Multivariate Bayesian Forecasting under Functional Distortions in the χ^2-metric

Alexey Kharin   Pavel Shlyk

Dept. of Probability Theory & Mathematic Statistics
Faculty of Applied Mathematics & Computer Science
Belarusian State University
Minsk, Belarus

The 8th Tartu Conference on Multivariate Statistics, 2007

Outline

Introduction

The main spheres of Bayesian framework application:

- medical sciences;
- financial markets;
- bio-informatics.

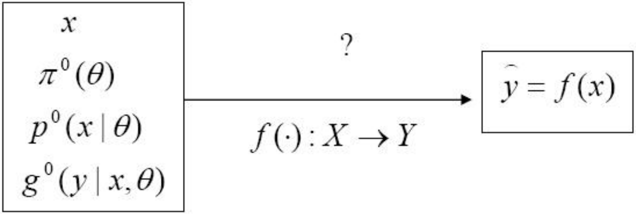
Investigations on robustness of Bayesian forecasting:

- *Bayesian Robustness* / ed. by G.O.Berger et al., 1995;
- *Problemy odporności w byesowskiej analizie statystycznej* / Marek Męczarski, 1998;
- *Minimax Robustness of Bayesian Forecasting under Functional Distortions of Probability Densities* / Alexey Kharin, 2002.

Hypothetical Model of Bayesian Forecasting

Let on a probability space $(\Omega, \mathcal{F}, P)$ be defined three random elements:

- the unobserved vector of model parameters $\theta \in \Theta \subseteq \mathbb{R}^m, \pi^0(\theta)$;
- the vector of observations $x = (x_t)_{t=1}^T \in X \subseteq \mathbb{R}^{n \times T}, p^0(x|\theta)$;
- the value to be forecasted $y \in Y \subseteq \mathbb{R}^n, g^0(y|x, \theta)$.



Hypothetical Risk Functional

The performance of a prediction statistics $f(\cdot) : X \to Y$ can be measured by hypothetical risk functional $r^0(f(\cdot))$:

$$r^0(f(\cdot)) = \mathrm{E}_0\left\{\rho^2(f(x),y)\right\} = \int\limits_X\int\limits_Y s^0(x,y)\rho^2(f(x),y)\, dydx, \quad (1)$$

$$s^0(x,y) = \int\limits_\Theta g^0(y|x,\theta)p^0(x|\theta)\pi^0(\theta)d\theta. \quad (2)$$

$\rho(\cdot,\cdot)$ is the Euclidean distance function in $\mathbb{R}^n$.

Bayesian Prediction Statistics

Hypothetical Bayesian Prediction Density:

$$q^0(y|x) = \int_{\Theta} g^0(y|x,\theta)\pi^0(\theta|x)\,dx. \tag{3}$$

$$\pi^0(\theta|x) = \frac{p^0(x|\theta)\pi^0(\theta)}{p^0(x)}, p^0(x) = \int_{\Theta} p^0(x|\theta)\pi^0(\theta)\,d\theta. \tag{4}$$

Bayesian Prediction Statistics:

$$\hat{y} = f^0(x) = \mathrm{E}_0\{y|x\} = \int_{Y} y \cdot q^0(y|x)\,dy. \tag{5}$$

Bayesian Prediction Statistics is optimal with respect to the hypothetical risk functional.

Distortions of the Hypothetical Model

We consider functional distortions of priors, defined using χ^2-metric:

$$\rho_{\chi^2}(h_1, h_2) = \int\limits_U \frac{(h_1(u) - h_2(u))^2}{h_1(u)}\, du, \tag{6}$$

where p.d.f.s $h_1(u), h_2(u)$ are defined on U.
Suppose that θ is distributed according to an unknown p.d.f. $\pi^\varepsilon(\cdot) \in \Pi$:

$$\pi^0(\theta) \rightsquigarrow \pi^\varepsilon(\theta) \in \Pi = \{\Pi_\varepsilon : 0 \le \varepsilon \le \varepsilon_+\}, \tag{7}$$

$$\Pi_\varepsilon = \left\{\pi^\varepsilon(\cdot) : \rho_{\chi^2}(\pi^0(\cdot), \pi^\varepsilon(\cdot)) = \varepsilon_+^2\right\}.$$

Risk Functionals as Robustness Characteristics

We measure the performance of a prediction statistics $f(\cdot) : X \to Y$ by the risk functional $r(\cdot,\cdot)$:

$$r\left(f(\cdot),\tilde{\pi}(\cdot)\right) = \mathrm{E}\left\{\rho^2\left(f(x),y\right)\right\} = \int\limits_X\int\limits_Y \tilde{s}(x,y)\rho^2\left(f(x),y\right)\,dxdy. \quad (8)$$

$$\tilde{s}(x,y) = \int\limits_\Theta g^0(y|x,\theta)p^0(x|\theta)\tilde{\pi}(\theta)\,d\theta.$$

We use the guaranteed upper risk functional $r_*(\cdot)$ to measure the robustness of $f(\cdot)$:

$$r_*\left(f(\cdot)\right) = \sup_{\tilde{\pi}(\cdot)\in\Pi} r\left(f(\cdot),\tilde{\pi}(\cdot)\right). \quad (9)$$

