

# Jining Qin Section 4-5 Comments

Brendan McVeigh

February 22, 2016

## General Comments:

As I have mentioned with other sections of this piece I really like have thorough most of the explanations are and the places where specific steps are listed. I think that this does a good job of making the procedural explanations very clear. However, I do think that I at least have a tendency to lose track of the big picture and get lost in the weeds, it might be helpful to add a sentence or two to some of the sections reiterating how the step described helps with the ultimate goal of the paper. This could also allow you to maybe remove some of the references to previous sections. Feel free to disagree but I worry about these because they encourage the reader to look back at the previous section and interrupt the flow of the reading. If you think they need a refresher on a term or definition it is good to just provide it again. This way if the reader already remembers it they feel smart.

Also on a flow note you use the word "Therefore" a lot. In general while being repetitive in technical writing is good for specific terms and definitions being too repetitive can be a really grating to the reader. Mixing in some synonyms (so, thus, hence,...) might be helpful. Not that you never use the other options just try not to do it twice in three or four sentences.

## Specific Comments:

- Page 7, sentence 2: I think you can delete "As mentioned earlier" if the definition is worth reiterating then just go for it.
- Page 7, start of paragraph 2: split into two sentences with the second one starting "There are several sets...". You are trying to make more than one point here so use more than one sentence
- Page 7, paragraph 1: Consider splitting into three paragraphs with the second starting on line 4 "In the discrete bin approach..." and the third starting on line 9 "In the continuous approach". As one paragraph it feels like there is too much material and I don't see a natural way to break it into just 2.
- Page 7, paragraph 2, last sentence: I would delete "can" in "We can then..."
- Page 7, section 3.1, first sentence: You could replace "As mentioned in 2.2 we have" with
- Page 7, section 3.1, first paragraph, last sentence: Delete this sentence, it feels like scaffolding to me, in its current form. Either that or expand to talk more about the properties.
- Page 8: the transition from page 7 seems very abrupt. Maybe add a sentence or two after the numbering at the end of page 7?
- Page 8, paragraph 2, sentence 2: "Further recall... in 2.1." again if you are re-describing things, which is often good, this can probably be taken out. If you think they might need to look back at the previous section then it would be good to add more detail here.
- Page 8, last sentence: I for this sentence a little confusing the first time or two I read it, an example might be helpful.

- Page 9, last sentence before 3.2.3, It might be more effective to point to some plots, or mentioned that they are in the appendix here. There are many ways of comparing things like the proportion of zeros, even just a table, so unless there is a reason to use a boxplot specifically I might be less specific about the method of comparison.
- Page 10, first sentence: you are trying to make too many different points in one sentence, Try splitting it into (with some revisions) "Under the null... the same distribution. We pool the statistics together... pooled data to create three new sets of feature statistics  $\{Y_i^{(1)} \dots Y_i^{(3)}\}$ . The new sets of feature statistics correspond..."
- Page 10, section 3.3, last paragraph: I am not sure this belongs here, the previous paragraph segues into the next section nicely. Would it be possible to move the either the end of 3.3.1 or the beginning of 3.3.2 since it refers to 3.3.2?
- Page 11, paragraph 2, I think you should switch the order of the first two sentences
- Page 11, first sentence after numbering: don't need "For example"
- Page 11, last paragraph, Table 4 should be referenced somewhere and I think there are 119 galaxies with more than one mode.
- Page 12, section 3.3.2 first sentence: Combine first two sentences into something like "retaining all... as in 3.2.1 means there is no longer..."
- Page 12, section 3.3.2, third sentence: Also combined 3rd and 4th sentence since it is the same idea "One way... under different filters, using the probability..."
- Page 13, Figure 8 description: I think it would be helpful to clarify that the spurious modes are the small ones, or ones close to larger modes (assuming I can correct about this), point to an example in one of the plots, maybe the one in the bottom right.
- Page 14, section 3.3.5, first line: I don't think it makes sense to have a paragraph break after the first sentence
- Page 14, last sentence: is  $\hat{t}$  suppose to be  $\hat{\lambda}$ ?
- Page 15, section 3.3.6: last paragraph: I think this could be put in a much more straight forward manner, it current as more of a stream on consciousness feel. Maybe something along the lines of "We estimate the 25th, percentile, the median, and the 75th percentile, the standard benchmark quantiles of a distribution."
- Page 15, section 3.3.7: You mention two different methods for getting confidence intervals, for me this naturally raises the question of have you looked at both of them and if so do they give similar results. If the answer is yes then great, otherwise give some sense of why not.