Prior and loss robustness for varoius loss functions

Agnieszka Kamińska and Zdzisław Porosiński

Institute of Mathematics and Computer Science,
Wroclaw University of Technology
Wybrzeze Wyspianskiego 27
50-370 Wroclaw

December 8, 2009

Model

Model

- Let $X_1, \ldots, X_n$ be i.i.d. random variables with a distribution P_ϑ indexed by a real parameter ϑ. We denote $X = (X_1, \ldots, X_n)$.

Model

- Let $X_1, \ldots, X_n$ be i.i.d. random variables with a distribution P_ϑ indexed by a real parameter ϑ.
  We denote $X = (X_1, \ldots, X_n)$.
- Let $(\mathcal{X}, \mathcal{B}, \mathcal{P})$ be a statistical space determined by X, where $\mathcal{X} \subset \mathbb{R}^n$, $\mathcal{B}$ is σ-field of $\mathcal{X}$ and $\mathcal{P} = \{P_\vartheta : \vartheta \in \Theta = \mathbb{R}\}$.

Model

- Let $X_1, \ldots, X_n$ be i.i.d. random variables with a distribution P_ϑ indexed by a real parameter ϑ.
  We denote $X = (X_1, \ldots, X_n)$.
- Let $(\mathcal{X}, \mathcal{B}, \mathcal{P})$ be a statistical space determined by X, where $\mathcal{X} \subset \mathbb{R}^n$, $\mathcal{B}$ is σ-field of $\mathcal{X}$ and $\mathcal{P} = \{P_\vartheta : \vartheta \in \Theta = \mathbb{R}\}$.
- Let $L(\vartheta, d)$ be a loss function.

Model

- Let $X_1, \dots, X_n$ be i.i.d. random variables with a distribution P_ϑ indexed by a real parameter ϑ.
  We denote $X = (X_1, \dots, X_n)$.
- Let $(\mathcal{X}, \mathcal{B}, \mathcal{P})$ be a statistical space determined by X, where $\mathcal{X} \subset \mathbb{R}^n$, $\mathcal{B}$ is σ-field of $\mathcal{X}$ and $\mathcal{P} = \{P_\vartheta : \vartheta \in \Theta = \mathbb{R}\}$.
- Let $L(\vartheta, d)$ be a loss function.
- Let ϑ have a prior distribution $\pi(\vartheta)$, defined on the measurable space (Θ, Ξ).

Model

- Let $X_1, \ldots, X_n$ be i.i.d. random variables with a distribution P_ϑ indexed by a real parameter ϑ.
  We denote $X = (X_1, \ldots, X_n)$.
- Let $(\mathcal{X}, \mathcal{B}, \mathcal{P})$ be a statistical space determined by X, where $\mathcal{X} \subset \mathbb{R}^n$, $\mathcal{B}$ is σ-field of $\mathcal{X}$ and $\mathcal{P} = \{P_\vartheta : \vartheta \in \Theta = \mathbb{R}\}$.
- Let $L(\vartheta, d)$ be a loss function.
- Let ϑ have a prior distribution $\pi(\vartheta)$, defined on the measurable space (Θ, Ξ).
- The posterior distribution has a form $\pi(\vartheta|x)$, for $X = x$.

Model

- Let $X_1, \ldots, X_n$ be i.i.d. random variables with a distribution P_ϑ indexed by a real parameter ϑ.
  We denote $X = (X_1, \ldots, X_n)$.
- Let $(\mathcal{X}, \mathcal{B}, \mathcal{P})$ be a statistical space determined by X, where $\mathcal{X} \subset \mathbb{R}^n$, $\mathcal{B}$ is σ-field of $\mathcal{X}$ and $\mathcal{P} = \{P_\vartheta : \vartheta \in \Theta = \mathbb{R}\}$.
- Let $L(\vartheta, d)$ be a loss function.
- Let ϑ have a prior distribution $\pi(\vartheta)$, defined on the measurable space (Θ, Ξ).
- The posterior distribution has a form $\pi(\vartheta|x)$, for $X = x$.
- We consider the problem of constructing the point Bayes estimator of ϑ under $L(\vartheta, d)$.

Bayesian estimation

- If $X = x$, then the posterior risk of d can be expressed as

$$\overline{R}_x(\pi, d) \quad = \quad E^{\pi|x}[L(\vartheta, d)] \,,$$

where $E^{\pi|x}[\cdot]$ denotes the expected value when $\vartheta \sim \pi(\vartheta|x)$.

Bayesian estimation

- If $X = x$, then the posterior risk of d can be expressed as

$$\overline{R}_x(\pi, d) \;=\; E^{\pi|x}[L(\vartheta, d)] \,,$$

where $E^{\pi|x}[\cdot]$ denotes the expected value when $\vartheta \sim \pi(\vartheta|x)$.

- The Bayes estimator $\widehat{\vartheta}^{\pi}$ satisfies

$$\overline{R}_x(\pi, \widehat{\vartheta}^{\pi}) \;=\; \inf_{d \in D} \overline{R}_x(\pi, d) \,.$$

Prior robustness

Information on the appropriate prior is often too inadequate to specify a prior distribution unambiguously.

Prior robustness

Information on the appropriate prior is often too inadequate to specify a prior distribution unambiguously.

The problem of expressing uncertainty regarding prior information can be solved by using a class Γ of prior distributions.

Prior robustness

Information on the appropriate prior is often too inadequate to specify a prior distribution unambiguously.

The problem of expressing uncertainty regarding prior information can be solved by using a class Γ of prior distributions.

Assume that the prior $\pi(\vartheta)$ belongs to the class Γ.

