Numerical Contributions to the Asymptotic Theory of Robustness

Matthias Kohl

Fakultät für Mathematik und Physik
Universität Bayreuth

Promotionskolloquium
15.12.2005

Outline

Outline

Outline

Ideal Model

- parametric family of probability measures

$$\mathcal{P} = \{P_\theta \,|\, \theta \in \Theta\} \qquad \Theta \subset \mathbb{R}^k \text{ (open)}$$

- smoothly parameterized; i.e., L_2 differentiable at $\theta \in \Theta$ with L_2 derivative $\Lambda_\theta \in L_2^k(P_\theta)$, $\mathrm{E}_\theta \Lambda_\theta = 0$ and

- Fisher information of full rank

$$\mathcal{I}_\theta = \mathrm{E}_\theta \Lambda_\theta \Lambda_\theta^\tau \qquad \mathcal{I}_\theta \succ 0$$

Ideal Model

- parametric family of probability measures

$$\mathcal{P} = \{\mathrm{P}_\theta \,|\, \theta \in \Theta\} \qquad \Theta \subset \mathbb{R}^{\mathrm{k}} \text{ (open)}$$

- smoothly parameterized; i.e., L_2 differentiable at $\theta \in \Theta$ with L_2 derivative $\Lambda_\theta \in \mathrm{L}_2^{\mathrm{k}}(\mathrm{P}_\theta)$, $\mathrm{E}_\theta\, \Lambda_\theta = 0$ and

- Fisher information of full rank

$$\mathcal{I}_\theta = \mathrm{E}_\theta\, \Lambda_\theta \Lambda_\theta^\tau \qquad \mathcal{I}_\theta \succ 0$$

Ideal Model

- parametric family of probability measures

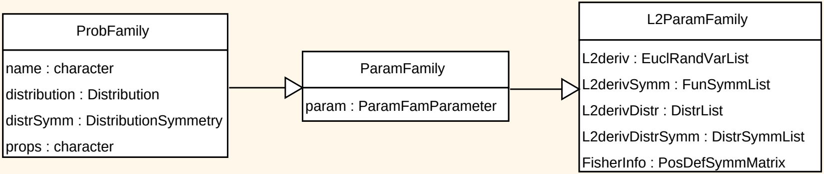
$$\mathcal{P} = \{\mathrm{P}_\theta \,|\, \theta \in \Theta\} \qquad \Theta \subset \mathbb{R}^{\mathrm{k}} \text{ (open)}$$

- smoothly parameterized; i.e., L_2 differentiable at $\theta \in \Theta$ with L_2 derivative $\Lambda_\theta \in \mathrm{L}_2^{\mathrm{k}}(\mathrm{P}_\theta)$, $\mathrm{E}_\theta\, \Lambda_\theta = 0$ and

- Fisher information of full rank

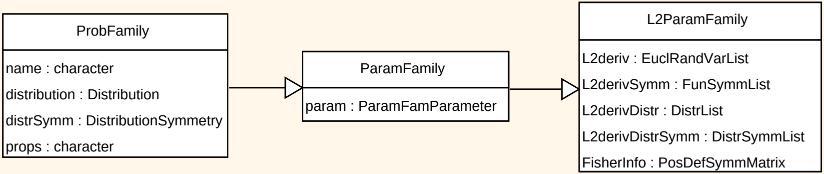
$$\mathcal{I}_\theta = \mathrm{E}_\theta\, \Lambda_\theta \Lambda_\theta^\tau \qquad \mathcal{I}_\theta \succ 0$$

Ideal Model in S4 Classes



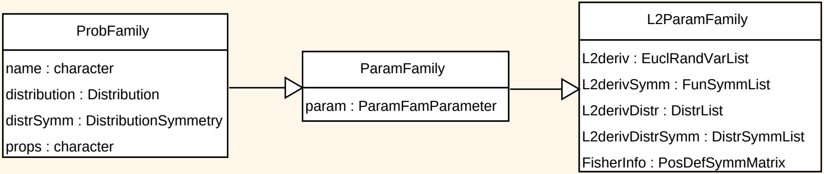
- meta-information slot `props`
- semi-symbolic calculus for symmetry properties
- generating functions for various L_2-families

