ROBUST BAYESIAN MULTIVARIATE FORECASTING UNDER DISTORTIONS OF PRIOR DENSITIES IN THE CHI-SQUARE METRIC

P.A.Shlyk, A.Yu.Kharin
Belarusian State University
Minsk, BELARUS
e-mail: `paul.shlyk@gmail.com`, `KharinAY@bsu.by`

Abstract

This paper investigates robustness of the multivariate Bayesian forecasting model under the chi-square metric distortions of priors. The explicit form for the guaranteed upper risk is obtained and an integral equation for the robust prediction statistics is given.

1 Introduction

Bayesian framework has been of considerable interest in the context of predicting future observations using a parametric model on the basis of previous ones. Due to incorporating *a priori* knowledge about the object under observation the Bayesian methods allow improving prediction quality, especially in case of a small size sample. This has implications in such spheres of computer data analysis as medical sciences, financial markets, bio-informatics. As in practice priors can be defined improperly, the robustness analysis of the model is required in order to make appropriate inferences under distortions of hypothetical assumptions. A detailed review of Bayesian robustness subject can be found in [1]. In this paper we explore minimax robustness of the multivariate Bayesian forecasting model under functional distortions of priors, defined using chi-squared metric. Similar results for the univariate model can be found in [2].

2 Forecasting Model under Functional Distortions

Suppose that the random vector of observations $x = (x_t)_{t=1}^{T} \in X \subseteq \mathbb{R}^{n\times T}$ stochastically depends on θ with the hypothetical conditional probability density function (p.d.f) $p^0(x|\theta)$, where $\theta \in \Theta \subseteq \mathbb{R}^m$ is the unobserved random vector of model parameters with the hypothetical p.d.f. $\pi^0(\theta)$. The problem is to forecast the random vector $y \in Y \subseteq \mathbb{R}^n$ that stochastically depends on x and θ with the hypothetical conditional p.d.f. $g^0(y|x,\theta)$. We explore model robustness in case of functional distortions, defined using χ^2-metric (pseudometric):

$$\rho_{\chi^2}(h_1,h_2) = \int\limits_U \frac{(h_1(u)-h_2(u))^2}{h_1(u)}\,du,$$

where p.d.f.s $h_1(u), h_2(u)$ are defined on U. Suppose that the parameters vector θ is distributed according to an unknown p.d.f. $\pi^\varepsilon(\theta) \in \Pi$, where Π is a set of admissible p.d.f.s of θ:

$$\pi^\varepsilon(\theta) \in \Pi = \{\Pi_\varepsilon : 0 \le \varepsilon \le \varepsilon_+\}, \Pi_\varepsilon = \left\{\pi^\varepsilon(\cdot) : \rho_{\chi^2}(\pi^0(\cdot), \pi^\varepsilon(\cdot)) = \varepsilon^2\right\}. \tag{1}$$

The performance of a prediction statistics (p.s.) $f(\cdot) : X \to Y$ is characterized by the risk functional:

$$r(f(\cdot); s^\varepsilon(\cdot)) = \int_X \int_Y \rho^2(f(x), y) s^\varepsilon(x, y)\, dy dx, \tag{2}$$

where $\rho(\cdot, \cdot)$ is the Euclidean distance function in $\mathbb{R}^n$ and $s^\varepsilon(\cdot, \cdot)$ is the join p.d.f. of x and y:

$$s^\varepsilon(x, y) = \int_\Theta g^0(y|x, \theta) p^0(x|\theta) \pi^\varepsilon(\theta) d\theta.$$

The guaranteed upper risk functional $r_*(\cdot)$ is used to analyze the robustness of a p.s. $f(\cdot)$:

$$r_*(f(\cdot)) = \sup_{s^\varepsilon(\cdot) \in S} r(f(\cdot), s^\varepsilon(\cdot)), \tag{3}$$

where S is a set of admissible p.d.f.s $s^\varepsilon(\cdot)$. First, we aim to find the explicit expression for the guaranteed upper risk functional under distortions (1). Our second objective is to find the robust p.s. $f_*(\cdot)$:

$$r_*(f_*(\cdot)) = \inf_{f(\cdot)} r_*(f(\cdot)).$$

3 The Guaranteed Upper Risk Functional

As a Borelean function $\pi^\varepsilon(\cdot)$ from Π should be a p.d.f., the following ratios are valid:

$$\pi^\varepsilon(\theta) \ge 0, \theta \in \Theta, \int_\Theta \pi^\varepsilon(\theta)\, d\theta = 1.$$

Denote the mathematical expectation and variance calculated for the hypothetical model as $E_0\{\cdot\}$, $D_0\{\cdot\}$ respectively. The risk functional (2) can be represented in the following form:

$$r(f(\cdot); \pi^\varepsilon(\cdot)) = \int_\Theta \pi^\varepsilon(\theta) r_1(f(\cdot); \theta)\, d\theta, \tag{4}$$

where $r_1(f(\cdot); \theta)$ is the conditional risk functional for the fixed parameters vector θ:

$$r_1(f(\cdot); \theta) = \int_X \int_Y \rho^2(f(x), y) s^0(x, y|\theta)\, dy dx;\ \ s^0(x, y|\theta) = g^0(y|x, \theta) p^0(x|\theta).$$

Then the guaranteed upper risk functional (3) can be represented as

$$r_*(f(\cdot)) = \sup_{\pi^\varepsilon(\cdot) \in \Pi} r(f(\cdot); \pi^\varepsilon(\cdot)).$$

Denote the conditional risk bias as $\overset{\circ}{r}(f(\cdot);\theta)$:

$$\overset{\circ}{r}(f(\cdot);\theta) = r_1(f(\cdot);\theta) - E_0\{r_1(f(\cdot);\theta)\}.$$

