Mathematical
Inequalities
& Applications
Volume 13, Number 2 (2010), 437–448

MINIMAX INEQUALITY FOR THE PREDICTION RISK

L. GAJEK AND V. LIPIŃSKA

(*Communicated by J. Pečarić*)

Abstract. This is a continuation of the article "Sharp inequality for the Bayes prediction risk" (2007). But, instead of the Bayes prediction problem, the minimax one is considered in this paper. The lower bound for the minimax prediction risk is derived by using the information inequality. Moreover it is shown that a predictor $\delta(X)$ for which the lower bound is attained is a minimax one. Sufficient conditions for the admissibility of predictors are also given.

1. Introduction

Assume that $\mathbf{X}$ is an observable random variable (or vector) 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 the function $\delta(\mathbf{X})$. This function will be called *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 $\theta \in \Theta$.

It is worth mentioning that the model we are dealing with in this paper admits also another interpretation: Y may be an unobservable quantity depending statistically on $\mathbf{X}$ which can be observed. Thus, knowing $\mathbf{X}$, we try to evaluate Y by a function $\delta(\mathbf{X})$. In this interpretation a crucial assumption is that $\mathbf{X}$ and Y are somehow statistically dependent though they may have completely different distributions.

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

$$\sup_{\theta \in \Theta} R(\delta^*, \theta) = \inf_{\delta} \sup_{\theta \in \Theta} R(\delta, \theta). \tag{1}$$

In that case the predictor δ^* will be called *minimax*.

To prove a predictor is minimax one can construct a sequence of priors that asymptotically minimizes the corresponding Bayes risk. In this paper we develop an alternative approach. Instead of looking for such a sequence of priors, which usually is a

Mathematics subject classification (2010): 62F15, 62M20.
Keywords and phrases: Prediction; minimax risk function; information inequality.
This article contains the first author's private opinions which should not be attributed to the PFSA.

Paper MIA-13-33

problem difficult in itself, we provide a lower limit of the minimax prediction risk, which is easy to compute (see Trybuła (2000) and (2003)).

Gajek (1987) and (1988) obtained the lower bound for the minimax estimation risk. In Section 3 we have obtained, using Cramér–Rao inequality, similar lower bound for the minimax prediction risk. In Section 4 we will show that the lower bound can be useful in proving minimaxity of the predictors (Theorem 4.1). Theorem 4.2 gives an admissible bound for the class of regular predictors (in the sense of the assumption **A1**) and can be easily adapted for proving admissibility (Theorem 4.3). Some examples of showing the minimaxity and admissibility of predictors are given.

2. Assumptions

Let $\mathrm{E}_\theta Y^2 < \infty$, $g(\mathbf{X},\theta) = \mathrm{E}_\theta[Y|\mathbf{X}]$ be the conditional expectation of Y given $\mathbf{X}$ (and given θ), $b(\theta) = \mathrm{E}_\theta[\delta(\mathbf{X}) - Y]$ be the bias of δ, $\phi'(\theta) = \mathrm{E}_\theta\,[(\partial/\partial\theta)g(\mathbf{X},\theta)]$ and ϕ be a real function defined on Θ such that $\frac{d}{d\theta}\phi(\theta) = \phi'(\theta)$. Let $I(\theta)$ be the Fisher information that $\mathbf{X}$ contains about parameter θ defined by $I(\theta) = \mathrm{E}_\theta\,[(\partial/\partial\theta)\log f(\mathbf{X}|\theta)]^2 < \infty$ and $V(\theta) = 1/I(\theta)$.

Let us make the following assumption (the same assumption is made in [10]):

A1 For all $\theta \in \Theta$ such that $R(\theta,\delta) < \infty$, the function $b(\theta)$ is differentiable at θ and

$$b'(\theta) + \phi'(\theta) = \mathrm{E}_\theta\left\{[\delta(\mathbf{X}) - b(\theta) - g(\mathbf{X},\theta)]\,\frac{\partial}{\partial\theta}\log f(\mathbf{X}|\theta)\right\}.$$

It is easy to show that

$$\mathrm{E}_\theta[Y - \delta(\mathbf{X})]^2 = \mathrm{E}_\theta[Y - g(\mathbf{X},\theta)]^2 + \mathrm{E}_\theta[\delta(\mathbf{X}) - b(\theta) - g(\mathbf{X},\theta)]^2 + b^2(\theta). \quad (2)$$

For simplicity let us use the notation

$$R_1(\delta,\theta) = \mathrm{E}_\theta[\delta(\mathbf{X}) - b(\theta) - g(\mathbf{X},\theta)]^2 m(\theta) + b^2(\theta)m(\theta). \quad (3)$$