Ideal Model in S4 Classes



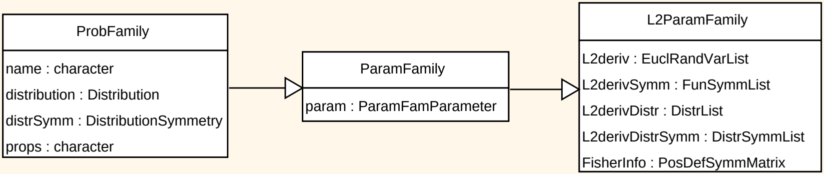
- meta-information slot `props`
- semi-symbolic calculus for symmetry properties
- generating functions for various L_2-families

Ideal Model in S4 Classes



- meta-information slot `props`
- semi-symbolic calculus for symmetry properties
- generating functions for various L_2-families

Ideal Model in S4 Classes



- meta-information slot `props`
- semi-symbolic calculus for symmetry properties
- generating functions for various L_2-families

Influence Curves (ICs) and AL Estimators

Definition

The set $\Psi_2(\theta)$ *of all* **square integrable ICs at P_θ** *is*

$$\Psi_2(\theta) = \left\{ \psi_\theta \in L_2^k(P_\theta) \,\middle|\, E_\theta\, \psi_\theta = 0,\ E_\theta\, \psi_\theta \Lambda_\theta^\tau = \mathbb{I}_k \right\}$$

Definition

An asymptotic estimator $S_n : (\Omega^n, \mathcal{A}^n) \to (\mathbb{R}^k, \mathbb{B}^k)$ *is called* **asymptotically linear at P_θ** *if there is an IC* $\psi_\theta \in \Psi_2(\theta)$ *with*

$$\sqrt{n}\,(S_n - \theta) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \psi_\theta(y_i) + o_{P_\theta^n}(n^0)$$

Influence Curves (ICs) and AL Estimators

Definition

The set $\Psi_2(\theta)$ of all **square integrable ICs at P_θ** *is*

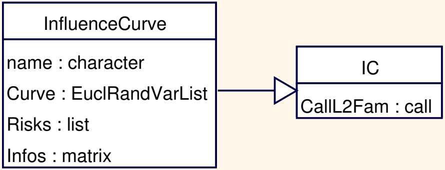
$$\Psi_2(\theta) = \left\{\psi_\theta \in L_2^k(P_\theta) \,\middle|\, E_\theta\, \psi_\theta = 0,\; E_\theta\, \psi_\theta \Lambda_\theta^\tau = \mathbb{I}_k\right\}$$

Definition

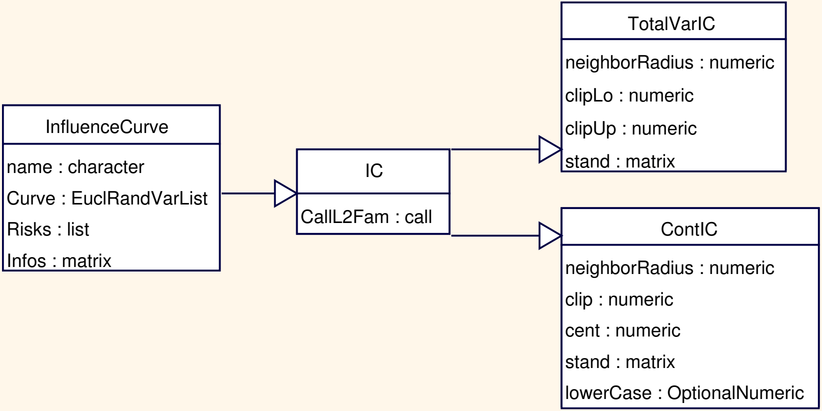
An asymptotic estimator $S_n : (\Omega^n, \mathcal{A}^n) \to (\mathbb{R}^k, \mathbb{B}^k)$ is called **asymptotically linear at P_θ** *if there is an IC $\psi_\theta \in \Psi_2(\theta)$ with*

$$\sqrt{n}\,(S_n - \theta) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \psi_\theta(y_i) + o_{P_\theta^n}(n^0)$$

Influence Curves in S4 Classes



Influence Curves in S4 Classes



Outline

Neighborhoods

Convex contamination neighborhood of radius $r \in [0, \infty)$

$$U_c(\theta, r) = \big\{(1-r)_+ P_\theta + (1 \wedge r)\, Q \,\big|\, Q \in \mathcal{M}_1(\mathcal{A})\big\}$$