Γ-minimax estimators

Let $\overline{F}_x(\pi, d)$ be a posterior functional. The optimal decision $\widehat{\vartheta}$ satisfies

$$\sup_{\pi\in\Gamma} \overline{F}_x(\pi, \widehat{\vartheta}) = \inf_{d\in D} \sup_{\pi\in\Gamma} \overline{F}_x(\pi, d)\,.$$

Γ-minimax estimators

Let $\overline{F}_x(\pi, d)$ be a posterior functional. The optimal decision $\widehat{\vartheta}$ satisfies

$$\sup_{\pi\in\Gamma} \overline{F}_x(\pi, \widehat{\vartheta}) = \inf_{d\in D} \sup_{\pi\in\Gamma} \overline{F}_x(\pi, d) .$$

- the conditional Γ-minimax estimator

$$\overline{F}_x(\pi,d)=\overline{R}_x(\pi,d),$$

Γ-minimax estimators

Let $\overline{F}_x(\pi, d)$ be a posterior functional. The optimal decision $\widehat{\vartheta}$ satisfies

$$\sup_{\pi\in\Gamma} \overline{F}_x(\pi, \widehat{\vartheta}) = \inf_{d\in D} \sup_{\pi\in\Gamma} \overline{F}_x(\pi, d)\,.$$

- the conditional Γ-minimax estimator

$$\overline{F}_x(\pi,d)=\overline{R}_x(\pi,d),$$

- the posterior regret Γ-minimax estimator

$$\overline{F}_x(\pi,d)=\overline{R}_x(\pi,d)-\overline{R}_x(\pi,\widehat{\vartheta}^{\pi}),$$

Γ-minimax estimators

Let $\overline{F}_x(\pi, d)$ be a posterior functional. The optimal decision $\widehat{\vartheta}$ satisfies

$$\sup_{\pi\in\Gamma} \overline{F}_x(\pi, \widehat{\vartheta}) = \inf_{d\in D} \sup_{\pi\in\Gamma} \overline{F}_x(\pi, d).$$

- the conditional Γ-minimax estimator

$$\overline{F}_x(\pi,d)=\overline{R}_x(\pi,d),$$

- the posterior regret Γ-minimax estimator

$$\overline{F}_x(\pi,d)=\overline{R}_x(\pi,d)-\overline{R}_x(\pi,\widehat{\vartheta}^{\pi}),$$

- the most stable estimator

$$\overline{F}_x(\pi,d)=\sup_{\pi\in\Gamma} \overline{R}_x(\pi,d)-\inf_{\pi\in\Gamma} \overline{R}_x(\pi,d).$$

Prior and loss robustness

Prior and loss robustness

It could be also interesting to take into consideration a sensitivity analysis with uncertainty in both: the prior distribution and the loss function.

Prior and loss robustness

It could be also interesting to take into consideration a sensitivity analysis with uncertainty in both: the prior distribution and the loss function.

Assume that the prior $\pi(\vartheta)$ belongs to the class Γ and the loss function $L(\vartheta, d)$ is in the class $\mathcal{L}$.

Prior and loss robustness

It could be also interesting to take into consideration a sensitivity analysis with uncertainty in both: the prior distribution and the loss function.

Assume that the prior $\pi(\vartheta)$ belongs to the class Γ and the loss function $L(\vartheta, d)$ is in the class $\mathcal{L}$.

A review of available robust estimators in $\mathcal{L} \times \Gamma$ can be found in Arias et al. (2009).

$\mathcal{L} \times \Gamma$-minimax estimators

Let $\overline{F}_x(\pi, d)$ be a posterior functional. The optimal decision $\widehat{\vartheta}_L$ satisfies

$$\sup_{(L,\pi)\in\mathcal{L}\times\Gamma} \overline{F}_x(\pi, \widehat{\vartheta}) = \inf_{d\in D} \sup_{(L,\pi)\in\mathcal{L}\times\Gamma} \overline{F}_x(\pi, d)\,.$$

$\mathcal{L} \times \Gamma$-minimax estimators

Let $\overline{F}_x(\pi, d)$ be a posterior functional. The optimal decision $\widehat{\vartheta}_L$ satisfies

$$\sup_{(L,\pi)\in\mathcal{L}\times\Gamma} \overline{F}_x(\pi, \widehat{\vartheta}) = \inf_{d\in D} \sup_{(L,\pi)\in\mathcal{L}\times\Gamma} \overline{F}_x(\pi, d).$$

- the conditional $\mathcal{L} \times \Gamma$-minimax estimator

$$\overline{F}_x(\pi,d)=\overline{R}_x(\pi,d),$$

$\mathcal{L} \times \Gamma$-minimax estimators

Let $\overline{F}_x(\pi, d)$ be a posterior functional. The optimal decision $\widehat{\vartheta}_L$ satisfies

$$\sup_{(L,\pi)\in\mathcal{L}\times\Gamma} \overline{F}_x(\pi, \widehat{\vartheta}) = \inf_{d\in D} \sup_{(L,\pi)\in\mathcal{L}\times\Gamma} \overline{F}_x(\pi, d).$$

- the conditional $\mathcal{L} \times \Gamma$-minimax estimator

$$\overline{F}_x(\pi,d)=\overline{R}_x(\pi,d),$$

- the posterior regret $\mathcal{L} \times \Gamma$-minimax estimator

$$\overline{F}_x(\pi,d)=\overline{R}_x(\pi,d)-\overline{R}_x(\pi,\widehat{\vartheta}^{\pi}),$$

$\mathcal{L} \times \Gamma$-minimax estimators

Let $\overline{F}_x(\pi, d)$ be a posterior functional. The optimal decision $\widehat{\vartheta}_L$ satisfies

$$\sup_{(L,\pi)\in\mathcal{L}\times\Gamma} \overline{F}_x(\pi, \widehat{\vartheta}) = \inf_{d\in D} \sup_{(L,\pi)\in\mathcal{L}\times\Gamma} \overline{F}_x(\pi, d) .$$

- the conditional $\mathcal{L} \times \Gamma$-minimax estimator

$$\overline{F}_x(\pi,d)=\overline{R}_x(\pi,d),$$

- the posterior regret $\mathcal{L} \times \Gamma$-minimax estimator

$$\overline{F}_x(\pi,d)=\overline{R}_x(\pi,d)-\overline{R}_x(\pi,\widehat{\vartheta}^\pi),$$

- the most stable $\mathcal{L} \times \Gamma$-minimax estimator

$$\overline{F}_x(\pi,d)=\sup_{(L,\pi)\in\mathcal{L}\times\Gamma} \overline{R}_x(\pi,d)-\inf_{(L,\pi)\in\mathcal{L}\times\Gamma} \overline{R}_x(\pi,d) .$$