Using (2) and (3) we can rewrite the risk function as follows

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

After using **A1**, (2) and Cauchy–Schwarz inequality we have

$$\begin{aligned}\mathrm{E}_\theta[Y - \delta(\mathbf{X})]^2 &\geqslant \mathrm{E}_\theta[Y - g(\mathbf{X},\theta)]^2 \\ &\quad + \frac{1}{I(\theta)}\mathrm{E}^2_\theta\left\{[\delta(\mathbf{X}) - b(\theta) - g(\mathbf{X},\theta)]\frac{\partial}{\partial\theta}\log f(\mathbf{X}|\theta)\right\} + b^2(\theta) \\ &= \mathrm{E}_\theta[Y - g(\mathbf{X},\theta)]^2 + V(\theta)[b'(\theta) + \phi'(\theta)]^2 + b^2(\theta), \quad (4)\end{aligned}$$

where the right side of (4) depends on the predictor δ through the bias $b(\theta)$. In the next section we present a lower bound for the right side of (4) not depending on the predictor δ.

3. Lower bounds

First of all we present two lemmas (see [4]) which will be useful to prove the main results.

Let a, b and y be real functions defined on an open interval $J \subset \mathbb{R}$, such that $b(x) > a(x) > 0$ for all $x \in J$, and y is differentiable on J. We will call $x^* \in \mathbb{R} \cup \{-\infty, \infty\}$ a point adherent to J if $x^* \notin J$ and there is a sequence $\{x_n\} \subseteq J$ such that $x_n \to x^*$.

Consider the following differential inequality:

$$\frac{a(x)}{b(x)} \geqslant \frac{y^2(x)}{x^2} + a(x)\frac{[1+y'(x)]^2}{b(x)-a(x)}. \tag{5}$$

LEMMA 3.1. *Suppose that the real function $y \in C^1(J)$ satisfies (5) on the interval of reals J and x^* is a point adherent to J. If $x^* \in \{-\infty, 0, \infty\}$ and there exists $\lim_{x\to x^*} a(x)/b(x)$ then there exists a sequence $\{x_n\}$ in J such that $x_n \to x^*$ and*

$$\lim_{n\to\infty} y'(x_n) = \lim_{n\to\infty} \frac{y(x_n)}{x_n} = -\lim_{n\to\infty} \frac{a(x_n)}{b(x_n)}, \tag{6}$$

$$\lim_{n\to\infty} \frac{a(x_n)}{b(x_n)} = \lim_{n\to\infty} \left\{ \frac{y^2(x_n)}{x_n^2} + \frac{a(x_n)}{b(x_n)-a(x_n)} \left[1+y'(x_n)\right]^2 \right\}. \tag{7}$$

If, additionally, there exist $\lim\limits_{x\to x^*} b(x) = \infty$ *and* $\lim\limits_{x\to x^*} y(x) = 0$ *then*

$$\lim_{n\to\infty} a(x_n) = \lim_{n\to\infty} \left\{ \frac{y^2(x_n)}{x_n^2} b(x_n) + a(x_n)[1+y'(x_n)]^2 \right\}. \tag{8}$$

LEMMA 3.2. *Suppose that the real function $y \in C^1(J)$ satisfies (5) on the interval $J = (\underline{x}, \overline{x})$. If $J = (-\infty, 0)$ or $J = (0, +\infty)$ and there exist $\lim_{x\to\underline{x}} a(x)/b(x) = \lim_{x\to\overline{x}} a(x)/b(x)$, then*

$$\frac{a(x)}{b(x)} = \frac{y^2(x)}{x^2} + \frac{a(x)}{b(x)-a(x)}[1+y'(x)]^2 \quad \textit{for all} \quad x \in J. \tag{9}$$

THEOREM 3.1. *Suppose that the parameter space Θ contains a subset Θ^* on which ϕ is a diffeomorphism, the assumption* **A1** *holds and θ^* is a point adherent to Θ^* such that $\lim\limits_{\theta\to\theta^*} \phi(\theta)$ exists and is equal to 0 or $\pm\infty$. If there exists a function $r(\theta)$ such that $0 < r(\theta) < \phi^2(\theta)m(\theta)$ on Θ^* and the limits $\lim\limits_{\theta\to\theta^*} \phi^2(\theta)m(\theta)$ and $\lim\limits_{\theta\to\theta^*} r(\theta) < \infty$ exist, and*

$$1 < \lim_{\theta\to\theta^*} \left[\frac{\phi^2(\theta)m(\theta)}{r(\theta)} - 1\right] \left[\frac{\phi'(\theta)}{\phi(\theta)}\right]^2 I^{-1}(\theta) < \infty, \tag{10}$$

then there exists no regular (in the sense of assumption **A1***) predictor* $\delta(\boldsymbol{X})$ *such that*

$$\limsup_{\theta\to\theta^*} R_1(\delta,\theta) < \lim_{\theta\to\theta^*} r(\theta).$$