Robust Prediction Statistics

We aim to find the robust bayesian prediction statistics $f_*(\cdot)$:

$$r_* \left(f_*(\cdot)\right) = \inf_{f(\cdot)} r_* \left(f(\cdot)\right). \tag{10}$$

The Conditional Risk Functional

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

$$\pi^{\varepsilon}(\theta) \geq 0, \theta \in \Theta, \int\limits_{\Theta} \pi^{\varepsilon}(\theta)\, d\theta = 1.$$

The risk functional can be represented as:

$$r(f(\cdot); \pi^{\varepsilon}(\cdot)) = \int\limits_{\Theta} \pi^{\varepsilon}(\theta) r_1(f(\cdot); \theta)\, d\theta, \tag{11}$$

$$r_1(f(\cdot); \theta) = \int\limits_{X}\int\limits_{Y} \rho^2(f(x), y) s^0(x, y|\theta)\, dy dx; \tag{12}$$

$$s^0(x, y|\theta) = g^0(y|x, \theta) p^0(x|\theta). \tag{13}$$

Critical Distortions Level

The guaranteed upper risk functional can be represented as

$$r_*(f(\cdot)) = \sup_{\pi^\varepsilon(\cdot)\in\Pi} r(f(\cdot);\pi^\varepsilon(\cdot)). \tag{14}$$

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

$$\overset{\circ}{r}(f(\cdot);\theta) = r_1(f(\cdot);\theta) - E_0\{r_1(f(\cdot);\theta)\}. \tag{15}$$

Introduce the critical distortions level:

$$\varepsilon^*(f(\cdot)) = \frac{\sqrt{D_0\{r_1(f(\cdot);\theta)\}}}{\sup_{\theta\in\Theta}|\overset{\circ}{r}(f(\cdot);\theta)|}. \tag{16}$$

Extreme Probability Density Function

Theorem

Let the hypothetical forecasting model be distorted according to (7) 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 can be represented as

$$r_*(f(\cdot)) = r(f(\cdot); \pi^*(\cdot)), \tag{17}$$

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

$$\pi^*(\theta) = \pi^0(\theta) \left(1 + \varepsilon_+ \frac{\mathring{r}\,(f(\cdot); \theta)}{\sqrt{D_0\{r_1(f(\cdot); \theta)\}}} \right). \tag{18}$$

The Guaranteed Upper Risk Functional

Corollary

Under the theorem conditions the guaranteed upper risk can be represented as

$$r_*(f(\cdot)) = r_0(f(\cdot)) + \varepsilon_+ \sqrt{D_0\{r_1(f(\cdot);\theta)\}}, \tag{19}$$

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

$$r_0(f(\cdot)) = \int_\Theta r_1(f(\cdot);\theta)\pi^0(\theta)\,d\theta. \tag{20}$$

The Robust Prediction Statistics

Denote for $x \in X, y \in Y, \theta \in \Theta$:

$$F_{\varepsilon}(f(\cdot); x, y, \theta) = s^0(x, y|\theta) + \frac{\varepsilon \left(s^0(x, y|\theta) - s^0(x, y)\right) \overset{\circ}{r} (f(\cdot); \theta)}{\sqrt{D_0\{r_1(f(\cdot); \theta)\}}}. \quad (21)$$

$$\varepsilon^{**} = \inf_{f(\cdot)} \varepsilon^*(f(\cdot)). \quad (22)$$

Theorem

Let the hypothetical forecasting model be distorted according to (7) and the distortion level $\varepsilon_+ \in [0, \varepsilon^{**}]$*. Then the robust p.s.* $f_*(\cdot)$ *satisfies the following integral equation:*

$$f_*(x) = \frac{\iint\limits_{Y\Theta} y \cdot \pi^0(\theta) F_{\varepsilon_+}(f_*(\cdot); x, y, \theta)\, d\theta dy}{\iint\limits_{Y\Theta} \pi^0(\theta) F_{\varepsilon_+}(f_*(\cdot); x, y, \theta)\, d\theta dy}. \quad (23)$$

Conclusion

- The explicit expression (19) of the guaranteed upper risk allows calculating its deviation from the hypothetical risk for any p.s. $f(\cdot)$ and this deviation is at most of order $\mathcal{O}(\varepsilon_+)$.
- The integral equation (23) allows building iterative procedures for calculating the robust p.s. $f_*(\cdot)$:

$$f_{(0)} := f_0(x),$$

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

$$x \in X; i \in \mathbb{N}.$$

References I

- A.Kharin, P.Shlyk *On Robustness of Multivariate Bayesian Forecasting under Functional Distortions of Priors*, BSU Bulletin, Minsk, vol.2, pp.103-107, 2006 (in russian)
- A.Kharin, P.Shlyk *On Robustness of Multivariate Bayesian Forecasting*, Proceedings for RobHD2004, Vorau, Austria, 2004.
- A.Kharin *Minimax Robustness of Bayesian Forecasting under Functional Distortions of Probability Densities,* Austrian Journal of Statistics, vol.31(2&3), pp.177-188, 2002.

Acknowledgement I

The authors would like to thank the Organizing Committee for the invitation and for the financial support making participation in the Conference possible.