Symmetric loss function

The symmetric loss functions

- The unbounded square-error loss function SE

$$L_{SE}(\vartheta, d) = \gamma \left(d - \vartheta\right)^2$$

The symmetric loss functions

- The unbounded square-error loss function SE

$$L_{SE}(\vartheta, d) = \gamma\,(d - \vartheta)^2$$

- The bounded reflected normal loss function RN

$$L_{RN}(\vartheta, d) = K\,\left[\,1\, -\, e^{-\gamma\,(\,d - \,\vartheta)^2}\,\right]$$

The symmetric loss functions

- The unbounded square-error loss function SE

$$L_{SE}(\vartheta, d) = \gamma\,(d - \vartheta)^2$$

- The bounded reflected normal loss function RN

$$L_{RN}(\vartheta, d) = K\,\left[\,1 \,-\, e^{-\gamma\,(\,d - \vartheta)^2}\,\right]$$

The symmetric loss functions

- The unbounded square-error loss function SE

$$L_{SE}(\vartheta, d) = \gamma\,(d - \vartheta)^2$$

- The bounded reflected normal loss function RN

$$L_{RN}(\vartheta, d) = K\left[\,1 \,-\, e^{-\gamma\,(\,d - \vartheta)^2}\,\right]$$

γ - shape parameter, K - maximum loss.

Normal model

- $X = (X_1, \ldots, X_n) \backsim \mathcal{N}(\vartheta, \tau^2)$, where $\tau^2 > 0$ is known.

Normal model

- $X = (X_1, \ldots, X_n) \backsim \mathcal{N}(\vartheta, \tau^2)$, where $\tau^2 > 0$ is known.
- $\vartheta \backsim \pi(\vartheta) = \mathcal{N}(\mu, \sigma^2)$.

Normal model

- $X = (X_1, \ldots, X_n) \backsim \mathcal{N}(\vartheta, \tau^2)$, where $\tau^2 > 0$ is known.
- $\vartheta \backsim \pi(\vartheta) = \mathcal{N}(\mu, \sigma^2)$.
- $\vartheta \backsim \pi(\vartheta|x) = \mathcal{N}(\mu_n, \sigma_n^2)$

$$\mu_n = r\bar{x} + (1 - r)\mu, \qquad \sigma_n^2 = \tau^2 r/n$$

where $r = n\sigma^2/(n\sigma^2 + \tau^2)$.

Class of prior

Let ϑ have a prior distribution in the following class

$$\Gamma_{\sigma_0} = \{ \, \pi(\vartheta) : \quad \pi(\vartheta) = \mathcal{N}(\mu, \sigma_0^2), \quad \mu \in (\underline{\mu}, \overline{\mu}) \, \}.$$

Class of prior

Let ϑ have a prior distribution in the following class

$$\Gamma_{\sigma_0} = \{\ \pi(\vartheta) : \quad \pi(\vartheta) = \mathcal{N}(\mu, \sigma_0^2), \quad \mu \in (\underline{\mu}, \overline{\mu})\ \}.$$

Class of loss for SE

We considered the following class of loss functions

$$\mathcal{L}_{SE} = \{\ L(\vartheta, d) : \quad L_{SE}(\vartheta, d) = \gamma\,(d - \vartheta)^2, \quad \gamma \in (\underline{\gamma}, \overline{\gamma})\ \}.$$

Class of prior

Let ϑ have a prior distribution in the following class

$$\Gamma_{\sigma_0} = \{\ \pi(\vartheta) : \quad \pi(\vartheta) = \mathcal{N}(\mu, \sigma_0^2), \quad \mu \in (\underline{\mu}, \overline{\mu})\ \}.$$

Class of loss for SE

We considered the following class of loss functions

$$\mathcal{L}_{SE} = \{\ L(\vartheta, d) : \quad L_{SE}(\vartheta, d) = \gamma\,(d - \vartheta)^2, \quad \gamma \in (\underline{\gamma}, \overline{\gamma})\ \}.$$

Class of loss for RN

We considered the following class of loss functions

$$\mathcal{L}_{RN} = \{\ L(\vartheta, d) :\ L_{RN}(\vartheta, d) = K\left[1 - e^{-\gamma\,(d - \vartheta)^2}\right],\ \gamma \in (\underline{\gamma}, \overline{\gamma})\ \}.$$

Prior and loss robustness

Results for SE

- Γ_{σ_0} (Boratyńska and Męczarski 1994)

$$\widetilde{\vartheta} = \widehat{\vartheta}^{PR} = \widehat{\vartheta}^{S} = \frac{\underline{\mu}_n + \overline{\mu}_n}{2}$$

Results for SE

- Γ_{σ_0} (Boratyńska and Męczarski 1994)

$$\tilde{\vartheta} = \hat{\vartheta}^{PR} = \hat{\vartheta}^{S} = \frac{\underline{\mu}_n + \overline{\mu}_n}{2}$$

- $\mathcal{L}_{SE} \times \Gamma_{\sigma_0}$

$$\tilde{\vartheta}_L = \hat{\vartheta}_L^{PR} = \hat{\vartheta}_L^{S} = \frac{\underline{\mu}_n + \overline{\mu}_n}{2}$$

Results for RN

- Γ_{σ_0} (Kamińska 2008a)

$$\widetilde{\vartheta} = \widehat{\vartheta}^{PR} = \widehat{\vartheta}^{S} = \frac{\underline{\mu}_n + \overline{\mu}_n}{2}$$

Results for RN

- Γ_{σ_0} (Kamińska 2008a)

$$\tilde{\vartheta} = \hat{\vartheta}^{PR} = \hat{\vartheta}^{S} = \frac{\underline{\mu}_n + \overline{\mu}_n}{2}$$

- $\mathcal{L}_{RN} \times \Gamma_{\sigma_0}$ (Kamińska and Porosiński 2009)

$$\tilde{\vartheta}_L = \frac{\underline{\mu}_n + \overline{\mu}_n}{2}$$

Results for RN

Theorem

If $\mathcal{L}_{RN} \times \Gamma_{\sigma_0}$ *is the class of loss functions and prior distributions, then the posterior regret* $\mathcal{L} \times \Gamma$*-minimax estimator under the RN loss function* **can not be always calculated analytically**.

