DISCUSSION OF "ASSUMPTION-LEAN INFERENCE FOR GENERALISED LINEAR MODEL PARAMETERS" BY VANSTEELANDT AND DUKES

H. S. Battey[1]
Department of Mathematics, Imperial College London

[Read before The Royal Statistical Society at a meeting organized by the Research Section on Tuesday, July 6th, 2021.]

For situations in which there is uncertainty over the underlying probabilistic model, there are at least three broad approaches. One is to seek reliable inference for interest parameters or perhaps, as the authors advocate, for quantities retaining at least a degree of interpretability under misspecification. Another is to acknowledge more explicitly the model uncertainty. A third approach, loosely connected to the first, is to encapsulate uncertainty over the model in a possibly large number of nuisance parameters, to be eliminated in the analysis by suitable conditioning arguments or other problem-specific manoeuvres (e.g. Bartlett, 1937). A helpful example is the use of partial likelihood to evade the baseline hazard function (an infinite-dimensional nuisance parameter) of a proportional hazards model. The appropriateness of each of the three approaches depends largely on context. I will constrain my discussion to the first two.

If the interpretation of an interest parameter is stable over models, it appears that first-order reliable inference via maximum likelihood estimation is possible in spite of considerable misspecification in the nuisance part of the model only when the interest parameter is orthogonal (in the sense of Jeffreys, 1948, pp. 158, 184) to the notional nuisance parameters, whose interpretation then has to be in terms of Kullback-Leibler projection. This would be a necessary condition rather than a sufficient one. Note that the true model is also implicit in the definition of parameter orthogonality. It is, as far as I am aware, an open problem to characterise the class of models whose interest and notional nuisance parameters are orthogonal under arbitrary model misspecification, perhaps after interest-respecting reparameterisation. The second-order properties are always affected, sometimes severely, which is of course problematic beyond point estimation. On a historical point, the limit in probability of the maximum likelihood estimator under model misspecification and its connection to the Kullback-Leibler divergence was derived by Cox (1961; 1962), who also noted the failure of Bartlett's second identity and gave a generalisation of the result (Cox, 1961, equations (28)-(43)) which later became known as the sandwich formula. A more rigorous discussion of regularity conditions was given by Huber (1967). Similar results were obtained independently by White (1982a,b).

It could be argued, contrary to the paper under discussion, that when the effects of interest are represented by parameters whose interpretations differ according to the model used, the appropriate approach is to acknowledge the model uncertainty rather than seek inference on a quantity whose interpretation is stable but perhaps only tangentially relevant when the assumed model is false. The role of sufficiency in assessment of model adequacy, implicit in R. A. Fisher's work, is perhaps best approached via Barndorff-Nielsen and Cox (1994, p.29). When the ideas can be operationalised, there are no difficulties associated with double use of the data for model assessment and parametric inference. The conclusion may be that some, all or none of the a priori plausible representations are compatible with the data. If multiple models with different interpretations are not significantly contradicted, it often seems appropriate to report as many as feasible, a point emphasised repeatedly by D. R. Cox (e.g. Cox, 1968; Cox and Snell, 1974, p.55, 1989, p.193; Cox, 1995). See also Davison (1995). This underlies the development of confidence sets of models (Cox and Battey, 2017).

[1]Email address: h.battey@imperial.ac.uk

REFERENCES

Barndorff-Nielsen, O. E. and Cox, D. R. (1994). *Inference and Asymptotics*. Chapman and Hall, London.

Bartlett, M. S. (1937). Properties of sufficiency and statistical tests. *Proc. R. Soc. Lond. A*, 160, 268–82.

Cox, D. R. (1961). Tests of separate families of hypotheses. In *Proceedings of the fourth Berkeley Symposium on Mathematical Statistics and Probability*, 1, edited by L.M. LeCam, J. Neyman and E.L. Scott. University of California Press, Berkeley, 105–123.

Cox, D. R. (1962). Further results on tests of separate families of hypotheses. *J. R. Statist. Soc. B*, 24, 406–424.

Cox, D. R. (1968). Notes on some aspects of regression analysis (with discussion). J. R. Statist. Soc. A, 131, 265—279.

Cox, D. R. and Snell, E. J. (1974). The choice of variables in observational studies. *J. R. Statist. Soc. C*, 23, 51–59.

Cox, D. R. and Snell, E. J. (1989). *Analysis of Binary Data*, 2nd edition. Chapman and Hall, London.

Cox, D. R. (1995). Discussion of the paper by Chatfield. *J. R. Statist. Soc. A*, 158, 455–456.

Cox, D. R. and Battey, H. S. (2017). Large numbers of explanatory variables, a semi-descriptive analysis. *Proc. Nat. Acad. Sci.* 114, 8592–8595.

Davison, A. C. (1995). Discussion of the paper by Chatfield. *J. R. Statist. Soc. A*, 158, 451–452.

Huber, P. J. (1967). The behavior of maximum likelihood estimates under non-standard conditions. *Proceedings of the fifth Berkeley Symposium on Mathematical Statistics and Probability*, 1, edited by L.M. LeCam and J. Neyman. University of California Press, Berkeley, 221–233.

Jeffreys, H. (1948). *Theory of probability*, 2nd ed. Oxford.

White, H. (1982a). Maximum likelihood estimation of misspecified models. *Econometrica*, 50, 1–25.

White, H. (1982b). Regularity conditions for Cox's test of non-nested hypotheses. *J. Econmetr.*, 19, 301–318.