INVESTIGATION OF PROBABILISTIC DEPENDENCY IN COMPLEX INFORMATION SYSTEMS BASED ON SCATTER PLOTS

Dmitry Bychkov

Saint-Petersburg State University of Aerospace Instrumentation
Saint-Petersburg, Russia

I. INTRODUCTION

The analysis of the linkages inherent in the studied processes or phenomena is the major task of research in various fields of knowledge. Structural complexity, the ambiguity of behavior, cyclic processes in biology and medicine, economics and technology determine the features of information being processed and impose certain requirements for methods of analysis. To date, the most common methods are correlation and regression analysis. This paper investigates the method of analysis of dependency based on the scatter plot [1]. In particular, the objective is to estimate a scatter plot for random processes of different probabilistic structure, namely: Normally distributed white noise and bimodal beta-distribution. A specific scatter plot is also considered for exponentially correlated sampling. Heart rate variability in the values of the durations of RR intervals is analyzed to illustrate the applied use of scatter plots.

II. SCATTER PLOT

A scatter plot is a graphical representation of the relationship between variables. One can conclude about the presence or absence of dependency between these values, and the nature of this dependency. Fig. 1 represents scatter plots for normally distributed white noise (a) and white noise having beta-distribution (b) with parameters $\alpha = \beta = 0{,}3$.

It is shown through the nature of the scatter plot that a change of one of the values does not result in changes of the second variable. There are no obvious signs of the grouping of data under study in the scatter plots.

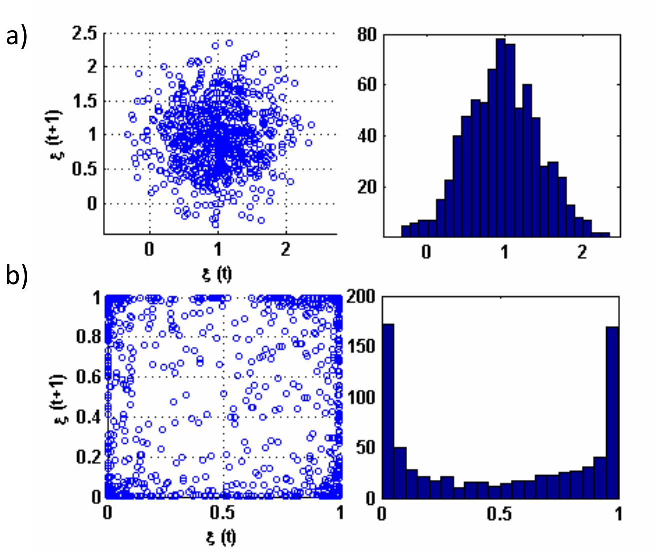

Fig. 1. Scatter plots and white noise histograms of different kind of distribution.

III. MODELING DEPENDENCIES

In order to understand the behavior of the scattering pattern when there is a dependency, we will simulate a sample with a given correlation function. Since the correlation function of a random process is bijectively related to its spectrum, we can take samples of white noise and let them pass through a filter with a given frequency response. A sample with an exponential correlation function is drawn below:

$$R(\tau) = \sigma^2 \alpha^{|\tau|}, \quad |\alpha| < 1.$$

Filter impulse response is as follows:

$$h(\tau) = \sigma\sqrt{1-\alpha^2} \cdot \alpha^{\tau}, \quad \tau \geq 0.$$

This impulse response corresponds to a recursive first order filter with α a coefficient of a feedback equal to [2]. The output of the shaping filter is presented below, whose input is fed with normally distributed white noise (Figures 2 to 4).