Simple perturbations for $\sqrt{n} \geq -r \inf_{P_\theta} q$, are defined as

$$dQ_n(q, r) = \left(1 + \frac{r}{\sqrt{n}} q\right) dP_\theta$$

where

$$\mathcal{G}_c(\theta) = \big\{q \in L_\infty(P_\theta) \,\big|\, E_\theta\, q = 0,\ \inf_{P_\theta} q \geq -1\big\}$$

Neighborhoods

Convex contamination neighborhood of radius $r \in [0, \infty)$

$$U_c(\theta, r) = \{(1-r)_+ P_\theta + (1 \wedge r)\, Q \,|\, Q \in \mathcal{M}_1(\mathcal{A})\}$$

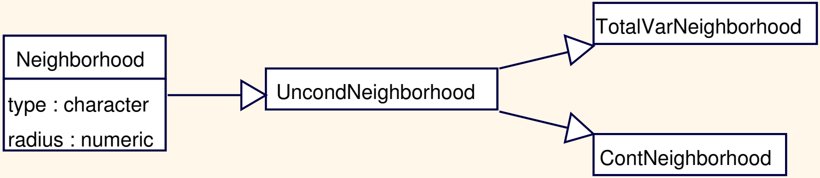
Simple perturbations for $\sqrt{n} \geq -r \inf_{P_\theta} q$, are defined as

$$dQ_n(q, r) = \left(1 + \frac{r}{\sqrt{n}} q\right) dP_\theta$$

where

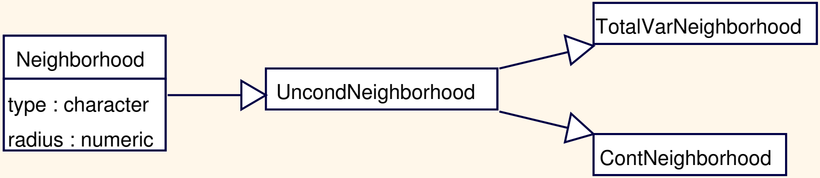
$$\mathcal{G}_c(\theta) = \{q \in L_\infty(P_\theta) \,|\, E_\theta\, q = 0,\ \inf_{P_\theta} q \geq -1\}$$

Neighborhoods in S4 Classes



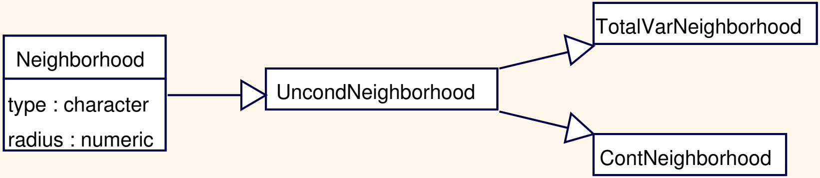
- may be of fixed size or
- of shrinking radius; i.e., $r \rightsquigarrow r/\sqrt{n}$

Neighborhoods in S4 Classes



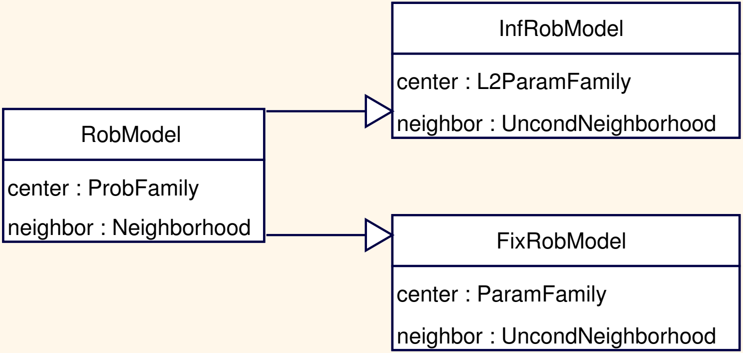
- may be of fixed size or
- of shrinking radius; i.e., $r \rightsquigarrow r/\sqrt{n}$

Neighborhoods in S4 Classes



- may be of fixed size or
- of shrinking radius; i.e., $r \rightsquigarrow r/\sqrt{n}$

