APPLICATION OF THE THEORY OF TIME SERIES FOR EVALUATION TECHNIQUES OF MEASURING AND DISTINGUISHING

Ing. Mansour ABAID, Doctoral Degree Programme (2)
Dept. of Telecommunications, FEEC, BUT
E-mail: Abaid57@hotmail.cz

Supervised by: Dr. Josef Zapletal

ABSTRACT

The article deals with one problem of evaluation technique of two time series obtained by measuring of two technical especially transmission channels using regression analysis. It shows that the value of regression coefficient near to one need not to “close” relation between them.

1 INTRODUCTION

In the introductory part of my article there I remember some notions from the theory of regression analysis and theory of time series.

The regression line will be understood the function

$$y = \beta_0 + \beta_1 x + e \ ,$$

where β_0, β_1 are regression parameters and e is a stochastic error. In the next part I remember the minimal squares method. Let us suppose that we evaluated $n > 2$ information measurements from which at least two are other. We inscribe them with

$$y_i = \beta_0 + \beta_1 x_i + e_i \ , \quad i = 1, 2, \ldots, n \quad .$$

Furthermore we denote estimations of β_0 and β_1 by b_0 and b_1. We receive them from the condition

$$S(\beta_0,\beta_1)=\sum_{i=1}^{n}\left[y_i-(\beta_0+\beta_1 x_i)\right]^2=\min$$

by the method of minimal squares. We calculate the partial derivatives according to β_0 and β_1 then we put them equal to zero. After the rearrangement of these two equations we obtain the normal equations for b_0 and b_1:

$$nb_0+b_1\sum_{i=1}^{n}x_i=\sum_{i=1}^{n}y_i$$

$$b_0\sum_{i=1}^{n}x_i+b_1\sum_{i=1}^{n}x_i^2=\sum_{i=1}^{n}x_i\,y_i$$

The solution of this normal equations is :

$$b_0=\frac{1}{n}\left(\sum_{i=1}^{n}y_i-b_1\sum_{i=1}^{n}x_i\right)$$

.

$$b_1=\frac{n\sum_{i=1}^{n}x_i\,y_i-\left(\sum_{i=1}^{n}x_i\right)\left(\sum_{i=1}^{n}y_i\right)}{n\sum_{i=1}^{n}x_i^2-\left(\sum_{i=1}^{n}x_i\right)^2}$$

The next notion which I apply in my accession to knowledge of interpretation of results of measurements is the notion of time series. Under the time series we understand a sequence of concrete comparable set of data, which is uniquely ordered with respect to time in the sense from bygone to presence. I understand their analysis optionally forecast the system of methods which are instrumental for description of dynamical aggregates.

In the concrete I study two time series representing information systems activity and I look for the dependence or independence of these systems in this paper. In the case of some dependence I continue in special studies of special close relation between these two systems. I can prove and predict from this “close” relation between two systems using the comparison of data from previous season and data of the first system in presence for a prediction for the second system. For the proof of close relation I need some results from the theory of correlation of time series, especially the notion of the coefficient of correlation. For the coefficient of correlation the following formula holds true:

$$r_{xy} = r_{yx} = \frac{n\sum_{i=1}^{n} x_i\, y_i - \sum_{t=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{\left[n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2\right]\left[n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2\right]}}$$

2 CORRELATION OF DATA OF TWO TIME SERIES

This article contains the measured data of two information systems taking the number of data transmissions. I show a very good dependence of both systems – the coefficient of correlation will be near to one but the next studies of proximity (close relation) of both the systems will be negative. The measured data and the following differences obtained as the difference of measured date and the value of line obtained by the least squares method.

x_i	y_i	Δx_i	Δy_i
108	332	-1.9273	4.4909
121	372	1.0121	14.5152
129	358	-1.0485	-29.4606
142	420	1.8909	2.5636
153	450	2.8303	2.5879
158	479	-2.2303	1.6121
172	508	1.7091	0.6364
179	540	-1.3515	2.6606
188	575	-2.4121	7.6848
202	590	1.5273	-7.2909

At first sight we see a relatively good relation between x_i and y_i. That proves also the value of correlation coefficient of according time series which is $r_{x\,y} = 0.9889842$.

As the next step we enumerate stochastic components. This operation chokes the trend component down. We perform the estimation of trend as the linear trend. The appropriate result is expressed by the following equations:

linear approximation for the first system $z_x = 10.06061\, x + 99.86617$

and

linear approximation for the second system $z_y = 29.97579\, y + 297.53373$.

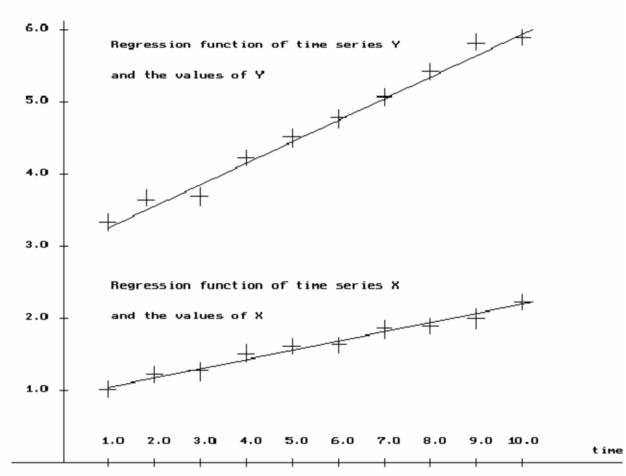

Fig. 1: *Regression function*

As the next step we calculate differences Δx_i and Δy_i . Their values are in the last two columns of the table. We calculate the coefficient of correlation and we have

$$\Delta r_{xy} = 0.0638198$$

We see that the value of this coefficient is very small – near to zero. This stands that the intensity of dependence between both systems is very small. The situation is described by the picture 1. The values of time series are distributed all round their regression lines in opposite sense. The high value of the coefficient of regression of prime two time series was very deceptive.

We obtain the same results using the parabolic approximation for our time series. In the concrete the coefficient of correlation for parabolic differences is: $\Delta^2 r_{xy} = 0.1112513$ which is also very small and we obtain that the intensity of dependence between both systems is small again.

3 CONCLUSIONS

It is very important the selection of the type of trend function describing the evolution of analyzed time series by inquiry into tightness of distance of these time series. Authenticity of the calculated coefficient of correlation of differences is qualified by the correct choice of trend function from which the differences are calculated. We see form the picture 1, that the choice of the trend function was correct. Applicability of the choice of trend function can be verified by Durbin – Watson test of autocorrelation.

REFERENCES

[1] Brownn, R.G.: Smoothing, Forecasting and Prediction of Discrete Time Series, Prentice Hall International, London 1963.

[2] Wei, W. W.: Time Series Analysis, John Wiley & Sons, New York 1990.

[3] Zapletal, J.: Základy počtu pravděpodobnosti a matematické statistiky, PCDIR Brno 1995, ISBN 80-214-0711-5.

[4] Zapletal, J.: Úvod do analýzy časových řad, PCDIR Brno 2000, ISBN 80-214-1719-6