Results for RN

Theorem

If $\mathcal{L}_{RN} \times \Gamma_{\sigma_0}$ *is the class of loss functions and prior distributions, then the posterior regret* $\mathcal{L} \times \Gamma$*-minimax estimator under the RN loss function* **can not be always calculated analytically**.

Proof

Posterior risk: $$\overline{R}_x(\pi,d)=1-\frac{1}{\sqrt{1+2\gamma\sigma_n^2}}\exp\left\{-\gamma\,\frac{(d-\mu_n)^2}{1+2\gamma\sigma_n^2}\right\}.$$

Results for RN

Theorem

If $\mathcal{L}_{RN} \times \Gamma_{\sigma_0}$ is the class of loss functions and prior distributions, then the posterior regret $\mathcal{L} \times \Gamma$-minimax estimator under the RN loss function **can not be always calculated analytically**.

Proof

Posterior risk: $$\overline{R}_x(\pi,d)=1-\frac{1}{\sqrt{1+2\gamma\sigma_n^2}}\exp\left\{-\gamma\,\frac{(d-\mu_n)^2}{1+2\gamma\sigma_n^2}\right\}.$$

Posterior regret: $$f(d,\gamma,\mu_n)=\frac{1}{\sqrt{1+2\gamma\sigma_n^2}}\left[1-\exp\left\{-\gamma\,\frac{(d-\mu_n)^2}{1+2\gamma\sigma_n^2}\right\}\right].$$

Results for RN

Theorem

If $\mathcal{L}_{RN} \times \Gamma_{\sigma_0}$ is the class of loss functions and prior distributions, then the posterior regret $\mathcal{L} \times \Gamma$-minimax estimator under the RN loss function **can not be always calculated analytically**.

Proof

Posterior risk: $$\overline{R}_x(\pi,d)=1-\frac{1}{\sqrt{1+2\gamma\sigma_n^2}}\exp\left\{-\gamma\,\frac{(d-\mu_n)^2}{1+2\gamma\sigma_n^2}\right\}.$$

Posterior regret: $$f(d,\gamma,\mu_n)=\frac{1}{\sqrt{1+2\gamma\sigma_n^2}}\left[1-\exp\left\{-\gamma\,\frac{(d-\mu_n)^2}{1+2\gamma\sigma_n^2}\right\}\right].$$

Our goal is to find:

$$\inf_{d\in D}\ \sup_{(\gamma,\mu_n)\in Q}\ f(d,\gamma,\mu_n),$$

where $Q=(\underline{\gamma},\overline{\gamma})\times(\underline{\mu}_n,\overline{\mu}_n)$.

$$f(d,\gamma,\mu_n)=\frac{1}{\sqrt{1+2\gamma\sigma_n^2}}\left[1-\exp\left\{-\gamma\ \frac{(d-\mu_n)^2}{1+2\gamma\sigma_n^2}\right\}\right].$$

$$f(d,\gamma,\mu_n)=\frac{1}{\sqrt{1+2\gamma\sigma_n^2}}\left[1-\exp\left\{-\gamma\ \frac{(d-\mu_n)^2}{1+2\gamma\sigma_n^2}\right\}\right].$$

For any fixed d, let $f(d,\gamma,\mu_n)=h(\gamma,\mu_n)$

- $\frac{\delta h}{\delta\mu_n}\geq 0 \ \Leftrightarrow\ \mu_n\geq d$, but $h(\gamma,d)=0$ thus h has no extremum

$$f(d,\gamma,\mu_n)=\frac{1}{\sqrt{1+2\gamma\sigma_n^2}}\left[1-\exp\left\{-\gamma\ \frac{(d-\mu_n)^2}{1+2\gamma\sigma_n^2}\right\}\right].$$

For any fixed d, let $f(d,\gamma,\mu_n) = h(\gamma,\mu_n)$

- $\frac{\delta h}{\delta \mu_n} \geq 0 \ \Leftrightarrow \ \mu_n \geq d$, but $h(\gamma, d) = 0$ thus h has no extremum
- $h(\gamma,\mu_n) > 0$ for $\gamma > 0$ and $d \neq \mu_n$.

$$f(d,\gamma,\mu_n)=\frac{1}{\sqrt{1+2\gamma\sigma_n^2}}\left[1-\exp\left\{-\gamma\ \frac{(d-\mu_n)^2}{1+2\gamma\sigma_n^2}\right\}\right].$$

For any fixed d, let $f(d,\gamma,\mu_n)=h(\gamma,\mu_n)$

- $\frac{\delta h}{\delta \mu_n} \geq 0 \ \Leftrightarrow\ \mu_n \geq d$, but $h(\gamma,d)=0$ thus h has no extremum
- $h(\gamma,\mu_n)>0$ for $\gamma>0$ and $d \neq \mu_n$.

Since $h(0,\mu_n)=0$ and $\lim_{\gamma\to\infty} h(\gamma,\mu_n)=0$ thus h has at least one local maximum as the function of γ.

$$f(d,\gamma,\mu_n)=\frac{1}{\sqrt{1+2\gamma\sigma_n^2}}\left[1-\exp\left\{-\gamma\ \frac{(d-\mu_n)^2}{1+2\gamma\sigma_n^2}\right\}\right].$$

For any fixed d, let $f(d,\gamma,\mu_n) = h(\gamma,\mu_n)$

- $\frac{\delta h}{\delta \mu_n} \geq 0 \ \Leftrightarrow\ \mu_n \geq d$, but $h(\gamma, d) = 0$ thus h has no extremum
- $h(\gamma,\mu_n) > 0$ for $\gamma > 0$ and $d \neq \mu_n$.

Since $h(0,\mu_n) = 0$ and $\lim_{\gamma\to\infty} h(\gamma,\mu_n) = 0$ thus h has at least one local maximum as the function of γ.

