ISSN 1973-2201 STATISTICA, anno LXXXIV, n. 2, 2024
DOI: https://doi.org/10.60923/issn.1973-2201/22027

DISCUSSION OF THE PAPER "CONNECTING MODEL-BASED AND MODEL-FREE APPROACHES TO LINEAR LEAST SQUARES REGRESSION" BY LUTZ DÜMBGEN AND LAURIE DAVIES (2024): THE MONOTONICITY PROBLEM

Alexandre G. Patriota [1]
Department of Statistics, University of São Paulo, São Paulo, Brazil
Andrey B. Sarmento
Department of Statistics, University of São Paulo, São Paulo, Brazil

Dümbgen and Davies (2024) provided analytical p-values for testing specific null hypotheses under linear models, which are of practical importance. In this discussion, we examine a feature of p-values that could lead to conflicting conclusions within the framework of linear models. The authors introduced the full model in their Equation (1) and demonstrated that the p-value to test the hypothesis

$$H_{p_o} : \beta_j = 0, \text{ for } p_o < j \leq p, \tag{1}$$

is p-value$(H_{p_o}, \boldsymbol{y}) = \mathrm{B}_{(n-p)/2,(p-p_o)/2}\left(\frac{||\boldsymbol{y}-\hat{\boldsymbol{y}}||^2}{||\boldsymbol{y}-\hat{\boldsymbol{y}}_o||^2}\right)$, where p_o, $\mathrm{B}_{a,b}$, $||\cdot||$, $\hat{\boldsymbol{y}}$ and $\hat{\boldsymbol{y}}_o$ are defined in the main paper. Consider two nested null hypotheses, namely, $H_{p_{o_1}}$ and $H_{p_{o_2}}$, where $p_{o_1} < p_{o_2}$. For instance, take $p = 3$, $p_{o_1} = 1$ and $p_{o_2} = 2$, then $H_{p_{o_1}} : \beta_2 = \beta_3 = 0$ and $H_{p_{o_2}} : \beta_3 = 0$. The hypothesis $H_{p_{o_1}}$ imposes a more stringent constraint than $H_{p_{o_2}}$, since if $\beta_2 = \beta_3 = 0$, then it necessarily follows that $\beta_3 = 0$ must also be true. That is, given the same data, the observed evidence against $H_{p_{o_1}}$ must be at least as strong as the observed evidence against $H_{p_{o_2}}$. Measures that adhere to this logical reasoning are said to be monotone. P-values often fail to uphold this reasoning (Schervish, 1996; Patriota, 2013, 2017). In the context of linear models discussed in Dümbgen and Davies (2024), we introduce a dataset (Table 1) with three regressors and ten observations. From our data, the p-values for testing $H_{p_{o_1}} : \beta_2 = \beta_3 = 0$ and $H_{p_{o_2}} : \beta_3 = 0$ are p-value$(H_{p_{o_1}}, \boldsymbol{y}) = 11.7\%$ and p-value$(H_{p_{o_2}}, \boldsymbol{y}) = 4.5\%$, respectively, illustrating a violation of monotonicity. These results paradoxically suggest stronger evidence against

[1] Corresponding Author. E-mail: patriota@ime.usp.br

TABLE 1
The dataset with three regressors and ten observations.

$\boldsymbol{y}$	0.55	0.26	3.93	1.21	1.89	3.09	2.46	1.55	2.63	1.86
$\boldsymbol{x}_1$	1	1	1	1	1	1	1	1	1	1
$\boldsymbol{x}_2$	2.37	-2.34	3.54	-0.26	5.69	-1.38	-1.34	3.03	-4.12	-5.19
$\boldsymbol{x}_3$	0.10	2.63	-2.35	3.55	1.08	-3.76	-4.76	1.78	0.88	0.84

reducing the model to '$\beta_1\boldsymbol{x}_1 + \beta_2\boldsymbol{x}_2$,' than to just '$\beta_1\boldsymbol{x}_1$', posing interpretative challenges for practitioners.

