

**Course: "Goodness of fit for regression models with applications"**

**Professor: Wenceslao González Manteiga**

**Motivation of the course:**

Since the seminal works by Pearson, who coined the term "goodness-of-fit" at the beginning of the 20th century to refer to the problem of fitting a distribution to a set of observations, much effort has been made in the statistical literature to develop goodness-of-fit techniques in an enormous variety of fields. In the particular context of nonparametric regression, the main contributions in testing problems have been developed in the last 25 years. These include testing for a parametric or a semiparametric specification for the regression function, testing for the equality of several regression functions, ,...etc

In this course, we will start with a review of classical techniques related to distribution problems, mainly based on the empirical distribution function or on estimators of density function. From there, we will explain how these ideas have been adapted to the regression context. Among others, we will cover the tests based on the comparison on nonparametric estimators and parametric estimators of the regression function, likelihood ratio tests, tests based on the integrated regression function and tests based on the empirical distribution of the residuals. Finally, we will cover very recent developments in more complicated settings, such as dependent data (temporal and spatio-temporal), directional data, incomplete data or functional data.

**Talk 1. Introduction to goodness-of-fit tests.** Classical ideas about goodness-of-fit tests. Desirable properties of a test. Tests based on the empirical distribution function. Tests based on estimators of the density function. Tests based on other functions that characterise the distribution.

**Talk 2. Introduction to goodness-of-fit tests in regression.** Main goodness-of-fit problems in regression. Description of the main methodologies. Basic tools in nonparametric regression: Nonparametric estimation, resampling methods.

**Talk 3. The goodness of fit in different setups.** Testing the equality of regression curves. Testing partial linearity. Testing generalized regression models. Testing significance. Testing additivity.

**Talk 4. The goodness of fit for models with dependent data.** Testing for time series. Testing for spatial/ temporal processes. Testing for point processes. Testing for diffusion models. Testing for functional data.

**Talk 5. The goodness of fit for models with complex data.** Testing for models with censored and /or truncated data. Testing for models with missing data. Testing for models with biased data. Testing for models with directional data.

References:

González Manteiga, W. and Crujeiras, R.M. (2013). Un adapted Review of goodness-of-fit tests for regression models. Invited paper with discussion. TEST, 22, 361-411.