Let suppose that for $\gamma \in (\underline{\gamma},\overline{\gamma})$ functions $h(\gamma,\underline{\mu}_n)$ and $h(\gamma,\overline{\mu}_n)$ have the maxima at points γ_1 and γ_2, **respecivly**. Then

$$f(d,\gamma,\mu_n)=\frac{1}{\sqrt{1+2\gamma\sigma_n^2}}\left[1-\exp\left\{-\gamma\ \frac{(d-\mu_n)^2}{1+2\gamma\sigma_n^2}\right\}\right].$$

For any fixed d, let $f(d,\gamma,\mu_n)=h(\gamma,\mu_n)$

- $\frac{\delta h}{\delta \mu_n} \geq 0 \ \Leftrightarrow\ \mu_n \geq d$, but $h(\gamma,d)=0$ thus h has no extremum
- $h(\gamma,\mu_n)>0$ for $\gamma>0$ and $d \neq \mu_n$.

Since $h(0,\mu_n)=0$ and $\lim_{\gamma\to\infty} h(\gamma,\mu_n)=0$ thus h has at least one local maximum as the function of γ.

Let suppose that for $\gamma \in (\underline{\gamma},\overline{\gamma})$ functions $h(\gamma,\underline{\mu}_n)$ and $h(\gamma,\overline{\mu}_n)$ have the maxima at points γ_1 and γ_2, **respecivly**. Then

$$\inf_{d\in D}\ \sup_{(\gamma,\mu_n)\in Q}\ f(d,\gamma,\mu_n)=\inf_{d\in D}\begin{cases} f(d,\gamma_2,\overline{\mu}_n) & d\leq(\underline{\mu}_n+\overline{\mu}_n)/2 \\ f(d,\gamma_1,\underline{\mu}_n) & d>(\underline{\mu}_n+\overline{\mu}_n)/2\end{cases}$$

$$f(d,\gamma,\mu_n)=\frac{1}{\sqrt{1+2\gamma\sigma_n^2}}\left[1-\exp\left\{-\gamma\ \frac{(d-\mu_n)^2}{1+2\gamma\sigma_n^2}\right\}\right].$$

For any fixed d, let $f(d,\gamma,\mu_n)=h(\gamma,\mu_n)$

- $\frac{\delta h}{\delta \mu_n} \geq 0 \ \Leftrightarrow \ \mu_n \geq d$, but $h(\gamma, d) = 0$ thus h has no extremum
- $h(\gamma,\mu_n) > 0$ for $\gamma > 0$ and $d \neq \mu_n$.

Since $h(0,\mu_n) = 0$ and $\lim_{\gamma\to\infty} h(\gamma,\mu_n) = 0$ thus h has at least one local maximum as the function of γ.

Let suppose that for $\gamma \in (\underline{\gamma}, \overline{\gamma})$ functions $h(\gamma, \underline{\mu}_n)$ and $h(\gamma, \overline{\mu}_n)$ have the maxima at points γ_1 and γ_2, **respecivly**. Then

$$\inf_{d\in D}\ \sup_{(\gamma,\mu_n)\in Q}\ f(d,\gamma,\mu_n) = \inf_{d\in D} \begin{cases} f(d,\gamma_2,\overline{\mu}_n) & d \leq (\underline{\mu}_n+\overline{\mu}_n)/2 \\ f(d,\gamma_1,\underline{\mu}_n) & d > (\underline{\mu}_n+\overline{\mu}_n)/2 \end{cases}$$

is reached for d that is solution of $f(d,\gamma_1,\overline{\mu}_n) = f(d,\gamma_2,\underline{\mu}_n)$.

Results for RN

Theorem

If $\mathcal{L}_{RN} \times \Gamma_{\sigma_0}$ is the class of loss functions and prior distributions, then the most stable $\mathcal{L} \times \Gamma$-minimax estimator under the RN loss function **does not exist**.

Results for RN

Theorem

If $\mathcal{L}_{RN} \times \Gamma_{\sigma_0}$ is the class of loss functions and prior distributions, then the most stable $\mathcal{L} \times \Gamma$-minimax estimator under the RN loss function **does not exist**.

Proof

Posterior risk: $\overline{R}_x(\pi,d)=1-\frac{1}{\sqrt{1+2\gamma\sigma_n^2}}\exp\left\{-\gamma\,\frac{(d-\mu_n)^2}{1+2\gamma\sigma_n^2}\right\}=f(d,\gamma,\mu_n)\,.$

Results for RN

Theorem

If $\mathcal{L}_{RN} \times \Gamma_{\sigma_0}$ is the class of loss functions and prior distributions, then the most stable $\mathcal{L} \times \Gamma$-minimax estimator under the RN loss function **does not exist**.

Proof

Posterior risk: $\overline{R}_x(\pi,d) = 1 - \frac{1}{\sqrt{1+2\gamma\sigma_n^2}} \exp\left\{ -\gamma \frac{(d-\mu_n)^2}{1+2\gamma\sigma_n^2} \right\} = f(d,\gamma,\mu_n)$.

Our goal is to find:

$$\inf_{d\in D} \left[\sup_{(\gamma,\mu_n)\in Q} f(d,\gamma,\mu_n) - \inf_{(\gamma,\mu_n)\in Q} f(d,\gamma,\mu_n) \right],$$

where $Q = (\underline{\gamma}, \overline{\gamma}) \times (\underline{\mu}_n, \overline{\mu}_n)$.

$$f(d,\gamma,\mu_n) = 1 - \frac{1}{\sqrt{1+2\gamma\sigma_n^2}} \exp\left\{-\gamma \,\frac{(d-\mu_n)^2}{1+2\gamma\sigma_n^2}\right\} .$$

$$f(d,\gamma,\mu_n) = 1 - \frac{1}{\sqrt{1+2\gamma\sigma_n^2}} \exp\left\{-\gamma \frac{(d-\mu_n)^2}{1+2\gamma\sigma_n^2}\right\}.$$

For any fixed d, let $f(d,\gamma,\mu_n) = h(\gamma,\mu_n)$

- $\frac{\delta h}{\delta \mu_n} \geq 0 \;\Leftrightarrow\; \mu_n \geq d$, but $h(\gamma, d) = 0$ thus h has no extremum

$$f(d,\gamma,\mu_n) = 1 - \frac{1}{\sqrt{1+2\gamma\sigma_n^2}} \exp\left\{-\gamma \frac{(d-\mu_n)^2}{1+2\gamma\sigma_n^2}\right\} .$$

