ISSN 1973-2201 *STATISTICA, anno LXXXIV, n. 2, 2024*
DOI: https://doi.org/10.60923/issn.1973-2201/22046

REJOINDER

Lutz Dümbgen [1]
Department of Mathematics and Statistics, University of Bern, Bern, Switzerland

I am very grateful to all discussants for their interesting and thoughtful comments. Here are some replies to the single contributors in alphabetical order.

Pietro Coretto

You are right that the present paper is related to ideas in Davies (1995) and Davies (2014) in the sense that the data are treated as fixed objects, and randomization is used to draw conclusions about the data. While in Davies (1995), the given data are compared with complete data sets generated from a stochastic model, the present paper is focussing on a more specific aspect, the relation of a response to covariates, and the original data are compared with partially randomized versions.

The question whether there is a unifying model-free framework which also incorporates the prediction problem is intriguing. One could argue that all test statistics we look at measure how well the response can be *approximated* by a linear function of the covariates, and that approximation is closely related to point prediction. If one has probabilistic forecasts in mind, it is less clear whether there is an entirely model-free approach. If one is willing to assume some stochastic modelling, maybe one could try to combine the idea of Gaussian covariate vectors or random rotations with the paradigm of conformal inference (Vovk *et al.*, 2022) for that purpose.

Efthymios Costa and Ioanna Papatsouma

Thank you for the additional references to the lasso method and selective inference methods. The work of Zrnic and Fithian (2024) is indeed intriguing. One can certainly adapt our methods for "local inference" in the sense that one restricts attention to a subset of the covariates and then computes, say, an equivalence region based on these preselected covariates only. In the model-free approach, the preselection could even be data-driven, anything is allowed. A more sophisticated answer might be to design a test statistic

[1] Corresponding Author. E-mail: lutz.duembgen@unibe.ch

$S(\boldsymbol{y}) = S(\boldsymbol{y}, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_p)$ which involves a preselection of covariates as a first step and then focuses on the selected ones. This is, admittedly, rather vague, but note that the restrictions on $S(\cdot)$ are rather mild.

Jan Hannig

Some "magic beta formula" in the context of matrix-denoising and low-rank approximations would definitely be very useful. Thank you for pointing out a challenging and highly relevant topic for future research.

Christian Hennig

Thank you for your comments on confidence and equivalence regions. It is true that this aspect of our paper is a bit more experimental, and some aspects of equivalence regions are easier to understand in a model-based context, using the statistical models as "test beds" in the sense of Tukey. These considerations are closely related to the ones of Davies (1995) and Davies (2014).

Moritz Herrmann and Michael Herrmann

Thank you very much for your thoughtful remarks on different scientific cultures and communities. Indeed, when working on the present paper and beyond, Laurie Davies and I realized that even our two views on statistical modelling differ in some aspects. The paper discussed here should invite scientists to appreciate both viewpoints and shed new light on connections and differences.

Alexandre G. Patriota and Andrey B. Sarmento

The non-monotonicity problem you point out is notorious in linear models. Presumably anyone working with regression analyses and comparing different models has been confused at some point by the fact that adding or removing a single covariate can have unforeseen effects on the p-values for the other variables. The s-values you propose are tempting indeed, but the resulting conclusions tend to could be rather conservative.

For the little data example you provided, I tried a different test statistic $S(\boldsymbol{y}) = S(\boldsymbol{y}, \boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3)$, namely,

$$S(\boldsymbol{y}) = \max_{j=2,3} \frac{(\boldsymbol{y}^\top \tilde{\boldsymbol{x}}_j)^2}{\|\boldsymbol{y} - \hat{\boldsymbol{y}}_o\|^2},$$

where $\tilde{\boldsymbol{x}}_j$ is the orthogonal projection of $\boldsymbol{x}_j$ onto $\boldsymbol{x}_1^\perp$, divided by its norm. With this test statistic, I obtained a (Monte Carlo) p-value of 6.4%, rather than the 11.7% from the F test statistic. This does not resolve the monotonicity problem, of course, but

indicates that our standard choices of test statistics may be suboptimal. Do there exist test statistics which resolve the monotonicity problem?

LARRY WASSERMAN

The model lean approach to statistical inference is intriguing. The particular procedure you describe reminds me of ideas from conformal inference. However, these ideas seem to be very much depending on an assumed supermodel with i.i.d. or at least exchangeable observations. In typical regression contexts, this would often be too restrictive, whereas the model-free approach proposed here avoids such assumptions.

REFERENCES

L. DAVIES (2014). *Data analysis and approximate models*, vol. 133 of *Monographs on Statistics and Applied Probability*. CRC Press, Boca Raton, FL.

P. L. DAVIES (1995). *Data features*. Statist. Neerlandica, 49, no. 2, pp. 185–245. URL https://doi.org/10.1111/j.1467-9574.1995.tb01464.x.

V. VOVK, A. GAMMERMAN, G. SHAFER (2022). *Algorithmic learning in a random world (2nd ed.)*. Springer, New York.

T. ZRNIC, W. FITHIAN (2024). *Locally simultaneous inference*. Ann. Statist., 52, no. 3, pp. 1227–1253.