# Testing the Equality of Nonparametric Regression Curves based on Fourier Coefficients

**Zaher Mohdeb**

University Constantine 1,
Laboratory of Mathematics and Sciences of the Decision,
Algeria

In this work we propose a new methodology for the comparison of two regression functions $f_1$ and $f_2$ in the case of homoscedastic error structure and a fixed design. Our approach is based on the empirical Fourier coefficients of the regression functions $f_1$ and $f_2$ respectively. More precisely, consider two regression models in fixed design given by

$$Y_{l,j} = f_l(t_j) + \varepsilon_{l,j}\,, \qquad l = 1,2 \ \text{ and } \ j = 1,\ldots,n\,,$$

where the design points $t_j$ are equispaced points and rescaled into the unit interval, $t_j = j/n$; for $l = 1,2$, $f_l$: $[0,1] \longrightarrow \mathbb{R}$, is unknown function and the errors $(\varepsilon_{l,j})_{j=1,\ldots,n}$, are i.i.d. random variables with mean zero and variance $\sigma_l^2$. We are interested in the problem of testing the equality of the regression functions $f_1$ and $f_2$, that is,

$$H_0 : \quad f_1 = f_2 \qquad \text{against} \qquad H_1 : \quad f_1 \neq f_2\,. \tag{1}$$

The problem of testing the equality of regression functions has been broadly studied in the literature. Recently, Neumeyer and Dette (2003) proposed a test for comparison of two regression curves that is based on a difference of two marked empirical processes based on residuals. Vilar-Fernández and González-Mantega (2003) studied the problem of checking the equality of $k$ regressions with dependent errors in a general context. Vilar-Fernández, Vilar-Fernández and González-Mantega (2007) studied the problem of testing the equality of regression curves with dependent data using bootstrap algorithm to approximate the distribution of the test statistics.

Our aim is to construct hypothesis test (1) based on the Fourier coefficients of the regression functions $f_1$ and $f_2$. More precisely, let

$$c_{l,k} = \int_0^1 e^{-2\pi i k t} f_l(t)\,dt\,, \quad l = 1,2 \ \text{ and } \ k \in \mathbb{Z}$$

be the Fourier coefficients of $f_l$, $l = 1,2$. The test (1) reduces to a test

$$H_0 : \ c_{1,k} = c_{2,k}\,,\ \forall k \in \mathbb{Z} \quad \text{against} \quad H_1 : \ \exists k \in \mathbb{Z} \ \text{ such that } \ c_{1,k} \neq c_{2,k}\,.$$

Since for $l = 1, 2$, $f_l$ is assumed to be a real-valued function, the null hypothesis $H_0$ is equivalent to $c_{1,k} = c_{2,k}$, $|k| \in \mathbb{N}$. It follows clearly that $H_0$ is true if and only if $\sum_{|k|\in\mathbb{N}} |c_{1,k} - c_{2,k}|^2 = 0$ .

The test statistic we use, is based on the empirical estimators $\hat{c}_{l,k}$ of $c_{l,k}$ defined by

$$\hat{c}_{l,k} = \frac{1}{n}\sum_{j=1}^{n} Y_{l,j} e^{-2\pi ikj/n}\,, \quad l = 1, 2$$

and we consider a sequence $p = p(n)$ of integers chosen such that $\lim_{n\to\infty} p(n) = \infty$. The test statistic is then based upon

$$\hat{T}_{n,p} = \sum_{|k|\le p} |\hat{c}_{1,k} - \hat{c}_{2,k}|^2$$

and we reject $H_0$ if $\hat{T}_{n,p} > t$ for some $t > 0$.

In our approach, the test statistic is used to test hypothesis on the Fourier coefficients of $f_l$, $l = 1, 2$. In the effective construction of the test, an estimation of the variance $\sigma_l^2$, $l = 1, 2$ is needed. This can be done in many ways. In the present work, we consider the Gasser, Sroka, and Jennen-Steinmetz (1986) estimator.

As our main results we obtain the asymptotic distribution of the test statistic under the null hypothesis, local and global alternatives. In order to investigate the finite sample performance and to study the power properties of the test, we conduct some Monte Carlo simulations.

**Keywords:** Nonparametric regression, Test of equality, Empirical Fourier coefficient.

## References:

Gasser, T., Sroka L. and Jennen-Steinmetz, C. (1986). Residual variance and residual pattern in nonlinear regression. *Biometrika*, **73**, 625-633.

Neumeyer N. and Dette, H. (2003). Nonparametric comparison of regression curves: an empirical process approach. *Ann. Statist.*, **31**, 3, 880-920.

Vilar-Fernández J. M. and González-Manteiga, W. (2003). Nonparametric comparison of curves with dependent errors. *Statistics*, **38**, 2, 81-99.

Vilar-Fernández, J. M., Vilar-Fernández J. A. and González-Manteiga, W. (2007). Bootstrap tests for nonparametric comparison of regression curves with dependent errors. *Test*, **16**, 123-144.