Optimal Estimation of Multidimensional Normal Means with Unknown Variances

Xu Han

Department of Statistics
The Wharton School
University of Pennsylvania

Joint work with Professor Lawrence D. Brown

Yale University
May 16, 2009

The Normal Mean Estimation Problem

$$X \sim N_p(\theta, \eta^{-2} I) \qquad W \sim \eta^{-2}\chi^2_m$$

- θ and η unknown;
- W independent of X;
- $\eta = 1/\sigma$ is called precision;
- Squared Error Loss: $L(\theta, \eta^2; \delta(x, w)) = ||\delta(x, w) - \theta||^2\eta^2$;
- Risk function: $R(\theta, \eta^2; \delta(X, W)) = E_{\theta,\eta^2} L(\theta, \eta^2; \delta(X, W))$.

Use procedure $\delta(x, w)$ to estimate θ.

Bayes Estimator:

$\delta_G(x, w) = \frac{E(\theta\eta^2|x,w)}{E(\eta^2|x,w)}$ with prior G on θ, η.

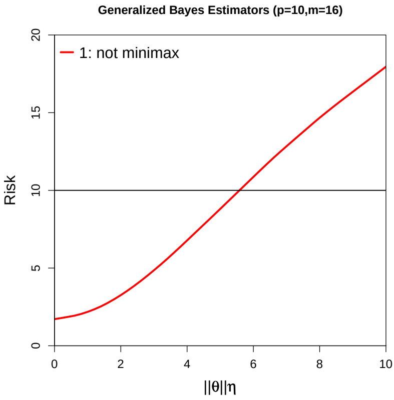
- G proper, proper Bayes (Strawderman (1973));
- G improper, generalized Bayes (Maruyama & Strawderman (2005)).

Benchmark for Evaluation of Estimators: Admissibility and Minimaxity

δ_1 dominates δ_2 if $R(\delta_1; \theta, \eta^2) \leq R(\delta_2; \theta, \eta^2)$ for all θ, η^2, and with strict $<$ for some θ, η^2.

Benchmark:

- Admissible estimator: no other estimators can dominate it;
- Minimax estimator: if it dominates MLE x.
- This talk will focus on admissible estimators.

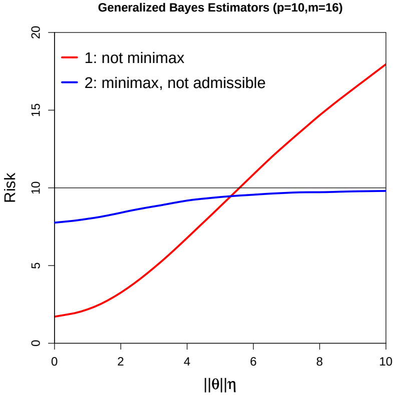


Benchmark for Evaluation of Estimators: Admissibility and Minimaxity

δ_1 dominates δ_2 if $R(\delta_1; \theta, \eta^2) \leq R(\delta_2; \theta, \eta^2)$ for all θ, η^2, and with strict $<$ for some θ, η^2.

Benchmark:

- Admissible estimator: no other estimators can dominate it;
- Minimax estimator: if it dominates MLE x.
- This talk will focus on admissible estimators.

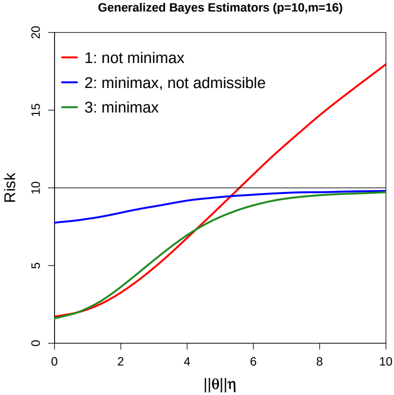


Benchmark for Evaluation of Estimators: Admissibility and Minimaxity

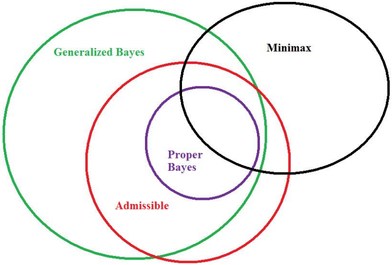
δ_1 dominates δ_2 if $R(\delta_1; \theta, \eta^2) \leq R(\delta_2; \theta, \eta^2)$ for all θ, η^2, and with strict $<$ for some θ, η^2.

Benchmark:

- Admissible estimator: no other estimators can dominate it;
- Minimax estimator: if it dominates MLE x.
- This talk will focus on admissible estimators.



