MINIMAX INEQUALITY FOR THE PREDICTION RISK

L. GAJEK[1] AND V. LIPIŃSKA[2]

[1]Institute of Mathematics, Technical University of Łódź, Poland, and Polish Financial Supervision Authority, Warszawa, Poland

[2]Institute of Mathematics, Technical University of Łódź, Poland

Abstract. This is a continuation of the article by L. Gajek and V. Lipińska, *Sharp inequality for the Bayes prediction risk*, Inequality Theory and Applications, **5**, (2007), 97-106. But, instead of the Bayes prediction problem, the minimax one is considered. The lower bound for the minimax prediction risk is derived by using a kind of information inequality.

Assume that $\mathbf{X}$ is an observable random vector (or variable) with the conditional probability density function $f(\cdot|\theta)$, indexed by a parameter $\theta \in \Theta$, relative to some σ-finite measure μ and $\Theta \subseteq \mathbb{R}$ is an open interval $(\underline{\theta}, \overline{\theta})$. Let Y be a random variable. We want to predict its value using a function $\delta(\mathbf{X})$. This function will be called **a predictor of** Y. Assume also that $(\mathbf{X}, Y)$ is a random vector with the joint (conditional) distribution $f(\mathbf{x}, y|\theta)$, indexed by a parameter $\theta \in \Theta$.

Let $R(\delta, \theta) = \mathrm{E}_\theta \left\{ [Y - \delta(\mathbf{X})]^2 m(\theta) \right\}$ be the risk function of the predictor δ with a positive weight function $m(\theta)$, where the expectation is taken over both $\mathbf{X}$ and Y. We want to obtain a predictor δ^* for which $\sup_{\theta\in\Theta} R(\delta^*, \theta) = \inf_\delta \sup_{\theta\in\Theta} R(\delta, \theta)$. In that case the predictor δ^* will be called **minimax**.

Gajek ((1987), An improper Cramer–Rao lower bound, *Applicationes Mathematicae* 19, 241−256) and Gajek ((1988), On the minimax value in the scale model with truncated data, *The Annals of Statistics* 16, 669 − 677) obtained the lower bound for the minimax estimation risk. We have obtained, using Cramér–Rao inequality, similar lower bound for the minimax prediction risk

$$\limsup_{\theta\to\theta^*} R_1(\delta, \theta) \geq \lim_{\theta\to\theta^*} \frac{\phi^2(\theta) m(\theta)}{\left(\frac{\phi(\theta)}{\phi_1(\theta)}\right)^2 I_{\mathbf{X}}(\theta) + 1},$$

where

$$R_1(\delta, \theta) = \mathrm{E}_\theta[\delta(\mathbf{X}) - B(\theta) - g(\mathbf{X}, \theta)]^2 m(\theta) + B^2(\theta) m(\theta)$$

and $R(\delta, \theta) = \mathrm{E}_\theta[Y - g(\mathbf{X}, \theta)]^2 m(\theta) + R_1(\delta, \theta)$, $g(\mathbf{X}, \theta) = \mathrm{E}_\theta(Y|\mathbf{X})$, $\phi_1(\theta) = \mathrm{E}_\theta \left[\frac{\partial}{\partial\theta} g(\mathbf{X}, \theta)\right]$, $\phi \in \int \phi_1(\theta) d\theta$ and $I_{\mathbf{X}}(\theta)$ be the so called Fisher information in $\mathbf{X}$ about the parameter θ, defined by $I_{\mathbf{X}}(\theta) = \mathrm{E}_\theta \left[\frac{\partial}{\partial\theta} \log f(\mathbf{X}|\theta)\right]^2$.

Moreover, we will show the way how the lower bound can be used in proving minimaxity of the predictors. An admissible bound for the class of regular predictors is also included. It can be easily adapted for proving admissibility. Illustrative examples of applications will be presented. This talk contains the first author's private opinions which should not be attributed to the PFSA.