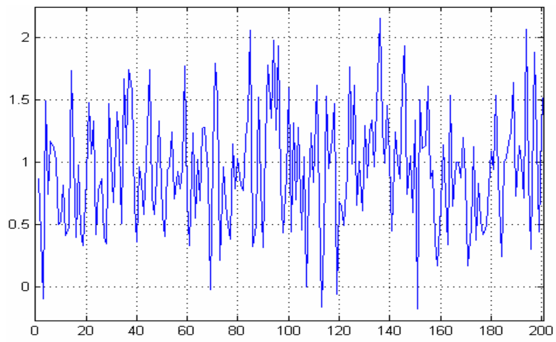

Fig. 2. White noise

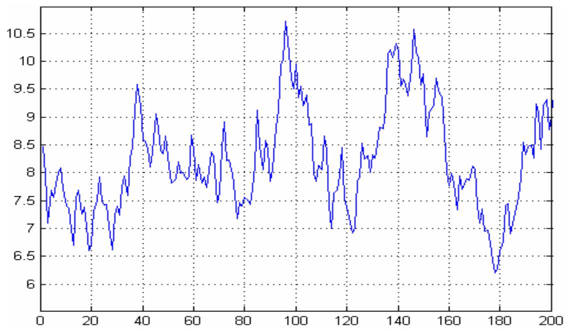

Fig. 3. Correlated Gaussian process

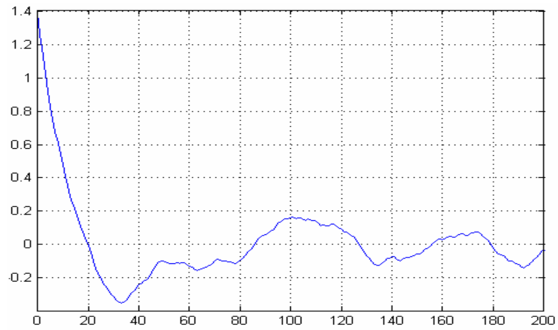

Fig. 4. Estimation of the correlation function of the formed process

A similar filtering was performed for the realizations of white noise with beta-distribution. The results of compiling scatter plots for processes obtained are shown in Fig. 5.

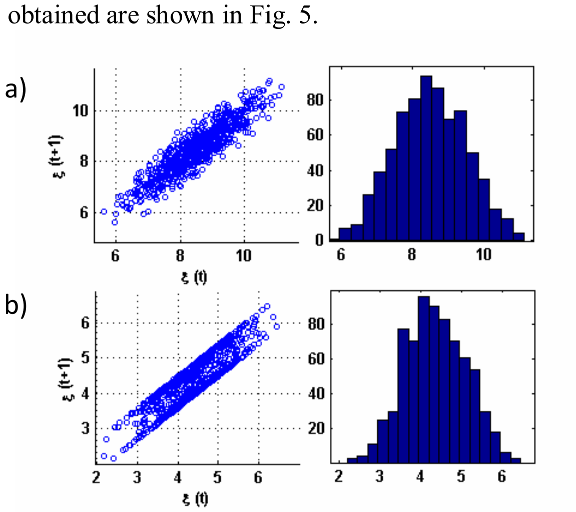

Fig. 5. Scatter plots and histograms for correlated processes

Figure 5 clearly observes direct linear dependency. An increase in one value increases the other. The degree of correlation in this case depends on the characteristics of the shaping filter. Obviously, normal (Fig. 5, a) distribution retains its properties under linear transformations, which is not true for beta-distribution, whose form varies greatly after filtration (Fig. 5, b). The effect of normalization of non-Gaussian distributions under linear transformations is quite clearly shown here [3].

It is known that the correlation function characterizes only linear dependencies. Let us consider the nature of the scattering diagram for a functionally transformed random process. Let us place the values of normally distributed white noise on one axis of the scatter plot and a square of this process on the other (Fig. 6). The resulting scatter plot clearly reflects the nonlinear dependency of the transformed process.

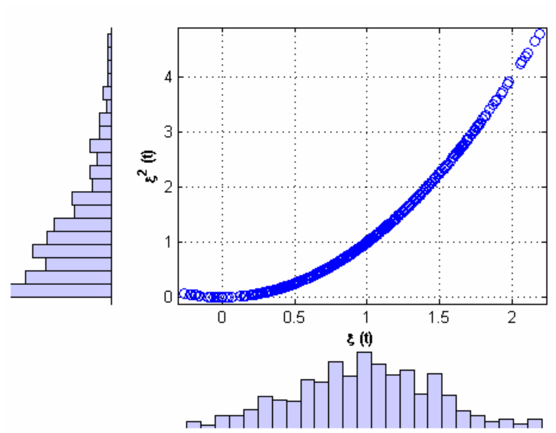

Fig. 6. A nonlinear dependency on the scatter plot

IV. PRACTICAL APPLICATION: ANALYSIS OF HEART RATE VARIABILITY

The values of time intervals between each contraction of the left ventricle are used in medicine to study a patient's heart rate variability (prong R corresponds to a contraction of the left ventricle in the ECG). These intervals in the scatter plot present arrhythmic events of various kinds. Normally, the scatter plot is supposed to show a tilting oval “cloud”. Restraint “clouds” may be indicative of the predominance of the sympathetic part of the autonomic nervous system and, by contrast, a considerable scatter of points on the scatter plot indicates the predominance of the ANS. This method of assessing heart rate variability is particularly useful for cases where rare and sudden cardiac arrhythmias occur against the monotony of the rhythm. The severe irregularity of a “cloud” with uneven seals or detached groups of points can also refer to the rhythm disturbances. Figures 7 – 9 shows different presentations of data obtained using electrocardiography (ECG) from a patient who shows arrhythmic events.

