Construction of the criterion of a checking of the presence of structural shifts in researching of time series

Sergey E. Khrushchev[1], Artem V. Logachov[1,2,3], and Olga M. Logachova[1,2]

[1] Department of Mathematics and Natural Sciences, and Department of Statistics,
Novosibirsk State University of Economics and Management,
Lomonosova st. 56, 630099 Novosibirsk, Russia
s.e.hrushchev@edu.nsuem.ru

[2] Department of Higher Mathematics,
Siberian State University of Geosystems and Technologies,
Plakhotnogo st. 10, 630108 Novosibirsk, Russia
omboldovskaya@mail.ru

[3] Laboratory of Probability and Mathematical Statistics,
Sobolev Institute of Mathematics,
Academy Koptyug av. 4, 630090 Novosibirsk, Russia
omboldovskaya@mail.ru

Abstract. Well discuss a construction of the criterion (test) that allows us to check the hypothesis on the homogeneity and independence of sampling elements of random variables having the continuous distribution. The constructed criterion is exact and, in contrast to the various criteria of the series, does not require the imposition of conditions on the sample size and moments of random variables. The criterion does not depend on the distribution of the observed random variables and can be applied for samples of small volume also. This test is suitable for testing the hypothesis of homogeneity and independence of perturbations (errors) of regression models. The methodology for applying the developed criterion for revealing the structural shifts observed in time series is also described in the article. The structural shift in 2008 is revealed by revising of the dynamics of the Russian Federation gross domestic product in the period from 2000 to 2008 years.

Keywords. Sample, Criterion, Regression, Structural shift, Time series.

M.S.C. classification. 62J05, 62M10.
J.E.L. classification. C12, C22, C32.

Mathematical Methods in Economics and Finance – m²ef
Vol. 11/12, No. 1, 2016/2017

1 Introduction

We are interesting in the criterion that allows us to check the hypothesis on the homogeneity and independence of sampling elements (further, samples possessing these two properties will be called simple) of a small volume consisting of random variables having the continuous distribution. Note that classical regression models are constructed assuming that remainders of the regression are homogeneous and independent [7]. As a rule, the independence of remainders of the regression is verified using the Durbin-Watson criterion [9]. But this criterion gives reliable results only for samples of large volume.

Various series criterions are used most often for checking the simplicity of the sample in cases where nothing is known about the distribution of its elements [2]. Note that these criterions are asymptotic, it also imposes certain conditions on moments of the random variables included the sample in addition to the condition on the sample size, see, for example [2].

The criterion is obtained is independent of the distribution of random variables and can be used for small-volume samples.

Moreover, the obtained criterion can be used to identify structural shifts in the study of time series. Most often the identification of structural shifts occurs with the help of the Chow test [4], and CUSUM & CUSUMSQ tests also, [3]. Structural shifts in time series were also studied in papers [8], [1], [6].

Let $\hat{y}(x)$ denotes the function that defines the equation of a selective pairwise linear regression, e_i denotes residuals of a regression, α is the level of significance (the probability of making an error of the 1st kind), $\mathbf{Z}^+$ is the set of non-negative integers.

The paper is constructed according to the following plan: Review of well-known criterions is done and notations are introduced in section 1; The main result is proved in section 2 (Theorem 1); The received criterion for the identification of structural shifts in the study of time series is tested by some example in section 3.

2 The criterion

A sequence of random variables $\{X_n\},\ n \in \mathbf{Z}^+$ is considered.

We introduce the notation

$$N_{\min} := \min\{n \geq 1 : X_n < X_0\},\ N_{\max} := \min\{n \geq 1 : X_n > X_0\},$$

where $\min \emptyset = \infty$.

Lemma 1. *Let random variables $X_0, X_1, \ldots, X_n, \ldots$ are independent copies of a random variable X, which has a continuous distribution function. Then*

$$\mathbf{P}(N = n) = \frac{2}{n} - \frac{2}{n+1}, \quad n \geq 2, \tag{1}$$

where $N := \max\{N_{\min}, N_{\max}\}$.

Proof. It's obvious that $\mathbf{P}(N = 1) = 0$. By virtue of incompatibility of events $\{\omega : N_{\min} = n\}$ and $\{\omega : N_{\max} = n\}$ for $n \geq 2$ we have

$$\mathbf{P}(N = n) = \mathbf{P}(\{N_{\min} = n\} \cup \{N_{\max} = n\}) = \mathbf{P}(N_{\min} = n) + \mathbf{P}(N_{\max} = n). \quad (2)$$

Due to the fact that random variables $X_0, X_1, \ldots, X_n, \ldots$ are independent and identically distributed for any $n \geq 1$ we have

$$\mathbf{P}(N_{\min} > n) = \mathbf{P}(\min_{0 \leq k \leq n} X_k = X_0) = \frac{1}{n+1}.$$

Therefore

$$\mathbf{P}(N_{\min} = n) = \mathbf{P}(N_{\min} > n-1) - \mathbf{P}(N_{\min} > n) = \frac{1}{n} - \frac{1}{n+1}. \quad (3)$$

Similarly the following equality is obtained, see also [5]

$$\mathbf{P}(N_{\max} = n) = \frac{1}{n} - \frac{1}{n+1}. \quad (4)$$

Formula (1) is follows from the equalities (2)–(4). □

Remark 1. From Lemma 1 it is obviously that

$$\mathbf{P}(N > n) = \frac{2}{n+1}, \quad n \geq 2.$$

Let $X_0, X_1, \ldots, X_n$ is a sample.