To circumvent this issue, one might consider the s-value proposed by Patriota (2013). This discussion outlines a partial solution where practitioners avoid joint testing of $\boldsymbol{\beta}$ with σ^2, treating σ^2 as a nuisance parameter. From the derivation in Dümbgen and Davies (2024), the p-value for testing the null hypothesis

$$H_{\{\boldsymbol{\beta}^0\}} : \boldsymbol{\beta} = \boldsymbol{\beta}^0 \tag{2}$$

is given by p-value$(H_{\{\boldsymbol{\beta}^0\}}, \boldsymbol{y}) = \mathrm{B}_{(n-p)/2,p/2}\left(\frac{||\boldsymbol{y}-\hat{\boldsymbol{y}}||^2}{||\boldsymbol{y}-\hat{\boldsymbol{y}}_{\boldsymbol{\beta}^0}||^2}\right)$, where $\hat{\boldsymbol{y}}_{\boldsymbol{\beta}^0} = \sum_{j=1}^{p} \beta_j^0 \boldsymbol{x}_j$. The following alternative measure for testing the general null $H_C : \boldsymbol{\beta} \in C$, which is called s-value in (Patriota, 2013), resolves the above logical contradiction:

$$\text{s-value}(H_C, \boldsymbol{y}) = \sup_{\boldsymbol{\beta}^0 \in C} \text{p-value}(H_{\{\boldsymbol{\beta}^0\}}, \boldsymbol{y}), \quad \text{with } C \subseteq \mathbb{R}^p,\ C \neq \varnothing.$$

Observe that s-value$(H_C, \boldsymbol{y}) = \mathrm{B}_{(n-p)/2,p/2}\left(\frac{||\boldsymbol{y}-\hat{\boldsymbol{y}}||^2}{||\boldsymbol{y}-\hat{\boldsymbol{y}}_C||^2}\right)$, where $\hat{\boldsymbol{y}}_C = \arg\min_{\boldsymbol{\beta} \in C} || \boldsymbol{y} - \sum_{j=1}^{p} \beta_j^0 \boldsymbol{x}_j ||^2$. The s-value was originally derived from confidence regions, representing the smallest significance level (one minus the confidence) at which C intersects with the confidence region. It is a possibility measure over the class of hypotheses and, in general, under any hypothesis involving only $\boldsymbol{\beta}$, the probability that the s-value is less than $\alpha \in (0,1)$ does not exceed α (see, e.g., Patriota and Alves, 2021). S-values are generally more conservatives than p-values, i.e., s-value$(H_C, \boldsymbol{y}) \geq$ p-value$(H_C, \boldsymbol{y})$. From the above data, the s-values for testing $H_{p_{o_1}}$ and $H_{p_{o_2}}$ are s-value$(H_{p_{o_1}}, \boldsymbol{y}) = 20.6\%$ and s-value$(H_{p_{o_2}}, \boldsymbol{y}) = 20.7\%$, respectively, maintaining a coherent conclusion. Given its broad applicability to other scenarios, what could prevent practitioners from adopting this measure?

References

L. Dümbgen, L. Davies (2024). *Connecting model-based and model-free approaches to linear least squares regression*. Statistica, 84, no. 2, pp. 65–81.

A. Patriota (2013). *A classical measure of evidence for general null hypotheses*. Fuzzy Sets and Systems, 233, pp. 74–88.

A. PATRIOTA (2017). *On some assumptions of the null hypothesis statistical testing*. Educational and Psychological Measurement, 77, pp. 507–528.

A. PATRIOTA, J. ALVES (2021). *A monotone frequentist measure of evidence for testing variance components in linear mixed models*. Journal of Statistical Planning and Inference, 219, pp. 43–62.

M. SCHERVISH (1996). *P values: what they are and what they are not*. The American Statistician, 50, pp. 203–296.