Review of *Validation of Visual Statistical Inference, Applied to Linear Models*

Grace Yi Chen

Graphics are used in many places like exploratory data analysis, model diagnosis, and statistical inference. This article introduced a visual statistical inference method using graphics. Compared with the traditional hypothesis test, the visual inference statistic is not a single value. The authors used the lineup protocol, which places the actual data plot among several plots of null data and asks a group of participants to pick the plot that is different from the rest. The authors demonstrated the advantages of the visual statistical inference method using three experiments. The first two experiments have ideal scenarios for conventional testing and the third one has a scenario where the assumptions for conventional testing are violated. Using the lineup protocol, the visual method performs comparably with the conventional tests in the first two experiments, and outperforms the conventional tests in the third experiment.

I think the authors explained the advantages of using the visual statistical inference method well with the lineup protocol. As we don't know if the data is contaminated or not, the visual statistical inference method will always work under both scenarios. I think it is a good try although I am not quite sure how useful it would be in real-world settings. Visual testing is different from conventional tests because the lineups will depend on participants' evaluation. For the lineup protocol, we first need to generate several plots from the null and mix them with one plot from the alternative. Then we will need to recruit a group of subjects from Amazon Mechanical Turk to evaluate if there is a difference. It is a novel way to test the hypothesis but might be very complicated and costly to do it. Also, as the authors mentioned, if there is a null plot with a small p-value, or the significance level is close to 0.05, it will be hard for participants to detect the actual data plot.

Question:
What would be an ideal real-world example to implement this method?