Robust Models in S4 Classes



Outline

Asymptotic Mean Square Error Problem

Choosing quadratic loss, one obtains for fixed $r \in (0, \infty)$

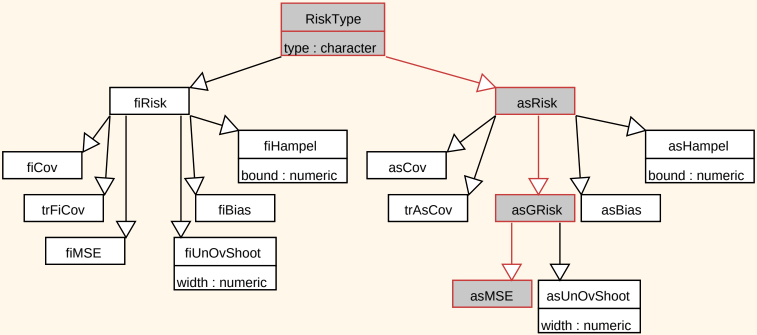
$$\mathit{maxMSE}_{\theta}(\eta_\theta, r) = E_\theta|\eta_\theta|^2 + r^2\omega_{c,\theta}(\eta_\theta)^2 = \min!$$

with $\eta_\theta \in \Psi_2^D(\theta)$ and

$$\omega_{c,\theta}(\eta_\theta) = \sup_{P_\theta} |\eta_\theta|$$

by Proposition 5.3.3 (a) of Rieder (1994).

Risks in S4 Classes



Unique MSE Solution

Theorem 5.5.7 (b), Rieder (1994)

$$\tilde{\eta}_\theta = (A_\theta \Lambda_\theta - a_\theta) w \qquad w = \min\left\{1, \frac{b_\theta}{|A_\theta \Lambda_\theta - a_\theta|}\right\}$$

with Lagrange multipliers A_θ, a_θ and b_θ determined by

$$0 = E_\theta (\Lambda_\theta - z_\theta) w \qquad a_\theta = A_\theta z_\theta$$

$$\mathbb{I}_k = A_\theta E_\theta (\Lambda_\theta - z_\theta)(\Lambda_\theta - z_\theta)^\tau w$$

$$r^2 b_\theta = E_\theta \left(|A_\theta \Lambda_\theta - a_\theta| - b_\theta\right)_+$$

Outline

Properties of MSE Solution I

- Classical Cramér-Rao bound:

$$Cov(\eta_\theta) \succeq Cov(\hat{\psi}_\theta) = \mathcal{I}_\theta^{-1} \qquad \text{where } \hat{\psi}_\theta = \mathcal{I}_\theta^{-1}\Lambda_\theta$$

Hence

$$MSE\,(\eta_\theta) = tr\, Cov(\eta_\theta) \geq tr\, Cov(\hat{\psi}_\theta) = tr\, \mathcal{I}_\theta^{-1} = MSE\,(\hat{\psi}_\theta)$$

- Generalized by Proposition 2.1.1, Kohl (2005):

$$maxMSE(\eta_\theta, r) \geq maxMSE(\tilde{\eta}_\theta, r) = tr\, A_\theta$$

Properties of MSE Solution I

- Classical Cramér-Rao bound:

$$Cov(\eta_\theta) \succeq Cov(\hat{\psi}_\theta) = \mathcal{I}_\theta^{-1} \qquad \text{where } \hat{\psi}_\theta = \mathcal{I}_\theta^{-1}\Lambda_\theta$$

Hence

$$MSE\,(\eta_\theta) = tr\,Cov(\eta_\theta) \geq tr\,Cov(\hat{\psi}_\theta) = tr\,\mathcal{I}_\theta^{-1} = MSE\,(\hat{\psi}_\theta)$$

- Generalized by Proposition 2.1.1, Kohl (2005):

$$maxMSE(\eta_\theta, r) \geq maxMSE(\tilde{\eta}_\theta, r) = tr\,A_\theta$$

Properties of MSE Solution II

- Lagrange multipliers are bounded (Kohl (2005))
- Lagrange multipliers are not necessarily unique (Rieder (1994), Kohl (2005))
- Lagrange multipliers are continuous w.r.t. radius $r \in (0, \infty)$ (Kohl (2005))
- Lagrange multipliers are continuous w.r.t. parameter $\theta \in \Theta$ (Kohl (2005))

Properties of MSE Solution II

- Lagrange multipliers are bounded (Kohl (2005))
- Lagrange multipliers are not necessarily unique (Rieder (1994), Kohl (2005))
- Lagrange multipliers are continuous w.r.t. radius $r \in (0, \infty)$ (Kohl (2005))
- Lagrange multipliers are continuous w.r.t. parameter $\theta \in \Theta$ (Kohl (2005))

