Heather Battey's contribution to the discussion of 'Parameterizing and Simulating from Causal Models' by Evans and Didelez

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

My brief comment relates to the inferential discussion of section 5. The sample-space factorisation of the joint probability from the marginal structural model induces a parameter-space factorisation of the likelihood function, with the implication that θ_{ZX} is orthogonal to $(\theta^*_{Y|X}, \phi^*_{YZ|X})$ in the sense of e.g. Cox and Reid (1987), and leading to the simplified asymptotic covariance matrix of Theorem 5.1. The parameter cut discussed at the beginning of section 5 also suggests inference by partial likelihood (Cox, 1975), in which the part of the likelihood function involving the nuisance parameter $p_{X|Z}$ (or θ_{ZX}) is discarded. This raises the question of whether the propensity score can be completely bypassed in the paper under discussion, in an analogous way that partial likelihood evades the baseline hazard function in the proportional hazards model; both are in principle infinite-dimensional nuisance parameters. There may well be practical difficulties in the present context.

Evans and Didelez note (second paragraph of section 5) that Theorem 5.1 allows the propensity score model $p_{X|Z}$ to be misspecified thanks to the parameter cut. Recent work (Battey and Reid, 2024) has explored structure in parametric statistical models that guarantees consistency of the maximum likelihood estimator for a parameter of interest in spite of arbitrary misspecification of the nuisance part of the model, the interest parameter being common to both the true and the fitted models. The structure exploited in the paper under discussion is an example case, with $p_{X|Z}$ the nuisance component. The parameter cut is not a necessary condition for consistency; the latter can be ensured under one of two weaker conditions presented in Propositions 1.1 and 1.2 of Battey and Reid (2024) alongside example cases. This may shed some light on Evans's and Didelez's statement below Theorem 5.1 that "if the model is misspecified there is no guarantee that the estimator will be consistent or even close to the true value."

Although Theorem 5.1 looks to obviate the sandwich formula, $I(\theta^*)$ is the Fisher information under p_{ZXY}, which is assumed (first paragraph of section 5) to be correctly specified even if $p_{X|Z}$ is not. The implication, I think, is that $I(\theta^*)$ is not explicitly calculable when $p_{X|Z}$ is misspecified.

I enjoyed reading the authors' thought-provoking work.

REFERENCES

Battey, H. S. and Reid, N. (2024). On the role of parametrization in models with a misspecified nuisance component. *arXiv preprint arXiv:2402.05708*.

Cox, D. R. (1975). Partial likelihood. *Biometrika*, 62, 269–276.

Cox, D. R. and Reid, N. (1987). Parameter orthogonality and approximate conditional inference (with discussion). J. R. Statist. Soc. B, 49, 1–39.

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

© The Royal Statistical Society 2024. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com This article is published and distributed under the terms of the Oxford University Press, Standard Journals Publication Model (https://academic.oup.com/pages/standard-publication-reuse-rights)

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jrsssb/advance-article/doi/10.1093/jrsssb/qkae019/7608817 by guest on 23 March 2024