Proof. Let us notice that if $\lim_{\theta\to\theta^*} r(\theta)=0$, the assertion holds.
Suppose now that $\lim_{\theta\to\theta^*} r(\theta)\neq 0$ and there exists $\delta(\mathbf{X})$ such that

$$\limsup_{\theta\to\theta^*} R_1(\delta,\theta) < \lim_{\theta\to\theta^*} r(\theta).$$

Then there exists $\varepsilon_0>0$ and a set $\Theta_0\subseteq\Theta^*$ such that

$$R_1(\delta,\theta)\leqslant r(\theta)-\varepsilon_0 \tag{11}$$

for every $\theta\in\Theta_0$, and θ^* is adherent to Θ_0. Moreover, after taking into account (4), the following inequality

$$R_1(\delta,\theta)\geqslant b^2(\theta)m(\theta)+V(\theta)[b'(\theta)+\phi'(\theta)]^2m(\theta) \tag{12}$$

holds on the set Θ_0. From (10) it follows that there exists a set Θ_1 such that θ^* is adherent to Θ_1 and

$$\left[\frac{\phi^2(\theta)m(\theta)}{r(\theta)}-1\right]\left[\frac{\phi'(\theta)}{\phi(\theta)}\right]^2 I^{-1}(\theta)\geqslant 1 \tag{13}$$

for every $\theta\in\Theta_1$. By (11)–(13) for every $\theta\in\Theta_0\cap\Theta_1$ the following inequality holds:

$$\begin{aligned} r(\theta)-\varepsilon_0 &\geqslant b^2(\theta)m(\theta)+[b'(\theta)+\phi'(\theta)]^2V(\theta)m(\theta)\\ &= b^2(\theta)m(\theta)+\left[\frac{db}{d\phi}\phi'(\theta)+\phi'(\theta)\right]^2 m(\theta)\frac{1}{I(\theta)}\\ &\geqslant b^2(\theta)m(\theta)+\frac{r(\theta)}{\phi^2(\theta)m(\theta)-r(\theta)}\left[\frac{db}{d\phi}+1\right]^2\phi^2(\theta)m(\theta). \end{aligned} \tag{14}$$

Dividing (14) by $\phi^2(\theta)m(\theta)$, we obtain

$$\frac{r(\theta)-\varepsilon_0}{\phi^2(\theta)m(\theta)}\geqslant\frac{b^2(\theta)}{\phi^2(\theta)}+\frac{r(\theta)}{\phi^2(\theta)m(\theta)-r(\theta)}\left[\frac{db}{d\phi}+1\right]^2. \tag{15}$$

Thus

$$\frac{r(\theta)}{\phi^2(\theta)m(\theta)}\geqslant\frac{b^2(\theta)}{\phi^2(\theta)}+\frac{r(\theta)}{\phi^2(\theta)m(\theta)-r(\theta)}\left[\frac{db}{d\phi}+1\right]^2 \tag{16}$$

for every $\theta\in\Theta_0\cap\Theta_1$. Since ϕ is diffeomorphism on $\Theta^*\supseteq\Theta_0\cap\Theta_1$ and the limits $\lim_{\theta\to\theta^*}\phi^2(\theta)m(\theta)$ and $\lim_{\theta\to\theta^*}r(\theta)<\infty$ exist, by Lemma 3.1 (7) there exists a sequence $\{\theta_n\}$ such that $\theta_n\in\Theta_0\cap\Theta_1$ for all n, $\theta_n\to\theta^*$ $(n\to\infty)$ and the equality in (16) holds in the sense of the limit of the sequence $\{\theta_n\}$. So we can write

$$\lim_{\theta\to\theta^*}\frac{r(\theta)}{\phi^2(\theta)m(\theta)}=\lim_{\theta\to\theta^*}\left\{\frac{b^2(\theta)}{\phi^2(\theta)}+\frac{r(\theta)}{\phi^2(\theta)m(\theta)-r(\theta)}\left[\frac{db}{d\phi}+1\right]^2\right\}.$$

Using this limit in (15) we have

$$\lim_{\theta\to\theta^*}\frac{r(\theta)}{\phi^2(\theta)m(\theta)}-\lim_{\theta\to\theta^*}\frac{\varepsilon_0}{\phi^2(\theta)m(\theta)}\geqslant\lim_{\theta\to\theta^*}\left\{\frac{b^2(\theta)}{\phi^2(\theta)}+\frac{r(\theta)}{\phi^2(\theta)m(\theta)-r(\theta)}\left[\frac{db}{d\phi}+1\right]^2\right\}.$$

Observe that this contradicts (15) in the case $\lim_{\theta\to\theta^*}\phi^2(\theta)m(\theta)<\infty$.