Unknown Variance Case



Admissibility:

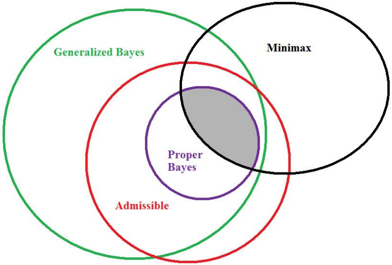
- Strawderman (1973): proper Bayes admissible and minimax estimators for $p \geq 5$.

Minimaxity:

- Maruyama & Strawderman (2005), Wells & Zhou (2007): minimax generalized Bayes estimators for $p \geq 3$.

- Admissibility of generalized Bayes estimators is an important and open problem.
- Goal: Provide sufficient conditions for prior density functions such that δ_G is admissible.

Unknown Variance Case



Admissibility:

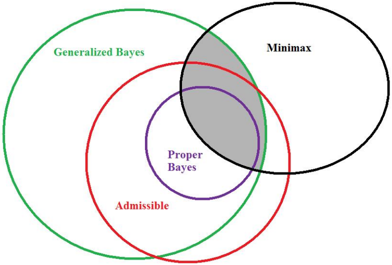
- Strawderman (1973): proper Bayes admissible and minimax estimators for $p \geq 5$.

Minimaxity:

- Maruyama & Strawderman (2005), Wells & Zhou (2007): minimax generalized Bayes estimators for $p \geq 3$.

- Admissibility of generalized Bayes estimators is an important and open problem.
- Goal: Provide sufficient conditions for prior density functions such that δ_G is admissible.

Unknown Variance Case



Admissibility:

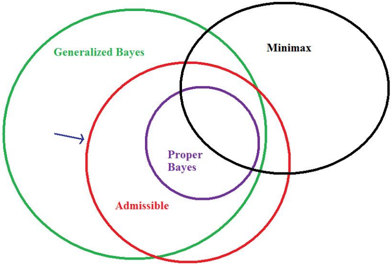
- Strawderman (1973): proper Bayes admissible and minimax estimators for $p \geq 5$.

Minimaxity:

- Maruyama & Strawderman (2005), Wells & Zhou (2007): minimax generalized Bayes estimators for $p \geq 3$.

- Admissibility of generalized Bayes estimators is an important and open problem.
- Goal: Provide sufficient conditions for prior density functions such that δ_G is admissible.

Unknown Variance Case



Admissibility:

- Strawderman (1973): proper Bayes admissible and minimax estimators for $p \geq 5$.

Minimaxity:

- Maruyama & Strawderman (2005), Wells & Zhou (2007): minimax generalized Bayes estimators for $p \geq 3$.

- Admissibility of generalized Bayes estimators is an important and open problem.
- Goal: Provide sufficient conditions for prior density functions such that δ_G is admissible.

Admissibility Theorem

Hierarchical Bayes Model:

$$\theta|\eta^2 \sim g(\theta|\eta^2) \qquad \eta^2 \sim \pi(\eta^2)$$

Admissibility Theorem:

Let $X \sim N_p(\theta, \eta^{-2} I_p), W \sim \eta^{-2}\chi^2_m$, where X and W are independent. If $g(\theta|\eta^2)$ and $\pi(\eta^2)$ satisfy Condition 1, 2 and 3, then δ_G is admissible.

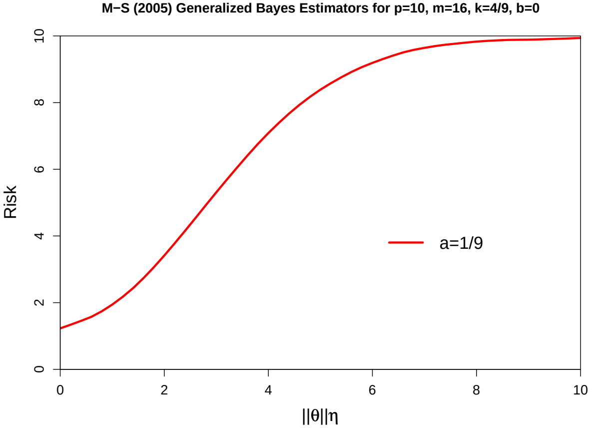
Application in Maruyama-Strawderman (2005) Hierarchical Bayes Model:

$$\begin{aligned} \theta|\nu,\eta &\sim N_p(0, \nu\eta^{-2} I) \\ h(\nu) &\propto \nu^b (1+\nu)^{-a-b-2} \qquad \nu \geq 0 \\ \pi(\eta) &\propto \eta^{-2k} \qquad \eta \geq 0 \end{aligned}$$

Corollary:

When $-a + 1/2 < k < 1/2$, and $b \geq 0$, δ_G is admissible.