For any fixed d, let $f(d,\gamma,\mu_n) = h(\gamma,\mu_n)$

- $\frac{\delta h}{\delta \mu_n} \geq 0 \Leftrightarrow \mu_n \geq d$, but $h(\gamma, d) = 0$ thus h has no extremum
- $\frac{\delta h}{\delta \gamma} > 0 \Leftrightarrow \sigma_n^2 (1+2\gamma\sigma_n^2)^2 + (\mu_n - d)^2 > 0$ for any γ.

$$f(d,\gamma,\mu_n) = 1 - \frac{1}{\sqrt{1+2\gamma\sigma_n^2}} \exp\left\{-\gamma\,\frac{(d-\mu_n)^2}{1+2\gamma\sigma_n^2}\right\}.$$

For any fixed d, let $f(d,\gamma,\mu_n) = h(\gamma,\mu_n)$

- $\frac{\delta h}{\delta \mu_n} \geq 0 \;\Leftrightarrow\; \mu_n \geq d$, but $h(\gamma, d) = 0$ thus h has no extremum
- $\frac{\delta h}{\delta \gamma} > 0 \;\Leftrightarrow\; \sigma_n^2(1+2\gamma\sigma_n^2)^2 + (\mu_n - d)^2 > 0$ for any γ.

Thus

$$\sup_{(\gamma,\mu_n)\in Q} h(\gamma,\mu_n) = \begin{cases} h(\overline{\gamma},\underline{\mu}_n), & d \leq (\underline{\mu}_n + \overline{\mu}_n)/2 \\ h(\overline{\gamma},\overline{\mu}_n), & d > (\underline{\mu}_n + \overline{\mu}_n)/2 \end{cases},$$

$$f(d, \gamma, \mu_n) = 1 - \frac{1}{\sqrt{1 + 2\gamma\sigma_n^2}} \exp\left\{-\gamma \frac{(d - \mu_n)^2}{1 + 2\gamma\sigma_n^2}\right\}.$$

For any fixed d, let $f(d, \gamma, \mu_n) = h(\gamma, \mu_n)$

- $\frac{\delta h}{\delta \mu_n} \geq 0 \Leftrightarrow \mu_n \geq d$, but $h(\gamma, d) = 0$ thus h has no extremum
- $\frac{\delta h}{\delta \gamma} > 0 \Leftrightarrow \sigma_n^2 (1+2\gamma\sigma_n^2)^2 + (\mu_n - d)^2 > 0$ for any γ.

Thus

$$\sup_{(\gamma,\mu_n)\in Q} h(\gamma, \mu_n) = \begin{cases} h(\overline{\gamma}, \underline{\mu}_n), & d \leq (\underline{\mu}_n + \overline{\mu}_n)/2 \\ h(\overline{\gamma}, \overline{\mu}_n), & d > (\underline{\mu}_n + \overline{\mu}_n)/2 \end{cases},$$

$$\inf_{(\gamma,\mu_n)\in Q} h(\gamma, \mu_n) = \begin{cases} h(\underline{\gamma}, \underline{\mu}_n), & d \leq \underline{\mu}_n \\ h(\underline{\gamma}, d), & \underline{\mu}_n < d \leq \overline{\mu}_n \\ h(\underline{\gamma}, \overline{\mu}_n), & d > \overline{\mu}_n \end{cases}.$$

$$f(d, \gamma, \mu_n) = 1 - \frac{1}{\sqrt{1 + 2\gamma\sigma_n^2}} \exp\left\{ -\gamma \frac{(d - \mu_n)^2}{1 + 2\gamma\sigma_n^2} \right\} .$$

$$f(d,\gamma,\mu_n) = 1 - \frac{1}{\sqrt{1+2\gamma\sigma_n^2}} \exp\left\{-\gamma \frac{(d-\mu_n)^2}{1+2\gamma\sigma_n^2}\right\}.$$

$$\inf_{d\in D} \begin{cases} f(d,\overline{\gamma},\overline{\mu}_n)-f(d,\underline{\gamma},\underline{\mu}_n), & d \le \underline{\mu}_n \\ f(d,\overline{\gamma},\overline{\mu}_n)-f(d,\underline{\gamma},d), & \underline{\mu}_n < d \le (\underline{\mu}_n+\overline{\mu}_n)/2 \\ f(d,\overline{\gamma},\underline{\mu}_n)-f(d,\underline{\gamma},d), & (\underline{\mu}_n+\overline{\mu}_n)/2 < d \le \overline{\mu}_n \\ f(d,\overline{\gamma},\underline{\mu}_n)-f(d,\underline{\gamma},\overline{\mu}_n), & d > \overline{\mu}_n \end{cases} = \inf_{d\in D} p(d).$$

$$f(d,\gamma,\mu_n) = 1 - \frac{1}{\sqrt{1+2\gamma\sigma_n^2}} \exp\left\{-\gamma \frac{(d-\mu_n)^2}{1+2\gamma\sigma_n^2}\right\} .$$

$$\inf_{d\in D} \begin{cases} f(d,\overline{\gamma},\overline{\mu}_n) - f(d,\underline{\gamma},\underline{\mu}_n), & d \le \underline{\mu}_n \\ f(d,\overline{\gamma},\overline{\mu}_n) - f(d,\underline{\gamma},d), & \underline{\mu}_n < d \le (\underline{\mu}_n + \overline{\mu}_n)/2 \\ f(d,\overline{\gamma},\underline{\mu}_n) - f(d,\underline{\gamma},d), & (\underline{\mu}_n + \overline{\mu}_n)/2 < d \le \overline{\mu}_n \\ f(d,\overline{\gamma},\underline{\mu}_n) - f(d,\underline{\gamma},\overline{\mu}_n), & d > \overline{\mu}_n \end{cases} = \inf_{d\in D} p(d).$$