Assume now that $\lim_{\theta\to\theta^*}\phi^2(\theta)m(\theta)=\infty$.

From (14) we have

$$r(\theta)-\varepsilon_0\geqslant b^2(\theta)m(\theta)+\frac{r(\theta)}{1-\frac{r(\theta)}{\phi^2(\theta)m(\theta)}}\left(\frac{db}{d\phi}+1\right)^2. \tag{17}$$

Because $0<\lim_{\theta\to\theta^*}r(\theta)<\infty$ we obtain $\lim_{\theta\to\theta^*}\frac{r(\theta)}{\phi^2(\theta)m(\theta)}=0$. After taking into account (16) and using Lemma 3.1 (6) we obtain

$$\lim_{\theta\to\theta^*}\frac{db}{d\phi}(\theta)=\lim_{\theta\to\theta^*}\frac{b(\theta)}{\phi(\theta)}=-\lim_{\theta\to\theta^*}\frac{r(\theta)}{\phi^2(\theta)m(\theta)}.$$

Consequently $\lim_{\theta\to\theta^*}\frac{db}{d\phi}(\theta)=0$. From (17) we have

$$\lim_{\theta\to\theta^*}r(\theta)-\varepsilon_0\geqslant\lim_{\theta\to\theta^*}b^2(\theta)m(\theta)+\lim_{\theta\to\theta^*}\frac{r(\theta)}{1-\frac{r(\theta)}{\phi^2(\theta)m(\theta)}}\left(\frac{db}{d\phi}+1\right)^2$$

and finally $-\varepsilon_0\geqslant\lim_{\theta\to\theta^*}b^2(\theta)m(\theta)$. It obviously leads to a contradiction.

The best lower bound for the upper limit in Theorem 3.1 is given in the following

THEOREM 3.2. *Suppose that the parameter space* Θ *contains a subset* Θ^* *on which* ϕ *is a diffeomorphism, the assumption* **A1** *holds and* θ^* *is a point adherent to* Θ^* *such that* $\lim_{\theta\to\theta^*}\phi(\theta)$ *exists and is equal to* 0 *or* $\pm\infty$*. If there exists* $\lim_{\theta\to\theta^*}\phi^2(\theta)m(\theta)$ *and*

$$\lim_{\theta\to\theta^*}\frac{\phi^2(\theta)m(\theta)}{\left(\frac{\phi(\theta)}{\phi'(\theta)}\right)^2 I(\theta)+1}<\infty$$

then for every predictor $\delta(\boldsymbol{X})$ *the following inequality holds:*

$$\limsup_{\theta\to\theta^*}R_1(\delta,\theta)\geqslant\lim_{\theta\to\theta^*}\frac{\phi^2(\theta)m(\theta)}{\left(\frac{\phi(\theta)}{\phi'(\theta)}\right)^2 I(\theta)+1}. \tag{18}$$

Proof. Consider the function

$$r_k(\theta)=\frac{\phi^2(\theta)m(\theta)}{k^{-1}\left(\frac{\phi(\theta)}{\phi'(\theta)}\right)^2 I(\theta)+1}\qquad\text{for any}\quad k\quad(0<k<1).$$

Obviously $k^{-1}\left[\phi(\theta)/\phi'(\theta)\right]^2 I(\theta)+1>1$ and consequently $0<r_k(\theta)<\phi^2(\theta)m(\theta)$.

It is easy to show that $\lim_{\theta\to\theta^*} r_k(\theta)<\infty$. Moreover, we can check that

$$\lim_{\theta\to\theta^*}\left\{\left[\frac{\phi^2(\theta)m(\theta)}{r_k(\theta)}-1\right]\left(\frac{\phi'(\theta)}{\phi(\theta)}\right)^2 I^{-1}(\theta)\right\}=k^{-1}\in(0,\infty).$$

We have shown the assumptions of Theorem 3.1 are satisfied and therefore

$$\limsup_{\theta\to\theta^*} R_1(\delta,\theta)\geqslant\lim_{\theta\to\theta^*} r_k(\theta) \tag{19}$$

for every regular predictor $\delta(\mathbf{X})$ and for every k $(0<k<1)$. Since

$$r_k(\theta)=\frac{k\phi^2(\theta)m(\theta)}{\left(\frac{\phi(\theta)}{\phi'(\theta)}\right)^2 I(\theta)+k}\geqslant\frac{k\phi^2(\theta)m(\theta)}{\left(\frac{\phi(\theta)}{\phi'(\theta)}\right)^2 I(\theta)+1}, \tag{20}$$

from (19) and (20) it follows that

$$\limsup_{\theta\to\theta^*} R_1(\delta,\theta)\geqslant\lim_{\theta\to\theta^*}\frac{\phi^2(\theta)m(\theta)}{\left(\frac{\phi(\theta)}{\phi'(\theta)}\right)^2 I(\theta)+1}.$$

This completes the proof.