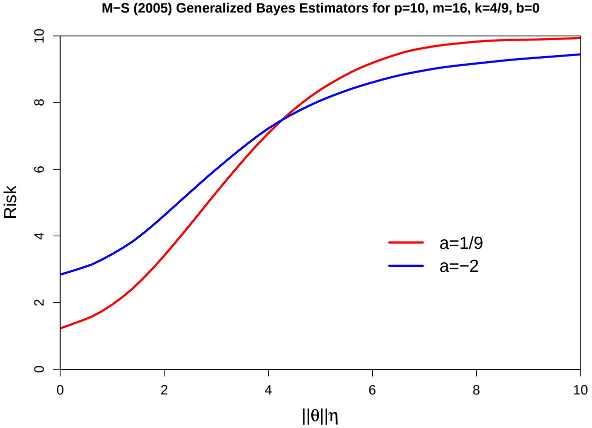
Performance of Generalized Bayes Estimators



- a=1/9 admissible
- a=-2 outside admissible range
- a=-5 far outside admissible range

$$\begin{aligned} X &\sim N_p(\theta, \eta^{-2} I) \qquad W \sim \eta^{-2}\chi^2_m \\ \theta|\nu,\eta &\sim N_p(0, \nu\eta^{-2} I) \\ h(\nu) &\propto \nu^b (1+\nu)^{-a-b-2} \quad \pi(\eta) \propto \eta^{-2k} \end{aligned}$$

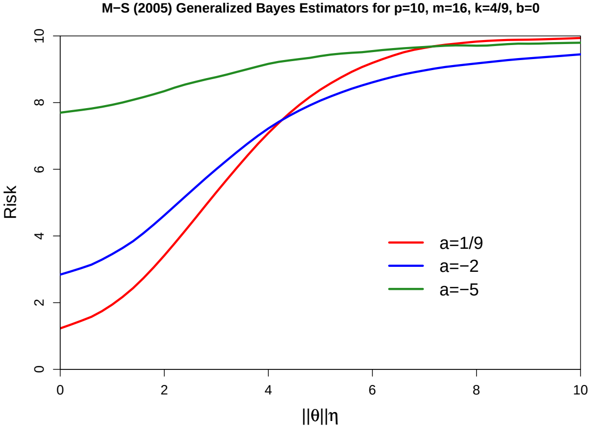
Performance of Generalized Bayes Estimators



- a=1/9 admissible
- a=-2 outside admissible range
- a=-5 far outside admissible range

$$
\begin{aligned}
X &\sim N_p(\theta, \eta^{-2} I) \qquad W \sim \eta^{-2}\chi^2_m \\
\theta|\nu,\eta &\sim N_p(0, \nu\eta^{-2} I) \\
h(\nu) &\propto \nu^b (1+\nu)^{-a-b-2} \qquad \pi(\eta) \propto \eta^{-2k}
\end{aligned}
$$

Performance of Generalized Bayes Estimators



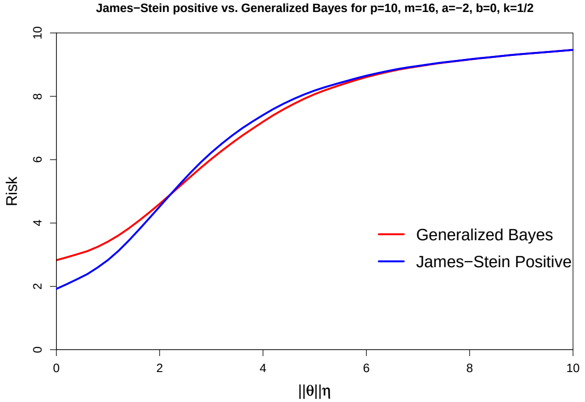
- a=1/9 admissible
- a=-2 outside admissible range
- a=-5 far outside admissible range

$$
\begin{aligned}
X &\sim N_p(\theta, \eta^{-2} I) \qquad W \sim \eta^{-2}\chi^2_m \\
\theta | \nu, \eta &\sim N_p(0, \nu\eta^{-2} I) \\
h(\nu) &\propto \nu^b (1+\nu)^{-a-b-2} \qquad \pi(\eta) \propto \eta^{-2k}
\end{aligned}
$$

Possible Improvement

Conjecture: When $-a - 3/2 < k < 1/2, b \geq 0$, $\delta_G(x, w)$ is admissible. This is the best sufficient condition for admissibility.

- $a = -2, b = 0, k = 1/2$ are on the boundary of sufficient conditions for admissibility in conjecture.



Summary

- Admissible estimation of normal mean with unknown variances is important problem;
- Admissibility and minimaxity provide benchmark for shrinkage estimators;
- Admissibility is powerful tool to select hierarchical priors;
- Admissibility Theorem gives sufficient conditions for admissibility;
- Corollary describes a subset of M-S hierarchical Bayes estimators that are minimax and admissible.

Thank You!