- $p(d) > 0$

$$f(d,\gamma,\mu_n) = 1 - \frac{1}{\sqrt{1+2\gamma\sigma_n^2}} \exp\left\{-\gamma\,\frac{(d-\mu_n)^2}{1+2\gamma\sigma_n^2}\right\}.$$

$$\inf_{d\in D}\begin{cases} f(d,\overline{\gamma},\overline{\mu}_n)-f(d,\underline{\gamma},\underline{\mu}_n), & d\le\underline{\mu}_n \\ f(d,\overline{\gamma},\overline{\mu}_n)-f(d,\underline{\gamma},d), & \underline{\mu}_n<d\le(\underline{\mu}_n+\overline{\mu}_n)/2 \\ f(d,\overline{\gamma},\underline{\mu}_n)-f(d,\underline{\gamma},d), & (\underline{\mu}_n+\overline{\mu}_n)/2<d\le\overline{\mu}_n \\ f(d,\overline{\gamma},\underline{\mu}_n)-f(d,\underline{\gamma},\overline{\mu}_n), & d>\overline{\mu}_n \end{cases} = \inf_{d\in D} p(d).$$

- $p(d) > 0$
- $\lim_{d\to-\infty} p(d) = 1 - 1 = 0$

$$f(d,\gamma,\mu_n) = 1 - \frac{1}{\sqrt{1+2\gamma\sigma_n^2}} \exp\left\{-\gamma \frac{(d-\mu_n)^2}{1+2\gamma\sigma_n^2}\right\}.$$

$$\inf_{d\in D} \begin{cases} f(d,\overline{\gamma},\overline{\mu}_n)-f(d,\underline{\gamma},\underline{\mu}_n), & d\le \underline{\mu}_n \\ f(d,\overline{\gamma},\overline{\mu}_n)-f(d,\underline{\gamma},d), & \underline{\mu}_n < d \le (\underline{\mu}_n+\overline{\mu}_n)/2 \\ f(d,\overline{\gamma},\underline{\mu}_n)-f(d,\underline{\gamma},d), & (\underline{\mu}_n+\overline{\mu}_n)/2 < d \le \overline{\mu}_n \\ f(d,\overline{\gamma},\underline{\mu}_n)-f(d,\underline{\gamma},\overline{\mu}_n), & d > \overline{\mu}_n \end{cases} = \inf_{d\in D} p(d).$$

- $p(d) > 0$
- $\lim_{d\to-\infty} p(d) = 1 - 1 = 0$
- $\lim_{d\to\infty} p(d) = 0$

$$f(d, \gamma, \mu_n) = 1 - \frac{1}{\sqrt{1 + 2\gamma\sigma_n^2}} \exp\left\{-\gamma \frac{(d - \mu_n)^2}{1 + 2\gamma\sigma_n^2}\right\}.$$

$$\inf_{d \in D} \begin{cases} f(d,\overline{\gamma},\overline{\mu}_n) - f(d,\underline{\gamma},\underline{\mu}_n), & d \leq \underline{\mu}_n \\ f(d,\overline{\gamma},\overline{\mu}_n) - f(d,\underline{\gamma},d), & \underline{\mu}_n < d \leq (\underline{\mu}_n + \overline{\mu}_n)/2 \\ f(d,\overline{\gamma},\underline{\mu}_n) - f(d,\underline{\gamma},d), & (\underline{\mu}_n + \overline{\mu}_n)/2 < d \leq \overline{\mu}_n \\ f(d,\overline{\gamma},\underline{\mu}_n) - f(d,\underline{\gamma},\overline{\mu}_n), & d > \overline{\mu}_n \end{cases} = \inf_{d \in D} p(d).$$

- $p(d) > 0$
- $\lim_{d \to -\infty} p(d) = 1 - 1 = 0$
- $\lim_{d \to \infty} p(d) = 0$

thus the most stable estimator does not exist.

Asymmetric loss function ABL

The asymmetric loss function

The bounded and asymmetric loss function ABL

$$L_{ABL}(\vartheta, d) = K \left[1 - \left(\frac{\vartheta}{d} \, e^{\,1-\frac{\vartheta}{d}} \right)^{\rho} \right],$$

where ρ is a shape parameter and K denotes the maximum loss.

Model

- Let $X \sim P_\vartheta \in \mathcal{P}$ with densities of the form

$$p_\vartheta(y) = c(y)\,\vartheta^{\,t(y)}\,e^{-s(y)\,\vartheta}, \qquad y \in \mathbb{R},\ \vartheta > 0$$

Model

- Let $X \sim P_\vartheta \in \mathcal{P}$ with densities of the form

$$p_\vartheta(y) = c(y)\, \vartheta^{\,t(y)}\, e^{-s(y)\,\vartheta}, \qquad y \in \mathbb{R},\ \vartheta > 0$$

- Representation of the family $\mathcal{P}$

Distribution	$t(y)$	$s(y)$	$p_\vartheta(y)$
Poisson $P(\vartheta)$	y	1	$\frac{\vartheta^y}{y!}\, e^{-\vartheta}$
Exponential $E(\vartheta)$	1	y	$\vartheta\, e^{-\vartheta y}$
Gamma $\mathcal{G}(\chi, \vartheta)$	χ	y	$\frac{\vartheta^\chi}{\Gamma(\chi)}\, y^{\chi-1}\, e^{-\vartheta y}$
Normal $\mathcal{N}(\mu, \frac{1}{\vartheta})$	$\frac{1}{2}$	$\frac{(y-\mu)^2}{2}$	$\sqrt{\frac{\vartheta}{2\pi}}\, e^{-\frac{(y-\mu)^2}{2}\vartheta}$
Pareto $\mathcal{P}a(\lambda, \vartheta)$	1	$\ln \frac{y}{\lambda}$	$\frac{\vartheta \lambda^\vartheta}{y^{\vartheta+1}}$

Model

Bayesian approach to a statistical problem requires defining a prior distribution over a parameter space. Let

- $\pi(\vartheta) = \mathcal{G}(\alpha, \beta)$

Model

Bayesian approach to a statistical problem requires defining a prior distribution over a parameter space. Let