4. Applications

4.1. Minimaxity

The inequality (18) can be useful in proving minimaxity. The following result is an immediate consequence of Theorem 3.2 and the definition of a minimax predictor.

THEOREM 4.1. *Suppose that the assumptions of Theorem* 3.2 *are satisfied and* $\delta^*(\boldsymbol{X})$ *is a predictor such that*

$$\sup_{\theta\in\Theta} R(\delta^*,\theta)=\liminf_{\theta\to\theta^*}\left[\mathrm{E}_\theta[Y-g(\boldsymbol{X},\theta)]^2 m(\theta)+\frac{\phi^2(\theta)m(\theta)}{\left(\frac{\phi(\theta)}{\phi'(\theta)}\right)^2 I(\theta)+1}\right]. \tag{21}$$

Then δ^* *is minimax.*

Proof. Let us notice that from (21) and Theorem 3.2 we have

$$\begin{aligned}\inf_\delta\sup_{\theta\in\Theta} R(\delta,\theta) &\geqslant \inf_\delta\limsup_{\theta\to\theta^*} R(\delta,\theta)\\ &\geqslant \liminf_{\theta\to\theta^*}\left[\mathrm{E}_\theta[Y-g(\mathbf{X},\theta)]^2 m(\theta)+\frac{\phi^2(\theta)m(\theta)}{\left(\frac{\phi(\theta)}{\phi'(\theta)}\right)^2 I(\theta)+1}\right]\\ &= \sup_{\theta\in\Theta} R(\delta^*,\theta). \end{aligned} \tag{22}$$

From the definition of supremum

$$\inf_{\delta}\sup_{\theta\in\Theta} R(\delta,\theta) \leqslant \sup_{\theta\in\Theta} R(\delta^*,\theta)$$

and consequently

$$\inf_{\delta}\sup_{\theta\in\Theta} R(\delta,\theta) = \sup_{\theta\in\Theta} R(\delta^*,\theta).$$

Obviously the right side of (21) is independent of δ. Moreover it is easy to check if equality in (21) is achieved. Consequently Theorem 4.1 may be very useful in proving minimaxity of a predictor δ.

If we reduce the class of predictors to the unbiased ones, we get from (4) an analogous inequality with $\phi^2 m/[(\phi/\phi')^2 I]$ instead of $\phi^2 m/[(\phi/\phi')^2 I+1]$. The quantities differ by exactly $+1$ in the denominators, providing bounds on the minimax prediction risk within the class of unbiased or all (possibly biased) predictors, respectively. What is interesting, the difference does not depend on the class of distributions but on the class of predictors only.

EXAMPLE 4.1. Let X and Y be random variables with the joint (conditional) density function given as follows:

$$f(x,y|\theta) = \frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[(x-\theta)^2 - 2\rho(x-\theta)(y-\theta)+(y-\theta)^2\right]\right\}$$

for $\rho\in(-1,1)$. Then the marginal and conditional density functions (given θ) are defined, respectively, as follows $f(x|\theta)=\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{(x-\theta)^2}{2}\right\}$, $f(y|\theta)=\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{(y-\theta)^2}{2}\right\}$ and $f(y|x,\theta) = \frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}}\exp\left\{-\frac{[y-\theta-\rho(x-\theta)]^2}{2(1-\rho^2)}\right\}$.

It is easy to see that $g(X,\theta)=\theta+\rho(X-\theta)$. Moreover $\mathrm{E}_\theta[Y-g(X,\theta)]^2 = 1-\rho^2$ and $\phi'(\theta)=1-\rho$. Thus $\phi(\theta)=(1-\rho)\theta+c$, for some $c\in\mathbb{R}$. Obviously $I(\theta)=1$.

Let $m(\theta)=1$. Then $\lim_{\theta\to\pm\infty}\frac{\phi^2(\theta)m(\theta)}{\left[\left(\frac{\phi(\theta)}{\phi'(\theta)}\right)^2 I(\theta)+1\right]} = (1-\rho)^2$, $\lim_{\theta\to\pm\infty}\phi^2(\theta)m(\theta) = +\infty$ and $\lim_{\theta\to\pm\infty}\phi(\theta)=\pm\infty$. Finally, the limit on the right side of (21) is equal to $2(1-\rho)$ (for $\theta^*\in\{-\infty,+\infty\}$).