Introduce the critical value of the distortions level:

$$\varepsilon^*(f(\cdot)) = \frac{\sqrt{D_0\{r_1(f(\cdot);\theta)\}}}{\sup\limits_{\theta\in\Theta} |\overset{\circ}{r}(f(\cdot);\theta)|}. \tag{5}$$

Theorem 1. *Let the hypothetical forecasting model be distorted according to (1) and for any p.s.* $f(\cdot): X \to Y$ *the distortion level* $\varepsilon_+ \in [0, \varepsilon^*(f(\cdot))]$. *Then the guaranteed upper risk functional (3) can be represented as*

$$r_*(f(\cdot)) = r(f(\cdot);\pi^*(\cdot)), \tag{6}$$

where the extreme p.d.f. $\pi^*(\cdot)$ *is defined as*

$$\pi^*(\theta) = \pi^0(\theta)\left(1 + \varepsilon_+ \frac{\overset{\circ}{r}(f(\cdot);\theta)}{\sqrt{D_0\{r_1(f(\cdot);\theta)\}}}\right). \tag{7}$$

Proof. The problem of the guaranteed upper risk finding under the theorem conditions is equivalent to the following variational calculus problem:

$$\begin{cases} \int\limits_\Theta \pi^\varepsilon(\theta) r_1(f(\cdot);\theta)\, d\theta \to \max\limits_{\pi^\varepsilon(\cdot)\in\Pi_\varepsilon, 0\le\varepsilon\le\varepsilon_+} \\ \int\limits_\Theta \pi^\varepsilon(\theta)\, d\theta = 1, \\ \int\limits_\Theta \frac{(\pi^\varepsilon(\theta)-\pi^0(\theta))^2}{\pi^0(\theta)} = \varepsilon^2. \\ \pi^\varepsilon(\theta) \ge 0, \theta \in \Theta. \end{cases} \tag{8}$$

Solving the problem (8), we obtain the extreme distorted p.d.f. $\pi^*(\cdot)$ (7). The restriction (5) on the distortion level ε_+ follows from the last restriction of (8). •

Corollary 1. *Under the theorem (1) conditions the guaranteed upper risk (3) can be represented as*

$$r_*(f(\cdot)) = r_0(f(\cdot)) + \varepsilon_+\sqrt{D_0\{r_1(f(\cdot);\theta)\}}, \tag{9}$$

where $r_0(f(\cdot))$ *is the hypothetical risk functional:*

$$r_0(f(\cdot)) = \int\limits_\Theta r_1(f(\cdot);\theta)\pi^0(\theta)\, d\theta. \tag{10}$$

Proof. Taking into account (6), (4) (7), (10) we obtain (9). •

4 The Robust Prediction Statistics

Denote

$$F_\varepsilon(f(\cdot); x, y, \theta) = s^0(x, y|\theta) + \frac{\varepsilon}{\sqrt{D_0\{r_1(f(\cdot);\theta)\}}}\left(s^0(x, y|\theta) - s^0(x, y)\right) \overset{\circ}{r}(f(\cdot); \theta),$$
$$x \in X, y \in Y, \theta \in \Theta;\ \varepsilon^{**} = \inf_{f(\cdot)} \varepsilon^*(f(\cdot)). \tag{11}$$

Theorem 2. *Let the hypothetical forecasting model be distorted according to (1) and the distortion level* $\varepsilon_+ \in [0, \varepsilon^{**}]$*. Then the robust p.s.* $f_*(\cdot)$ *satisfies the following integral equation:*

$$f_*(x) = \frac{\iint\limits_{Y\Theta} y \cdot \pi^0(\theta) F_{\varepsilon_+}(f_*(\cdot); x, y, \theta)\, d\theta dy}{\iint\limits_{Y\Theta} \pi^0(\theta) F_{\varepsilon_+}(f_*(\cdot); x, y, \theta)\, d\theta dy}. \tag{12}$$

Proof. The result is obtained by applying the method of variations to the following optimization problem:

$$r_*(f(\cdot)) \to \max_{f_i(\cdot), i=\overline{1,n}}, f(\cdot),$$

where

$$f_i(\cdot) : X \to Y_i,\ i = \overline{1,n};\ Y = Y_1 \times Y_2 \times ... \times Y_n,$$
$$f(x) = (f_1(x), f_2(x), ..., f_n(x)),\ x \in X. \quad \bullet$$

5 Conclusion

The explicit expression (9) of the guaranteed upper risk allows calculating its deviation from the hypothetical risk for any p.s. $f(\cdot)$, and the order of this deviation is $\mathcal{O}(\varepsilon_+)$.

The integral equation (12) allows building iterative procedures for calculating the robust p.s.$f_*(\cdot)$:

$$f_{(0)} := f_0(x), f_{(i)}(x) = \frac{\iint\limits_{Y\Theta} y \cdot \pi^0(\theta) F_{\varepsilon_+}(f_{(i-1)}(\cdot); x, y, \theta)\, d\theta dy}{\iint\limits_{Y\Theta} \pi^0(\theta) F_{\varepsilon_+}(f_{(i-1)}(\cdot); x, y, \theta)\, d\theta dy}, x \in X; i \in \mathbb{N}.$$

The research is partially supported by the grant No F06M-072 of the Belarusian National Science Foundation.

References

[1] Berger J.O. (1985). *Statistical Decision Theory and Bayesian Analysis*. 2nd ed. Springer-Verlag, New York.

[2] Kharin A. (2002). Minimax Robustness of Bayesian Forecasting under Functional Distortions of Probability Densities. *Austrian Journal of Statistics*. Vol. **31**, No. 2&3, pp. 177-188.