- $\pi(\vartheta) = \mathcal{G}(\alpha, \beta)$

We assume the conjugate family of prior distribution, thus

Model

Bayesian approach to a statistical problem requires defining a prior distribution over a parameter space. Let

- $\pi(\vartheta) = \mathcal{G}(\alpha, \beta)$

We assume the conjugate family of prior distribution, thus

- $\pi(\vartheta|x) = \mathcal{G}(\alpha + T, \beta + S) = \mathcal{G}(\alpha_n, \beta_n)$ for $X = x$

$$T = T(x) = \sum_{i=1}^{n} t(x_i), \quad S = S(x) = \sum_{i=1}^{n} s(x_i)$$

Class of prior

Let ϑ have a prior distribution in the following class

$$\Gamma_{\alpha_0} = \{ \pi(\vartheta) : \quad \pi(\vartheta) = \mathcal{G}(\alpha_0, \beta), \quad \beta \in (\underline{\beta}, \overline{\beta}), \quad \underline{\beta} < \overline{\beta} \},$$

Class of prior

Let ϑ have a prior distribution in the following class

$$\Gamma_{\alpha_0} = \{ \pi(\vartheta) : \quad \pi(\vartheta) = \mathcal{G}(\alpha_0, \beta), \quad \beta \in (\underline{\beta}, \overline{\beta}), \quad \underline{\beta} < \overline{\beta} \},$$

Class of loss

We considered the following class of loss functions

$$\mathcal{L}_{ABL} = \{ L(\vartheta, d) : \; L(\vartheta, d) = K \left[1 - \left(\frac{\vartheta}{d} \, e^{\,1-\frac{\vartheta}{d}} \right)^{\rho} \right], \; \rho \in (\underline{\rho}, \overline{\rho}) \}.$$

Results for ABL - the conditional estimator

- Γ_{α_0} (Kamińska and Porosiński 2008b)

$$\widetilde{\vartheta} = \widehat{\vartheta}^{PR} = \widehat{\vartheta}^{S} = \rho \cdot \frac{(\underline{\beta}+T(x))^{-\frac{\alpha_0+T(x)}{\alpha_0+T(x)+\rho}} - (\overline{\beta}+S(x))^{-\frac{\alpha_0+T(x)}{\alpha_0+T(x)+\rho}}}{(\overline{\beta}+S(x))^{1-\frac{\alpha_0+T(x)}{\alpha_0+T(x)+\rho}} - (\underline{\beta}+S(x))^{1-\frac{\alpha_0+T(x)}{\alpha_0+T(x)+\rho}}}$$

Results for ABL - the conditional estimator

- Γ_{α_0} (Kamińska and Porosiński 2008b)

$$\widetilde{\vartheta} = \widehat{\vartheta}^{\,PR} = \widehat{\vartheta}^{\,S} = \rho \cdot \frac{(\underline{\beta}+T(x))^{-\frac{\alpha_0+T(x)}{\alpha_0+T(x)+\rho}} - (\overline{\beta}+S(x))^{-\frac{\alpha_0+T(x)}{\alpha_0+T(x)+\rho}}}{(\overline{\beta}+S(x))^{1-\frac{\alpha_0+T(x)}{\alpha_0+T(x)+\rho}} - (\underline{\beta}+S(x))^{1-\frac{\alpha_0+T(x)}{\alpha_0+T(x)+\rho}}}$$

- $\mathcal{L}_{ABL} \times \Gamma_{\alpha_0}$ (Kamińska and Porosiński 2009)

$$\widetilde{\vartheta}_L = \overline{\rho} \cdot \frac{(\underline{\beta}+S(x))^{-\frac{\alpha_0+T(x)}{\alpha_0+T(x)+\overline{\rho}}} - (\overline{\beta}+S(x))^{-\frac{\alpha_0+T(x)}{\alpha_0+T(x)+\overline{\rho}}}}{(\overline{\beta}+S(x))^{1-\frac{\alpha_0+T(x)}{\alpha_0+T(x)+\overline{\rho}}} - (\underline{\beta}+S(x))^{1-\frac{\alpha_0+T(x)}{\alpha_0+T(x)+\overline{\rho}}}}$$

Results for ABL

Theorem

If $\mathcal{L}_{ABL} \times \Gamma_{\alpha_0}$ *is the class of loss functions and prior distributions, then the posterior regret* $\mathcal{L} \times \Gamma$*-minimax estimator under the ABL loss function* **can not be always calculated analytically**.

Results for ABL

Theorem

If $\mathcal{L}_{ABL} \times \Gamma_{\alpha_0}$ is the class of loss functions and prior distributions, then the posterior regret $\mathcal{L} \times \Gamma$-minimax estimator under the ABL loss function **can not be always calculated analytically**.

Theorem

If $\mathcal{L}_{ABL} \times \Gamma_{\alpha_0}$ is the class of loss functions and prior distributions, then the most stable $\mathcal{L} \times \Gamma$-minimax estimator under the ABL loss function **does not exist**.

Remrks

Remrks

Is it true?

If the posterior risk is strictly increasing function of parameter of the loss function, then the conditional $\mathcal{L} \times \Gamma$-minimax estimator has the same form as the conditional Γ-minimax estimator.

Remrks

Is it true?

If the posterior risk is strictly increasing function of parameter of the loss function, then the conditional $\mathcal{L} \times \Gamma$-minimax estimator has the same form as the conditional Γ-minimax estimator.

Is it true?

The most stable estimator does not exist for the bounded loss functions.

Bibliography

- Arias, P., Martin, J., Ruggeri, F., Suárez, A. (2009). Optimal actions in problems with convex loss function*International Journal of Approximate Reasoning*, 50, 303-314.
- Boratyńska, A., Męczarski, M. (1994). Robust Bayesian estimation in the one-dimensional normal model. *Statist. Dec.* 12, 221-230.
- Kamińska, A. (2008a), The equivalence of Bayes and robust Bayes estimators for various loss functions. To appear in *Statistical Papers*.
- Kamińska, A., Porosiński, Z. (2008b). On robust Bayesian estimation under some asymmetric and bounded loss function. To appear in *Statistics*.
- Kamińska, A., Porosiński, Z. (2009). Prior and loss robustness. Preprint.