STATISTICAL FORECASTING: OPTIMALITY AND ROBUSTNESS

Yu. S. Kharin

Research Institute for Applied Problems of Mathematics and Informatics,
Belarusian State University
Minsk, Belarus
E-mail: kharin@bsu.by

This presentation summarizes the results of the author on development of the theory of robust statistical forecasting.

Key words: robustness, forecasting, distortions, time series, risk.

Webster encyclopedic dictionary defines forecasting as "an activity aimed at computing or predicting some future events or conditions based on rational analysis of relevant data". Forecasting is widely used in applications, including decision-making and planning at top executive level.

Rigorous mathematical studies of stochastic forecasting were started in the 1930s by the founder of modern probability theory Andrey Kolmogorov. Two stages can be identified in the development of statistical forecasting techniques. The first stage took place before 1974 and was characterized by construction of forecasting statistics (algorithms or procedures) that minimized the forecast risk (the mean square error) for a number of simple time series models (such as stationary models with fixed spectral densities, stationary models with trends belonging to a given function family, autoregressive models, moving average models, etc.).

In the 1970s it was found that applying many of the developed "optimal" forecasting procedures to real-world data often resulted in much higher risks compared to the expected theoretical values. The reason for this phenomenon was pointed out by Peter Huber (Swiss Federal Institute of Technology) in his talk at the 1974 International Congress of Mathematicians (Vancouver, Canada): Statistical inferences are based only in part upon the observations. An equally important base is formed by prior assumptions about the underlying situation [1]. These assumptions form the hypothetical model M_0 of the process being investigated. In applications the behavior of the investigated processes often deviates from the model assumptions M_0, resulting in instability of forecasting statistics. The main types of deviations from the hypothetical model M_0 are the following: non-normal observation errors, dependence (or correlation) of observation errors, nonhomogeneous observation errors, model specification errors, presence of outliers, change points, or missing values in the time series [1–3]. It was suggested [1] that statisticians develop robust statistical procedures, which would have been affected only slightly by small deviations from the hypothetical model M_0. This marked the beginning of the second stage in the history of statistical forecasting.

In the recent years, the development of robust statistical algorithms has become one of the major research topics in mathematical statistics. New results in this field are presented each year at the International Conference on Robust Statistics (ICORS).

The mathematical substance of the time series forecasting problem is quite simple: to estimate the future value $x_{T+\tau} \in R^d$ of the d – variate time series in τ steps ahead based on T observations $X_T = (x_1, \ldots, x_T) \in R^{Td}$ and some hypothetical stochastic model M_0 of the time series $\{x_t\}$. In practice, the underlying hypothetical model M_0 is often distorted, and this fact leads to instability of the "optimal" forecasting statistics that are optimal under M_0 only [1, 2]. The lecture is devoted to sensitivity analysis of the risk for traditional forecasting statistics (based on M_0) and also to construction and analysis of new robust forecasting statistics (based on distorted model M_ε for some classes of distortions).

Let $\hat{x}_{T+\tau} = f(X_T) : R^{Td} \to R^d$, be any forecasting statistic,

$$r_\varepsilon = r_\varepsilon(f) = E_\varepsilon\left\{\left\|\hat{x}_{T+\tau} - x_{T+\tau}\right\|^2\right\} \geq 0$$

be the mean square forecast risk, $E_\varepsilon\{\cdot\}$ be the expectation symbol w.r.t. the probability distribution of the time series, $\varepsilon \in [0, \varepsilon_+]$ be the distortion level, e. g., the portion of outliers, (if $\varepsilon = 0$ we have the hypothetical model M_0). We use robustness characteristics [3–5] based on the risk functional: the guaranteed (upper) risk

$$r_+ = r_+(f) = \sup r_\varepsilon(f),$$

where the supremum is taken over all admissible distortions

$$\{M_\varepsilon : 0 \leq \varepsilon \leq \varepsilon_+\};$$

the hypothetical risk

$$r_0 = r_0(f_0) = \inf_f r_0(f) > 0;$$

the risk instability coefficient

$$\kappa = \kappa(f) = \left(r_+(f) - r_0\right)/r_0 \geq 0$$

equal to the relative increment of the guaranteed risk w.r.t. the hypothetical risk; δ - admissible distortion level

$$\varepsilon_* = \varepsilon_*(\delta) = \sup\{\varepsilon_+ \geq 0 : \kappa(f) \leq \delta\}$$

indicating the maximal distortion level such that the risk instability coefficient doesn't exceed a given positive constant δ; the risk-robust forecasting statistic

$$\hat{x}^*_{T+\tau} = f_*(X_T),$$

where $f_*(\cdot)$ minimizes the guaranteed risk:

$$\kappa(f_*) = \inf_f \kappa(f).$$

In this lecture we present our results [4–8] related to the following topical problems:

- construction of mathematical models and descriptions of typical distortions;
- evaluation of the risk-robustness for traditional forecasting procedures;
- evaluation of critical distortion levels;

- construction of new robust forecasting procedures.

The theoretical results are illustrated on simulated and real statistical data.

Bibliography

1. *Huber P.* Robust Statistics. N.Y. : Wiley, 1981.
2. *Hampel F., Ronchetti E., Rousseuw P., Stahe W.* Robust Statistics. N.Y. : Wiley, 1986.
3. *Maronna R., Martin R., Yohai V.* Robust Statistics. N.Y. : Wiley, 2006.
4. *Kharin Yu.* Robustness in Statistical Forecasting. N.Y. : Springer, 2013.
5. *Kharin Yu.* Robustness in Statistical Pattern recognition. Dordrecht : Kluwer, 1996.
6. *Kharin Yu.* Optimality and robustness in statistical forecasting // Int. Encyclopedia of Statistical science. N.Y. : Springer, 2011. P. 1034–1037.
7. *Kharin Yu.* Robustness of mean square risk in forecasting of regression time series // Commun. Stat. Theory. 2011. V. 40. P. 2893–2906.
8. *Kharin Yu., Voloshko V.* Robust estimation of AR coefficients under simultaneously influencing outliers and missing values // JSPI. 2011. V. 141. P. 3276–3288.