Assume now that $\delta(X)=aX$. Then

$$R(\delta,\theta) = \mathrm{E}_\theta[Y-\delta(X)]^2 = a^2 - 2a\rho + 1 + \theta^2(a^2-2a+1).$$

We can see at once that for $a=1$

$$\sup_{\Theta} R(\delta,\theta) = 2(1-\rho).$$

Consequently (21) is satisfied. Thus $\delta(X)=X$ is a minimax predictor of Y.

EXAMPLE 4.2. Let $\mathbf{X}=(X_1,\ldots,X_n)$ be a random vector with the density functions $f(x_i|\theta) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x_i-\theta)^2}{2\sigma^2}\right\}$ for every $i\in\{1,\ldots,n\}$, $x_i\in\mathbb{R}$ and $\theta>0$ and let Y be a random variable with the same density function as X_i. Assume that the random variables X_i and Y are conditionally independent (given θ). Let $m(\theta)=1$. In this example $g(\mathbf{X},\theta)=\theta$. Moreover $\mathrm{E}_\theta[Y-g(\mathbf{X},\theta)]^2m(\theta)=\sigma^2$, $\phi'(\theta)=1$ and $\phi(\theta)=\theta+c$ for some $c\in\mathbb{R}$. In this case $I(\theta)=\frac{n}{\sigma^2}$, so $\lim_{\theta\to+\infty}\frac{\phi^2(\theta)}{\left(\frac{\phi(\theta)}{\phi'(\theta)}\right)^2 I(\theta)+1}=\frac{\sigma^2}{n}$, $\lim_{\theta\to+\infty}\phi^2(\theta)m(\theta)=+\infty$ and $\lim_{\theta\to+\infty}\phi(\theta)=+\infty$. Thus the right side of (21) is equal to $\frac{\sigma^2}{n}+\sigma^2$.

Let $\delta(\mathbf{X})=\alpha\sum_{i=1}^n X_i+\beta$. Then

$$R(\delta,\theta)=\theta^2(n\alpha-1)^2+\theta(2n\beta\alpha-2\beta)+\beta^2+\alpha^2n\sigma^2+\sigma^2.$$

Let $\alpha=\frac{1}{n}$ and $\beta=0$. Then

$$\sup_\Theta R(\delta,\theta)=\frac{\sigma^2}{n}+\sigma^2.$$

Finally we conclude that $\delta^*(\mathbf{X})=\frac{1}{n}\sum_{i=1}^n X_i$ is a minimax predictor of Y.

EXAMPLE 4.3. Let $\mathbf{X}=(X_1,\ldots,X_n)$ be a random vector with the density functions $f(x_i|\theta)=\theta\exp(-x_i\theta)$ for every $i\in\{1,\ldots,n\}$, for all $x_i\geqslant 0$ and $\theta>0$ and let Y be a random variable with the same density function as X_i. Assume that the random variables X_i and Y are conditionally independent (given θ). Let $m(\theta)=\theta^2$. It is easy to show that $g(\mathbf{X},\theta)=1/\theta$ and $I(\theta)=n/\theta^2$. Moreover $\mathrm{E}_\theta[Y-g(\mathbf{X},\theta)]^2m(\theta)=1$, $\phi'(\theta)=-1/\theta^2$ and $\phi(\theta)=1/\theta+c$ for some $c\in\mathbb{R}$. In this case $\lim_{\theta\to0^+}\phi^2(\theta)m(\theta)=1$, $\lim_{\theta\to0^+}\phi(\theta)=+\infty$. Consequently $\lim_{\theta\to0^+}\frac{\phi^2(\theta)m(\theta)}{\left[\left(\frac{\phi(\theta)}{\phi'(\theta)}\right)^2 I(\theta)+1\right]}=\frac{1}{1+n}$. We have shown the assumptions of Theorem 4.1 are satisfied.

Thus the right side of (21) is equal to $1+\frac{1}{n+1}$.

Let $\delta(\mathbf{X})=a\sum_{i=1}^n X_i+b$. Then

$$R(\delta,\theta)=\theta^2b^2-2\theta b(1-na)+(n+n^2)a^2-2an+2.$$

Let $a=\frac{1}{n+1}$ and $b=0$. Thus

$$\sup_\Theta R(\delta,\theta)=1+\frac{1}{1+n}.$$

So $\delta(\mathbf{X})=\dfrac{1}{n+1}\sum_{i=1}^n X_i$ is a minimax predictor of Y.