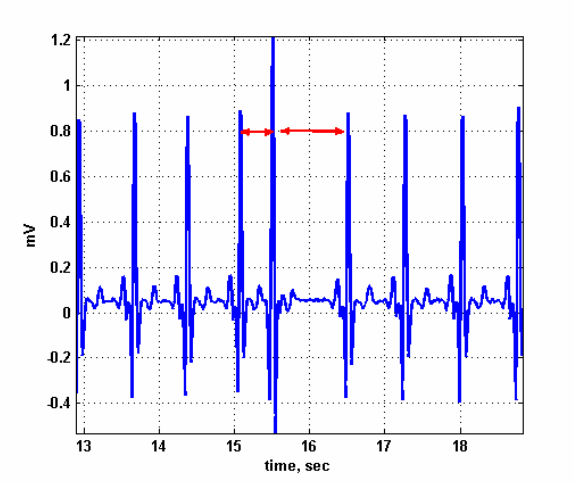

Fig. 7. ECG

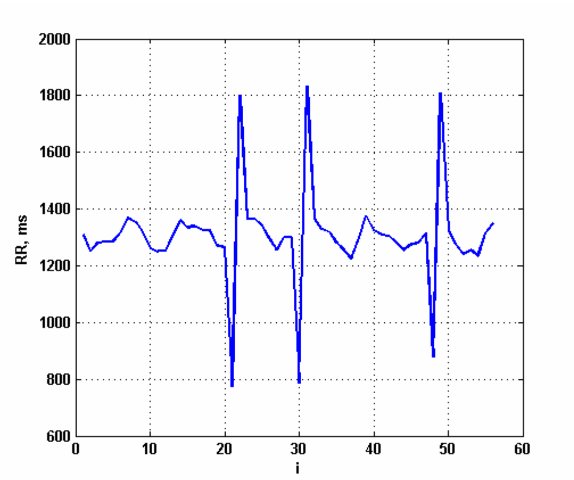

Fig. 8. RR – intervals duration values

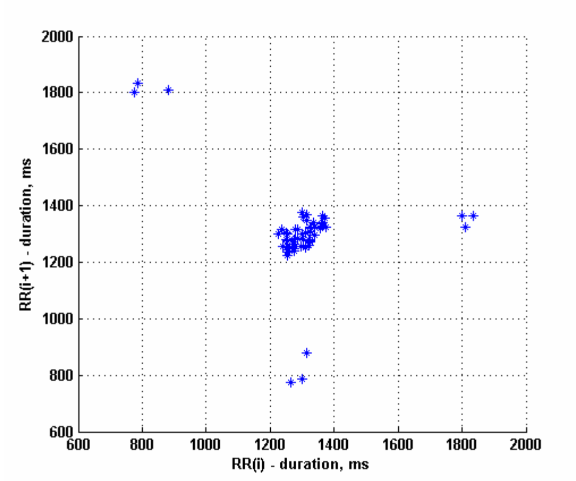

Fig. 9. RR-intervals scatter plot

V. CONCLUSIONS

By analyzing scatter plots, one can solve several problems associated with the monitoring and assessment of the system, detection of abnormal emissions, and the problem of forecasting the behavior of the system. Easy to chart and minimal use of computing resources make scatter plots a good alternative to classical methods of correlation analysis, especially when solving problems in which data representation is not informative in the time domain.

REFERENCES

[1] *Хименко В.И.* Модели и методы обработки экспериментальных данных: Курс лекций. СПб.: ГУАП, 2009.

[2] *Сергиенко А.Б.* Цифровая обработка сигналов. СПб.: Питер, 2002.

[3] *Тихонов В.И.* Статистическая радиотехника. М.: Радио и связь, 1982.

[4] *Р.М. Рангайян.* Анализ биомедицинских сигналов. М.:Физматлит, 2010.