Co-author comments

randy

2023-11-10

Stef's comments

I enjoyed reading this paper. I think it provides useful insights into the practical application of PLM approaches. Below I list some suggestions and comments to improve the manuscript. Please incorporate as you see fit. I trust you are able to deal with these points in the submitted version, so I do not need to see it again.

1. Readers not familiar with existing PLM approaches might find it difficult to see how the present manuscript extends current approaches. In the abstract and the introduction, make clear that **existing approaches are based on predictive mean matching of a single future outcome**. Your work extends matching to sequential multivariate future outcomes. I think adding "future" is important because a reader unfamiliar with PLM might think that matching is only possible using past measurements. Also, motivate the work, for example by saying that the new method is designed to predict well according to a range of future ages, rather than at a single future age.

This is an easy comments to solve, just add extra sentences and comments on

2. Page 2: Improve flow by holding back on the method details. Reserve "The specific steps are as follows: . . . " for the method section later on. Or: replace by a short introduction to finding matches if you need that to explain to objective of the paper.

- Move the details into methods section

3. Page 5: "Imputing" may not be the right term to describe step 1. "Imputing" suggests that you are filling up unobserved parts of the data, whereas I think what you are doing is replacing each individual trajectory by a small set of repeated measures that feeds into step 2. Step 1 feels more like summarizing and creating consistency across individuals to ease subsequent analyses and interpretation.

- What should we call this step? Augment, Supplement, ...

4. Page 5: Step 2 "linear model" could perhaps be labeled as something like "Add person baseline"

- Easy to fix with extra comments

5. Page 5: "Matching". There are two notations (y_1, y_2) and $(y_i, y_\star)$ to refer to the same operation. I suggest only using the second.

6. Page 6: The idea of the Chi-square is neat and an important contribution of the paper to set the number of matches. A reviewer might ask: Does line 175 hold if y_i and y_* are estimated and smoothed? (I didn't read Zhang) The correlations between estimated y's are higher than between observed y's, so I am not sure whether $D_M \tilde{\chi}^2$ holds in your case. Please check. It might be possible to repair it **by adding a ridge to Sigma** equal to **the variance of the error of the brokenstick model** ("impute" using brokenstick actually **adds noise to the broken stick predictions to get to an unbiased estimates of Sigma**). For your use, it actually doesn't matter whether it holds since you are using it as a tuning parameter rather than for testing.

7. I missed the connection between κ (number of matches) and alpha throughout the report. How large is κ for a given α? On page 8, you explain it varies by individual. Does is make sense to report average κ over individuals?

Need to address this problem.

8. Page 8: Step 4 uses a GAMLSS models. Such models can have many parameters. Can you still estimate these for 10 matches?

In fact, we use the GAMLSS only as a smoothing function... Stef may already noticed. as in comment 12

9. Page 9: Lines 219-221. You have restricted the paper to including only the baseline. Please motivate, and say that it is also possible to condition on more measurements and update using PLM, but that this was not done for brevity.

- Simply address this motivation

10. Table 2: The results on 90% CR are impressive (it is very hard to get these right). Could the slight undercoverage for $\kappa = 10$ perhaps be due to the fact that y's are estimated by brokenstick? (see also point 3 and 6).

- as well as the uncertainty in the estimating section. Which we also want to comment in the paper

11. Figure S4: All models miss the pubertal spurt and subsequently ceasing in growth around ages 14-15. It seems that a simple smooth of the median of the matches might be more accurate here than the GAMLSS model.

- Yes, we should have simply use a smoothing function, other than the GAMLSS