

AE comments on TAS-11-231

Both reviewers provide a number of useful comments as well as important criticisms that need to be addressed in an acceptable revision. I largely endorse these reviews and do not repeat what they say except in cases where I want to amplify.

Main comments

1. Figures 1 and 2 really display a bivariate example, e.g., suitable for the bag-plot section. As a motivating example, it would be better to have data that truly come from several populations, rather than having to be shoehorned into this framework by grouping. You could have grouped in the other direction and the example would be no less (or more) meaningful.
2. Echoing many comments from Referee 1. The “ideal tail area” column is useless beyond the 4th row or so—the same rows where “rough %” has exactly the same information. I suggest showing the concept: $2^{-1}, 2^{-2}, 2^{-3}, \dots$. Also, I do not see why “SEfactor” is displayed to 6 decimal places in the 1st entry and only 2 thereafter.
3. Discussion accompanying Table 1, as well as its caption, is very convoluted and certainly is aimed at a small group of insiders, not the general reader. Somewhere in there I get the idea that $n = 120$ is kind of a pivotal quantity; and maybe even that I shouldn’t even draw a boxplot if the sample size is smaller than that? Should some (all?) of this be moved forward to section 4? Do we even need Table 1, given what’s in Table 2?
4. The discussion on p.9 and Figure 3 seems tangential; I could have skipped that page and not missed a thing, thematically.
5. Pages 12–17. Lots of reading on relatively arcane points. Please condense. And as I said above, maybe bring in what’s needed from earlier material.
6. Referee 2 mentions that the “bag” may be incorrectly defined in Section 6. This definitely needs to be addressed, and could account for the differences in the inner portions of the two plots in Figure 8.
7. I think referee 2’s concern about being based on uncontaminated data arises in part from the fact that you do not illustrate any contaminated data (except perhaps in the bagplot, where the stray text on p.18 may be an abortive attempt to discuss that issue). (On the other hand, the outliers are clustered, not wildly scattered like we usually regard them to be.) We do not see how the ordinary letter-value plot behaves when there are a lot of outliers; examples are sorely needed.

Minor points/corrections

1. I agree with referee 1 that the title should say “datasets”
2. p.3, pos 52: ... from the same distribution ...

3. The term “outlier” appears in quotes every time it appears; this is quite distracting and unnecessary. Just use the word, no need to decorate it. Maybe the first time is OK, where you clarify what you’re talking about and hence don’t need to quote it any more.
4. p.5, pos 23, “free of tuning parameters.” Not quite: you still have to decide how many letters to display; that’s a tuning parameter.
5. Personally, I don’t like the plotting symbol used for outliers. They look like fireworks bursts, and are too big. I suggest small hollow circles, like in the boxplot examples, because you can more easily visualize two nearby ones. Another possibility (maybe even better) would be small hash marks.
6. p.10, pos 10: it’s vertical only if the data scale is horizontal.