Robust testing for superiority between two regression curves

G. Boente[1*] and J.C. Pardo–Fernández[2]

[1] *Departamento de Matemáticas, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires and IMAS, CONICET, Ciudad Universitaria, Pabellón 1, 1428, Buenos Aires, Argentina; gboente@dm.uba.ar.*
[2] *Departamento de Estatística e Investigación Operativa, Universidade de Vigo, Campus Universitario As Lagoas-Marcosende, Vigo, 36310, Spain; juancp@uvigo.es*
[*] *Presenting author*

Keywords. *Hypothesis testing; Nonparametric regression models; Robust inference; Smoothing techniques.*

Let us assume that the random vectors $(X_j, Y_j)^{\mathrm{T}} \in \mathbb{R}^2$, $j = 1, 2$, follow the homoscedastic nonparametric regression models given by

$$Y_j = m_j(X_j) + \varepsilon_j = m_j(X_j) + \sigma_j U_j, \tag{1}$$

where $m_j : \mathbb{R} \to \mathbb{R}$ is a nonparametric smooth function and the error ε_j is independent of the covariate X_j. As is usual in a robust framework, we will assume that the errors ε_j are such that $\varepsilon_j = \sigma_j\, U_j$, where U_j has a symmetric distribution $G_j(\cdot)$ with scale 1, so that we are able to identify the error's scale, σ_j. When second moments exist, as it is the case of the classical approach, these conditions imply that $\mathbb{E}(\varepsilon_j) = 0$ and $\mathrm{Var}(\varepsilon_j) = \sigma_j^2$, which means that m_j represents the conditional mean, while σ_j^2 equals the residuals variance, i.e., $\sigma_j^2 = \mathrm{Var}(Y_j - m_j(X_j))$. The nonparametric nature of model (**??**) offers more flexibility than the standard linear model when modelling a complicated relationship between the response variable and the covariate. In many situations, it is of interest to compare the regression functions m_1 and m_2 to decide if the same functional form appears in both populations. In particular, we will focus on testing the null hypothesis of equality of the regression curves versus a one-sided alternative. Let $\mathcal{R}$ be the common support of the covariates X_1 and X_2 where the comparison will be performed. The null hypothesis to be considered is

$$H_0 : m_1(x) = m_2(x) \text{ for all } x \in \mathcal{R},$$

while the alternative hypothesis is of the following one-sided type

$$\begin{aligned} H_1 :\ & m_1(x) \le m_2(x) \text{ for all } x \in \mathcal{R} \text{ and } m_1(x) < m_2(x) \text{ for } x \in \mathcal{A}, \\ & \text{where } \mathcal{A} \subset \mathcal{R} \text{ is such that } \mathbb{P}(X_j \in \mathcal{A}) > 0, \text{ for } j = 1, 2. \end{aligned} \tag{2}$$

To protect against atypical observations, the test statistic to be considered is based on the residuals obtained by using a robust estimate for the regression function under the null hypothesis. More precisely, our proposal combines the ideas of robust smoothing with those given in **?** to obtain a procedure detecting root$-n$ alternatives. The asymptotic distribution of the test statistic is studied under the null hypothesis and under root$-n$ contiguous alternatives.The results of a Monte Carlo study performed to compare the finite sample behaviour of the proposed tests with the classical one obtained using local averages will be described.

References

Neumeyer, N & Pardo-Fernández, J. C. (2009). A simple test for comparing regression curves versus one-sided alternatives. *Journal of Statistical Planning and Inference*, **139**, 4006-4016.