Theorem 1. *Let the level of significance α is given and the hypothesis*

$$H_0 : \textit{ the sample is simple}$$

is verified against a competing hypothesis

$$H_1 : \textit{ the sample is not simple.}$$

Then the following criterion is valid:

$$if\ N \leq \frac{2}{\alpha},\ \textit{then the hypothesis } H_0 \textit{ is accepted},$$

$$if\ N > \frac{2}{\alpha},\ \textit{then the hypothesis } H_0 \textit{ is rejected.}$$

Proof. It is obviously follows from the Remark 1.□

3 Identification of structural shifts in time series

Let there some time series is analyzed in which so-called structural shift take place (the fundamental characteristics of the studied system have changed over time). Generally speaking, it means that the model before the shift and the model after the shift is different. For example, the economy underwent structural changes in 1998-1999 and 2008-2009 due to the crisis phenomena, so the parameters of macroeconomic models can be different before and after these moments.

Suppose that the changes of some indicator are described by a linear time series model:

$$y_t = a + bt + \varepsilon_t,\ t = 0, ..., n.$$

We will assume that a structural shift occurred at the time $t = n$. This leads not only to a changing of the model parameters, but also to the inhomogeneity of the perturbations ε_t. If there were no structural shifts, then the homogeneity of the perturbations ε_t will be preserved.

As mentioned above, the economy underwent structural changes in 1998-1999 and 2008-2009 due to the crisis phenomena. In this regard, let's consider the changing of the dynamics of the gross domestic product (GDP) in the period from 2000 to 2008, using the data on the Russian Federation GDP given in Table 1.

Table 1. The Russian Federation GDP in the period from 2000 to 2008 (in 2008 prices, billion rubles)

2000	2001	2002	2003	2004	2005	2006	2007	2008
24799.9	26062.5	27312.3	29304.9	31407.8	33410.5	36134.6	39218.7	41276.8

In the period between the crises, from 2000 to 2007, the GDP dynamics was described by the following model:

$$y_t = 23802.7 + 2043.9t + \varepsilon_t,\ t = 0, ..., 7. \tag{5}$$

The determination coefficient R^2 equals 0.98, and values of the residues ε_t are equal 997.2; 215.9; -578.3; -629.5; -570.5; -611.8; 68.4; 1108.6.

In the period from 2000 to 2008, the GDP dynamics was described by the model

$$y_t = 23628.0 + 2118.8t + \varepsilon_t,\ t = 0, ..., 8. \tag{6}$$

The determination coefficient R^2 is the same as for the model of the period from 2000 to 2007, but the residual values ε_t are equal 1171.9; 315.7; -553.3; -679.4; -695.3; -811.4; -206.1; 759.2; 698.7.

Using the criterion constructed in Section 2, we check the remainders (perturbations) of models (5) and (6) on simplicity (homogeneity). For the model (5) $N_{\min} = 1$, $N_{\max} = 7$. Consequently, $N = \max\{N_{\min}, N_{\max}\} = 7$. This means that the constructed criterion accepts the hypothesis of the simplicity of the residues in the model (5) for any level of significance $\alpha \leq 0.29$, and, consequently, the hypothesis about the homogeneity of the residues is accepted in the the period between the crises.

For the model (6) $N_{\min} = 1$, $N_{\max} = \min \emptyset = \infty$. Consequently, $N = \max\{N_{\min}, N_{\max}\} = \infty$. This means that the constructed criterion rejects the hypothesis of the simplicity of the residues in the model (6) at any level of significance α, therefore, the hypothesis of the homogeneity of the remnants during the crisis period is rejected. This is why the hypothesis of a structural shift in 2008 is accepted.

References

1. Andrews, D.W.K.: Test for parameter instability and structural change with unknown change point. Econometrica **61** (1993) 821-856
2. Ayvazyan, S.A.: Applied Statistics. Finances and Statistics Publishers, (1983) [in Russian]
3. Brown, R.L., Durbin, J., and Evans, J.M.: Techniques for testing the constancy of regression relationship over time. Journal of the Royal Statistical Society. Series B (Methodological) **37** (1975) 149-192
4. Chow, G.C.: Tests of the equality between two sets of coefficients in two linear regressions. Econometrica **28** (1960) 561-605
5. Feller, V.: Introduction to the Theory of Probability and its Applications, Vol. II. Mir (1966) [in Russian]
6. Maddala, G.S., and Kim I.M.: Unit Roots, Cointegration and Structural Change. Cambridge University Press (1998)
7. Magnus, Ya.R., Katyshev, P.K., and Peresetskii, A.A.: Econometrics. A Beginner's Course. Delo (2007) [in Russian]
8. Nikiforov, I.V. Sequential Detection of Abrupt Changes in Time Series Properties. Nauka (1983) [in Russian]
9. Suslov, V.I., Ibragimov, N.M., Talysheva, L.P., and Tsyplakov, A.A.: Econometrics. Siberian Branch of the Russian Academy of Sciences (2005) [in Russian]

Mathematical Methods in Economics and Finance – m^{2}ef
Vol. 11/12, No. 1, 2016/2017
ISSN print edition: 1971-6419 – ISSN online edition: 1971-3878
Web page: `http://www.unive.it/m2ef/` – E-mail: `m2ef@unive.it`