4.2. Admissibility

THEOREM 4.2. *Suppose that* ϕ *is a diffeomorphism on* $\Theta = (\underline{\theta}, \overline{\theta})$, *assumption* **A1** *holds on* Θ, *and* $\phi(\Theta) = (-\infty, 0)$ *or* $\phi(\Theta) = (0, +\infty)$. *Assume that the limits below exist and the following equality holds*

$$\lim_{\theta \to \underline{\theta}} \left[\left(\frac{\phi(\theta)}{\phi'(\theta)} \right)^2 I(\theta) + 1 \right]^{-1} = \lim_{\theta \to \overline{\theta}} \left[\left(\frac{\phi(\theta)}{\phi'(\theta)} \right)^2 I(\theta) + 1 \right]^{-1}.$$

Then there exists no regular (in the sense of the assumption **A1***) predictor* $\delta(\boldsymbol{X})$ *such that*

$$R_1(\delta, \theta) \leqslant \frac{\phi^2(\theta) m(\theta)}{\left(\frac{\phi(\theta)}{\phi'(\theta)} \right)^2 I(\theta) + 1} \tag{23}$$

for every $\theta \in \Theta$ *with the sharp inequality holding for some* $\theta \in \Theta$.

Proof. Suppose that there exists predictor $\delta(\mathbf{X})$ such that (23) holds for every $\theta \in \Theta$ and the sharp inequality holds for some $\theta_1 \in \Theta$. Hence, by the (4), we have

$$\frac{\phi^2(\theta) m(\theta)}{\left(\frac{\phi(\theta)}{\phi'(\theta)} \right)^2 I(\theta) + 1} \geqslant R_1(\delta, \theta) \geqslant b^2(\theta) m(\theta) + V(\theta) [b'(\theta) + \phi'(\theta)]^2 m(\theta).$$

Thus, dividing both sides of this inequality by $\phi^2(\theta) m(\theta)$, we obtain

$$\frac{1}{\left(\frac{\phi(\theta)}{\phi'(\theta)} \right)^2 I(\theta) + 1} \geqslant \frac{b^2(\theta)}{\phi^2(\theta)} + V(\theta) \frac{[b'(\theta) + \phi'(\theta)]^2}{\phi^2(\theta)}. \tag{24}$$

Since ϕ is diffeomorphism on Θ, from (24) we obtain

$$\frac{a(\phi)}{b(\phi)} \geqslant \frac{b^2(\phi)}{\phi^2} + a(\phi) \frac{[1 + \frac{db}{d\phi}]^2}{b(\phi) - a(\phi)}, \tag{25}$$

where $a(\phi) \equiv 1$ and $b(\phi(\theta)) = \left(\frac{\phi(\theta)}{\phi'(\theta)} \right)^2 I(\theta) + 1$. By the assumptions there exist, and are equal to each other, both limits of $a(\phi)/b(\phi)$ at the endpoints of the interval $\phi(\Theta)$. Thus, by Lemma 3.2, the equality holds in (25) for every $\phi \in \phi(\Theta)$ but this is a contradiction to (23) for $\phi_1 = \phi(\theta_1)$.

The following result is the consequence of Theorem 4.1 and the definition of admissible predictors.

THEOREM 4.3. *Suppose that the assumptions of Theorem* 3.2 *are fulfilled and* $\delta(\boldsymbol{X})$ *is a predictor such that*

$$R(\delta, \theta) = \mathrm{E}_\theta [Y - g(\boldsymbol{X}, \theta)]^2 m(\theta) + \frac{\phi^2(\theta) m(\theta)}{\left(\frac{\phi(\theta)}{\phi'(\theta)} \right)^2 I(\theta) + 1}. \tag{26}$$

Then δ *is admissible.*

EXAMPLE 4.4. Let us take into account the model described in Example 4.3. Let $c = 0$. Then $\frac{\phi^2(\theta)m(\theta)}{\left[\left(\frac{\phi(\theta)}{\phi'(\theta)}\right)^2 I(\theta)+1\right]} = \frac{1}{1+n}$. Let $\delta(\mathbf{X}) = \frac{1}{n+1}\sum_{i=1}^n X_i$. For this predictor all the assumptions of Theorem 4.3 are satisfied. Finally, the predictor $\delta(\mathbf{X}) = \dfrac{1}{n+1}\sum\limits_{i=1}^n X_i$ is an admissible one.

