3. Probably $J - l | J > l$?"

- You are completely right again. This was also a typo. We changed this accordingly.

"P.4. In the first line of the theorem waiting times should be W_k. To avoid confusion, it would be better to use p_l instead of p and some other letter to denote the limit instead of W_β."

- We changed this accordingly.

"P.5. In the proof you use the fact that J_k have continuous distribution, but this has not been assumed. You also use the fact that J_k belong to some maximal domain of attraction - which is not assumed either."

- Thank you. We added an assumption accordingly.

"P.5. You should be careful about which parameter tends to infinity in your limiting relations, it seems that $c \to \infty$ is missing at several places, and $t \to \infty$ appears instead at one point."

- We fixed this.

"P.5. You seem to to assume that your l grows at a very precise rate - could this be a problem?"

- PETER TO DO

"The last two lines of the proof need more explanation, I think."

- We added some more explanation to the last two lines.

"P.6. threshold"

- Fixed.

"P.7. Could you comment worse performance of the estimator of the scale for small tail indices."

- TO DO

"P.7. the Hill or hill estimator?"

- Fixed.

"P.9. In step 3. could you explain how the plot"stabilizes" as l increases since one gets very few observations in that region? Argument behind step 4. should be written out, in my opinion."

- You are right. We should better explain this. The plot stabilizes in a region in the middle. We reformulated step 3 accordingly. We also clarified the argument for what was written before at the end of step 4 and added a few sentences for this, following the algorithm.

"p.11. Do you mean lower variance or lower bias?""

- In the pareto case the Hill estimator is less biased, but in all cases it has a higher variance compared to the other two estimators. We reformulated this part.

"P.11. I do not see how you derived the conclusion: Therefore, the mle is..."

- The formulation "Therefore" was wrong. We changed the text accordingly.

"P.14. Is that last formula an equality or an approximation?"

- This is an approximation. We replaced the equality sign by the approximately-equal sign.

"P.18. Space forgotten in 'assumption.For'"

- Fixed.

"P.19. and earlier, Mark. M Meerschaert seems to be the only author cited with his full name and initials - maybe one should strive for uniformity here."

- We changed this accordingly.