Properties of MSE Solution II

- Lagrange multipliers are bounded (Kohl (2005))
- Lagrange multipliers are not necessarily unique (Rieder (1994), Kohl (2005))
- Lagrange multipliers are continuous w.r.t. radius $r \in (0, \infty)$ (Kohl (2005))
- Lagrange multipliers are continuous w.r.t. parameter $\theta \in \Theta$ (Kohl (2005))

Properties of MSE Solution II

- Lagrange multipliers are bounded (Kohl (2005))
- Lagrange multipliers are not necessarily unique (Rieder (1994), Kohl (2005))
- Lagrange multipliers are continuous w.r.t. radius $r \in (0, \infty)$ (Kohl (2005))
- Lagrange multipliers are continuous w.r.t. parameter $\theta \in \Theta$ (Kohl (2005))

Outline

MSE–Inefficiency

Definition

The MSE–inefficiency *of* $\tilde{\eta}_{r_0}$ *w.r.t.* $\tilde{\eta}_r$ *is defined as*

$$relMSE(\tilde{\eta}_{r_0}, r) = \frac{maxMSE(\tilde{\eta}_{r_0}, r)}{maxMSE(\tilde{\eta}_r, r)}$$

where

$$maxMSE(\tilde{\eta}_{r_0}, r) = E\,|\tilde{\eta}_{r_0}|^2 + r^2\omega_*(\tilde{\eta}_{r_0})^2 \qquad * = c, v$$

Remark

The MSE–inefficiency was first considered and numerically evaluated in Rieder et al. (2001).

MSE–Inefficiency

Definition

The MSE–inefficiency *of* $\tilde{\eta}_{r_0}$ *w.r.t.* $\tilde{\eta}_r$ *is defined as*

$$relMSE(\tilde{\eta}_{r_0}, r) = \frac{maxMSE(\tilde{\eta}_{r_0}, r)}{maxMSE(\tilde{\eta}_r, r)}$$

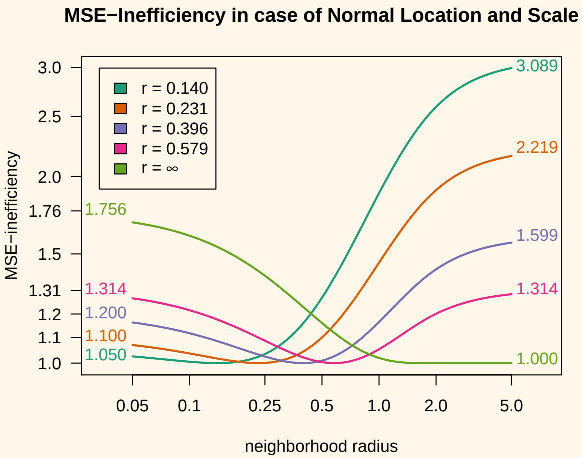
where

$$maxMSE(\tilde{\eta}_{r_0}, r) = E\,|\tilde{\eta}_{r_0}|^2 + r^2 \omega_*(\tilde{\eta}_{r_0})^2 \qquad * = c, v$$

Remark

The MSE–inefficiency was first considered and numerically evaluated in Rieder et al. (2001).

Maximum MSE–Inefficiency



Outline

One-Step Estimator

The one-step estimator $S = (S_n)$ is defined as

$$S_n = \hat{\theta}_n + \frac{1}{n}\sum_{i=1}^{n} \psi_{n,\hat{\theta}_n(y_1,\ldots,y_n)}(y_i)$$

where $\hat{\theta}_n$ is an appropriate initial estimate.

- works in case of Exponential families of full rank (cf. Lemma 2.3.6, Kohl (2005))
- initial estimator: Kolmogorov(-Smirnov) minimum distance estimator (cf. Theorem 6.3.7, Rieder (1994))

One-Step Estimator

The one-step estimator $S = (S_n)$ is defined as

$$S_n = \hat{\theta}_n + \frac{1}{n}\sum_{i=1}^{n} \psi_{n,\hat{\theta}_n(y_1,\ldots,y_n)}(y_i)$$

where $\hat{\theta}_n$ is an appropriate initial estimate.