EXAMPLE 4.5. Let us take into account the model described in Example 4.3 with the parameter space $\Theta = (0, \mu_0^{-1})$, where $\mu_0 > 0$. Obviously for the predictor $\delta(\mathbf{X}) = a\sum_{i=1}^n X_i + b$ we have $R(\delta,\theta) = \theta^2 b^2 - 2\theta b(1-na) + (n+n^2)a^2 - 2an + 2$. Moreover, for this predictor the right side of (21) is equal to $1 + \frac{1}{n+1}$. First of all we try to obtain such coefficients a and b for which

$$\sup_{\Theta} R(\delta,\theta) = 1 + \frac{1}{1+n}.$$

In order for the risk of the prediction to have supremum equal to $1 + \frac{1}{1+n}$, the following two conditions should be satisfied:

$$\begin{cases} R(\delta,\theta = 0) = a^2 n(1+n) - 2na + 2 \leqslant 1 + \frac{1}{n+1}, \\ R(\delta,\theta = \mu_0^{-1}) = b^2\mu_0^{-2} - 2(1-na)\mu_0^{-1} + a^2 n(1+n) - 2na + 2 \leqslant 1 + \frac{1}{n+1}. \end{cases}$$

The first condition is satisfied only when $a = \frac{1}{n+1}$. For this value of a the second condition is satisfied for $b \in \left[0, \frac{2}{n+1}\mu_0\right]$. For these values of a and b there are a lot of corresponding minimax predictors of Y.

Finally, every element of the class

$$\Xi = \left\{ \frac{1}{n+1}\sum_{i=1}^n X_i + b; b \in \left[0; \frac{2}{n+1}\mu_0\right] \right\}$$

is a minimax predictor of Y within the class of all regular predictors.

Secondly we try to find an admissible predictor of Y within the class Ξ of linear predictors. It is easy to see that for $a = \frac{1}{n+1}$ and $b = 0$ the predictor $\delta(\mathbf{X}) = \frac{1}{n+1}\sum_{i=1}^n X_i$ has a constant risk function. Because of that there are a lot of elements in the class Ξ which dominate this predictor. So this predictor is not admissible.

Assume now that $b \neq 0$. Because the risk function is a trinomial, the following additional condition should be satisfied:

$$\theta_w = \frac{1-na}{b} \leqslant \frac{1}{\mu_0},$$

where θ_w is the argmin of the risk function. After simple calculations we get $b \geqslant \frac{1}{n+1}\mu_0$. Hence we can conclude that every element of the class

$$\Xi_a = \left\{ \frac{1}{n+1}\sum_{i=1}^n X_i + b; b \in \left[\frac{1}{n+1}\mu_0; \frac{2}{n+1}\mu_0\right] \right\}$$

is a minimax predictor of Y within the class of all regular predictors and an admissible one within the class of linear predictors.

REFERENCES

[1] B. Z. BOBROVSKY, E. MAYER-WOLF, AND M. ZAKAI, Some classes of global Cramér–Rao bounds, *The Annals of Statistics* 15, 1421–1438, (1987).

[2] D. BROWN, AND L. GAJEK, Information inequalities for the Bayes risk, *The Annals of Statistics* 18, 1578–1594, (1990).

[3] P. DIACONIS, AND D. YLVISAKER, *Conjugate priors for exponential families*, The Annals of Statistics 7, 269–281, (1979).

[4] L. GAJEK, *An improper Cramer–Rao lower bound*, Applicationes Mathematicae 19, 241–256, (1987).

[5] L. GAJEK, *On the minimax value in the scale model with truncated data*, The Annals of Statistics 16, 669–677, (1988).

[6] L. GAJEK, AND M. KAŁUSZKA, *Nonexponential applications of a global Cramer–Rao inequality*, Statistics 26, 111–122, (1995).

[7] L. GAJEK, AND V. LIPIŃSKA, *Sharp inequality for the Bayes prediction risk*, Inequality Theory and Applications, Vol. 5, 97–106, (2007)

[8] E. L. LEHMANN, AND G. CASELLA, *Theory of Point Estimation, Second edition*, Springer, (1998)

[9] R. W. MILLER, AND C. B. CHANG, *A modified Cramér–Rao bound and its applications*, IEEE Trans. Inform. Theory, IT-24, 398–400, (1978).

[10] Y. TAKADA, *Lower bound on the Bayes risk for statistical prediction problems*, Comm. Statist. Theory Methods 28, 693–703, (1999).

[11] H. VAN TREES, *Detection, Estimation and Modulation Theory 1*, Wiley, New York, (1968).

[12] S. TRYBUŁA, *Minimax mutual prediction*, Applicationes Mathematicae XXVII, 4, 437–444, (2000).

[13] S. TRYBUŁA, *Minimax mutual prediction of multinomial random variables*, Applicationes Mathematicae XXX, 4, 371–377, (2003).

(Received October 30, 2008)

L. Gajek
Institute of Mathematics
Technical University of Łódź
ul. Wólczańska 215
90–924 Łódź
Poland
and
Polish Financial Supervision Authority
Plac Powstańców Warszawy 1
00–950 Warszawa
Poland

V. Lipińska
Institute of Mathematics
Technical University of Łódź
ul. Wólczańska 215
90–924 Łódź
Poland