- works in case of Exponential families of full rank (cf. Lemma 2.3.6, Kohl (2005))
- initial estimator: Kolmogorov(-Smirnov) minimum distance estimator (cf. Theorem 6.3.7, Rieder (1994))

One-Step Estimator

The one-step estimator $S = (S_n)$ is defined as

$$S_n = \hat{\theta}_n + \frac{1}{n}\sum_{i=1}^{n} \psi_{n,\hat{\theta}_n(y_1,\ldots,y_n)}(y_i)$$

where $\hat{\theta}_n$ is an appropriate initial estimate.

- works in case of Exponential families of full rank (cf. Lemma 2.3.6, Kohl (2005))
- initial estimator: Kolmogorov(-Smirnov) minimum distance estimator (cf. Theorem 6.3.7, Rieder (1994))

Optimally Robust Estimation – a Proposal

0. Choose an appropriate parametric family.
1. Choose and evaluate an appropriate initial estimate; e.g., Kolmogorov(–Smirnov) MD estimator.
2. Depending on the quality of the data, try to find a rough estimate for the amount $\varepsilon \in [0, 1]$ of gross errors such that $\varepsilon \in [\underline{\varepsilon}, \overline{\varepsilon}]$.
3. Estimate the parameter of interest by means of the corresponding radius-minimax estimator using the one-step construction.

Via our R package `ROptEst` this proposal so far works for all(!) L_2-differentiable parametric families which are based on a univariate distribution.

Optimally Robust Estimation – a Proposal

0. Choose an appropriate parametric family.
1. Choose and evaluate an appropriate initial estimate; e.g., Kolmogorov(–Smirnov) MD estimator.
2. Depending on the quality of the data, try to find a rough estimate for the amount $\varepsilon \in [0, 1]$ of gross errors such that $\varepsilon \in [\underline{\varepsilon}, \overline{\varepsilon}]$.
3. Estimate the parameter of interest by means of the corresponding radius-minimax estimator using the one-step construction.

Via our R package `ROptEst` this proposal so far works for all(!) L_2-differentiable parametric families which are based on a univariate distribution.

Optimally Robust Estimation – a Proposal

0. Choose an appropriate parametric family.
1. Choose and evaluate an appropriate initial estimate; e.g., Kolmogorov(–Smirnov) MD estimator.
2. Depending on the quality of the data, try to find a rough estimate for the amount $\varepsilon \in [0, 1]$ of gross errors such that $\varepsilon \in [\underline{\varepsilon}, \overline{\varepsilon}]$.
3. Estimate the parameter of interest by means of the corresponding radius-minimax estimator using the one-step construction.

Via our R package `ROptEst` this proposal so far works for all(!) L_2-differentiable parametric families which are based on a univariate distribution.

Optimally Robust Estimation – a Proposal

0. Choose an appropriate parametric family.
1. Choose and evaluate an appropriate initial estimate; e.g., Kolmogorov(–Smirnov) MD estimator.
2. Depending on the quality of the data, try to find a rough estimate for the amount $\varepsilon \in [0, 1]$ of gross errors such that $\varepsilon \in [\underline{\varepsilon}, \overline{\varepsilon}]$.
3. Estimate the parameter of interest by means of the corresponding radius-minimax estimator using the one-step construction.

Via our R package `ROptEst` this proposal so far works for all(!) L_2-differentiable parametric families which are based on a univariate distribution.

Optimally Robust Estimation – a Proposal

0. Choose an appropriate parametric family.
1. Choose and evaluate an appropriate initial estimate; e.g., Kolmogorov(–Smirnov) MD estimator.
2. Depending on the quality of the data, try to find a rough estimate for the amount $\varepsilon \in [0, 1]$ of gross errors such that $\varepsilon \in [\underline{\varepsilon}, \overline{\varepsilon}]$.
3. Estimate the parameter of interest by means of the corresponding radius-minimax estimator using the one-step construction.

Via our R package `ROptEst` this proposal so far works for all(!) L_2-differentiable parametric families which are based on a univariate distribution.

Normal Location and Scale

- Ideal Model:

$$\mathcal{P} = \left\{ P_\theta = \mathcal{N}(\mu, \sigma^2) \,\middle|\, \theta = (\mu, \sigma)^\tau \in \mathbb{R} \times (0, \infty) \right\}$$

- L_2 derivative and Fisher information at $\theta = (\mu, \sigma)^\tau$:

$$\Lambda_\theta(y) = \frac{1}{\sigma} \begin{pmatrix} (y-\mu)/\sigma \\ (y-\mu)^2/\sigma^2 - 1 \end{pmatrix} \qquad \mathcal{I}_\theta = \frac{1}{\sigma^2} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

- Invariant under the group of transformations

$$g_\theta(u) = \sigma u + \mu$$

i.e., $P_\theta = g_\theta(P_{\theta_0})$ where $\theta_0 = (0, 1)^\tau$

Normal Location and Scale

- Ideal Model:

$$\mathcal{P} = \left\{ P_\theta = \mathcal{N}(\mu, \sigma^2) \,\middle|\, \theta = (\mu, \sigma)^\tau \in \mathbb{R} \times (0, \infty) \right\}$$

- L_2 derivative and Fisher information at $\theta = (\mu, \sigma)^\tau$:

$$\Lambda_\theta(y) = \frac{1}{\sigma} \begin{pmatrix} (y-\mu)/\sigma \\ (y-\mu)^2/\sigma^2 - 1 \end{pmatrix} \qquad \mathcal{I}_\theta = \frac{1}{\sigma^2} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

- Invariant under the group of transformations

$$g_\theta(u) = \sigma u + \mu$$

i.e., $P_\theta = g_\theta(P_{\theta_0})$ where $\theta_0 = (0, 1)^\tau$

Normal Location and Scale

- Ideal Model:

$$\mathcal{P} = \left\{ P_\theta = \mathcal{N}(\mu, \sigma^2) \,\middle|\, \theta = (\mu, \sigma)^\tau \in \mathbb{R} \times (0, \infty) \right\}$$

- L_2 derivative and Fisher information at $\theta = (\mu, \sigma)^\tau$:

$$\Lambda_\theta(y) = \frac{1}{\sigma} \begin{pmatrix} (y-\mu)/\sigma \\ (y-\mu)^2/\sigma^2 - 1 \end{pmatrix} \qquad \mathcal{I}_\theta = \frac{1}{\sigma^2} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

- Invariant under the group of transformations

$$g_\theta(u) = \sigma u + \mu$$

i.e., $P_\theta = g_\theta(P_{\theta_0})$ where $\theta_0 = (0, 1)^\tau$

Outline

Poisson and Normal Approximation of Binomial Distribution

- Poisson approximation: Let $\lambda = mp$ (p small, m large). Then,

$$Binom\,(m, p) \approx Pois\,(\lambda)$$

- Normal approximation: ($mp(1-p) \geq 9$)

$$Binom\,(m, p) \approx \mathcal{N}\,(mp, mp(1-p))$$

Poisson and Normal Approximation of Binomial Distribution

- Poisson approximation: Let $\lambda = mp$ (p small, m large). Then,

$$\mathit{Binom}\,(m, p) \approx \mathit{Pois}\,(\lambda)$$

- Normal approximation: ($mp(1-p) \geq 9$)

$$\mathit{Binom}\,(m, p) \approx \mathcal{N}\left(mp, mp(1-p)\right)$$

Setup

- Let
$$\mathcal{P}_\nu = \{P_{\nu,\theta} \,|\, \theta \in \Theta\} \subset \mathcal{M}_1(\mathcal{A}_\nu) \qquad (\nu \in \mathbb{N}_0)$$
be a sequence of L_2-differentiable parametric families.
- In addition, let $r_n := r/\sqrt{n}$ and consider
$$U_{c,\nu}(\theta, r_n) = \big\{(1 - r_n)_+ P_{\nu,\theta} + (1 \wedge r_n)\, Q_\nu \,\big|\, Q_\nu \in \mathcal{M}_1(\mathcal{A}_\nu)\big\}$$
- Question:
$$\mathcal{P}_\nu \approx \mathcal{P}_0 \qquad \text{or even} \qquad U_{c,\nu}(\theta, r_n) \approx U_{c,0}(\theta, r_n)$$

Setup

- Let
$$\mathcal{P}_\nu = \{P_{\nu,\theta} \,|\, \theta \in \Theta\} \subset \mathcal{M}_1(\mathcal{A}_\nu) \qquad (\nu \in \mathbb{N}_0)$$
be a sequence of L_2-differentiable parametric families.

- In addition, let $r_n := r/\sqrt{n}$ and consider
$$U_{c,\nu}(\theta, r_n) = \left\{(1 - r_n)_+ P_{\nu,\theta} + (1 \wedge r_n)\, Q_\nu \,\middle|\, Q_\nu \in \mathcal{M}_1(\mathcal{A}_\nu)\right\}$$

- Question:
$$\mathcal{P}_\nu \approx \mathcal{P}_0 \qquad \text{or even} \qquad U_{c,\nu}(\theta, r_n) \approx U_{c,0}(\theta, r_n)$$

Setup

- Let
$$\mathcal{P}_\nu = \{P_{\nu,\theta} \,|\, \theta \in \Theta\} \subset \mathcal{M}_1(\mathcal{A}_\nu) \qquad (\nu \in \mathbb{N}_0)$$
be a sequence of L_2-differentiable parametric families.
- In addition, let $r_n := r/\sqrt{n}$ and consider
$$U_{c,\nu}(\theta, r_n) = \Big\{(1 - r_n)_+ P_{\nu,\theta} + (1 \wedge r_n)\, Q_\nu \,\Big|\, Q_\nu \in \mathcal{M}_1(\mathcal{A}_\nu)\Big\}$$
- Question:
$$\mathcal{P}_\nu \approx \mathcal{P}_0 \qquad \text{or even} \qquad U_{c,\nu}(\theta, r_n) \approx U_{c,0}(\theta, r_n)$$

Convergence of Experiments

Definition 2.2.1, Le Cam and Lo Yang (2000)

The deficiency $\delta(\mathcal{P}_\nu, \mathcal{P}_0)$ of $\mathcal{P}_\nu$ w.r.t. $\mathcal{P}_0$ is the smallest number $\delta \in [0,1]$ such that for every arbitrary loss function W with $0 \le W \le 1$ and every risk function r_2 there is an risk function r_1 such that $r_1(P_{\nu,\theta}, W) \le r_2(P_{0,\theta}, W) + \delta$ for all $\theta \in \Theta$.

Theorem 2.4.1, Kohl (2005)

Assume the laws of the corresponding L_2 derivatives as well as the trace of the corresponding Fisher information converge under suitable standardizations. Then, the suitable standardized maximum asymptotic MSE of the corresponding optimally robust estimators converges.

Convergence of Experiments

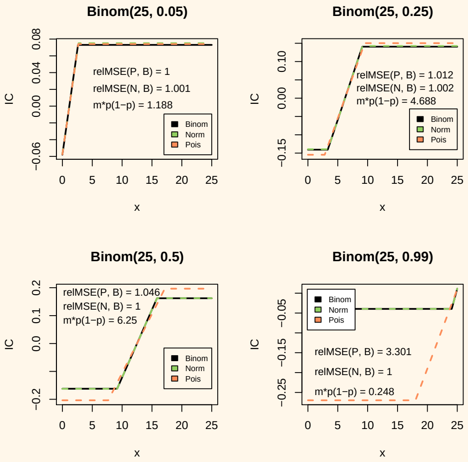
Definition 2.2.1, Le Cam and Lo Yang (2000)

The deficiency $\delta(\mathcal{P}_\nu, \mathcal{P}_0)$ of $\mathcal{P}_\nu$ w.r.t. $\mathcal{P}_0$ is the smallest number $\delta \in [0,1]$ such that for every arbitrary loss function W with $0 \le W \le 1$ and every risk function r_2 there is an risk function r_1 such that $r_1(P_{\nu,\theta}, W) \le r_2(P_{0,\theta}, W) + \delta$ for all $\theta \in \Theta$.

Theorem 2.4.1, Kohl (2005)

Assume the laws of the corresponding L_2 derivatives as well as the trace of the corresponding Fisher information converge under suitable standardizations. Then, the suitable standardized maximum asymptotic MSE of the corresponding optimally robust estimators converges.

Approximation of ICs for $r = 0.25$



Bibliography

- M. Kohl.
  Numerical Contributions to the Asymptotic Theory of Robustness. Dissertation, University of Bayreuth, 2005.
- L. Le Cam and G. Lo Yang.
  Asymptotics in Statistics. Springer, 2000.
- H. Rieder.
  Robust Asymptotic Statistics. Springer, 1994.
- H. Rieder, M. Kohl and P. Ruckdeschel.
  The Costs of not Knowing the Radius. Submitted. Appeared as discussion paper Nr. 81. SFB 373, Humboldt University, Berlin, 2001.