Optimally Bounding the GES (Hampel's Problem)

Let $\hat{\mu}_n$ be a (smooth) location M-estimate based on the monotone location score function ψ and the consistent dispersion estimate $\hat{\sigma}_n$. We have shown before that the asymptotic variance of $\hat{\mu}_n$ at the central normal model $H_0 = N\left(\mu, \sigma^2\right)$ is given by

$$AV(\psi, H_0) \quad = \quad \sigma^2 \frac{E_{H_0}\left\{\psi^2(\frac{y-\mu}{\sigma})\right\}}{\left[E_{H_0}\left\{\psi(\frac{y-\mu}{\sigma})\left(\frac{y-\mu}{\sigma}\right)\right\}\right]^2}. \tag{1}$$

We have also shown before that the gross error sensitivity of $\hat{\mu}_n$ at the central normal model $N\left(\mu, \sigma^2\right)$ is given by

$$\gamma(\psi, H_0) \quad = \quad \sigma \frac{\psi(\infty)}{E_{H_0}\left\{\psi(\frac{y-\mu}{\sigma})\left(\frac{y-\mu}{\sigma}\right)\right\}}. \tag{2}$$

In his Ph.D. dissertation (completed in 1967), Hampel posed and solved the following problem.

Hampel's Optimality Problem: Consider the class $\mathcal{C}$ of all location M-estimates with non-decreasing and bounded score functions ψ. We wish to find the *most efficient robust estimate* within this class. Suppose that we will measure efficiency by the asymptotic variance (at the target model H_0) and robustness by the gross error sensitivity. Therefore, we wish to choose the location M-estimate within $\mathcal{C}$ with the smallest asymptotic variance $AV(\psi, H_0)$ among all estimates in $\mathcal{C}$ with relatively small gross-error sensitivity $\gamma(\psi, H_0)$. Hampel's optimality problem can be mathematically stated as follows: find $\psi^* \in \mathcal{C}$ such that

$$\psi^* \quad = \quad \arg \min_{\gamma(\psi, H_0) \leq K} AV(\psi, H_0) \tag{3}$$

The change of variables $z = (y-\mu)/\sigma$ in (1) and (2) gives

$$AV(\psi, H_0) \quad = \quad \sigma^2 \frac{E_{H_0}\left\{\psi^2(\frac{y-\mu}{\sigma})\right\}}{\left[E_{H_0}\left\{\psi(\frac{y-\mu}{\sigma})\left(\frac{y-\mu}{\sigma}\right)\right\}\right]^2} = \sigma^2 \frac{E_\Phi\left\{\psi^2(z)\right\}}{\left[E_\Phi\left\{\psi(z)\,(z)\right\}\right]^2} \tag{4}$$

and

$$\gamma(\psi, H_0) \quad = \quad \sigma \frac{\psi(\infty)}{E_{H_0}\left\{\psi(\frac{y-\mu}{\sigma})\left(\frac{y-\mu}{\sigma}\right)\right\}} = \sigma \frac{\psi(\infty)}{E_\Phi\left\{\psi(z) z\right\}}. \tag{5}$$

Therefore, except for the fixed scaling factors (σ^2 and σ) the asymptotic variance and the GES of M-estimates essentially depend on the shape of the distribution H_0 and not on the particular values of its location and dispersion parameters. Therefore, without loss of generality we can assume that the fixed location and dispersion parameters are $\mu = 0$ and $\sigma = 1$ (and so $H_0 = \Phi$). Moreover, since (4) and (5) are invariant under multiplication of ψ by a non-zero constant, we can assume without loss of generality that

$$E_\Phi \{\psi(y)y\} \quad = \quad 1. \tag{6}$$

Therefore, problem (3) can be re-stated as follows.

$$\begin{array}{lll} \text{Minimize} & (\text{in } \psi) & E_\Phi \left\{\psi^2(z)\right\} \\ \text{Subject to} & & \\ & & (\text{i}) \quad E_\Phi \{\psi(y)y\} = 1. \\ \text{and} & & \\ & & (\text{ii}) \quad \psi(\infty) \le K \end{array}$$

For future reference let

$$\mathcal{C}_K \quad = \quad \{\psi \in \mathcal{C} : (\text{i}) \text{ and } (\text{ii}) \text{ above hold}\}. \tag{7}$$

We have the following result:

Theorem 1 *Let $H_0 = N\left(\mu, \sigma^2\right)$ and consider the family of Huber's score functions:*

$$\psi_c(y) = \begin{cases} y/[2\Phi(c) - 1] & |y| \le c \\ \\ c \; sign(y)/[2\Phi(c) - 1] & |y| > c. \end{cases} \tag{8}$$

Then

(a) *If $K < 1/[2\varphi(0)]$, then $\mathcal{C}_K = \phi$ (the empty set). In other words,*

$$\gamma(\psi, \Phi) \quad \ge \quad \sqrt{\frac{\pi}{2}}, \qquad \textit{for all } \psi. \tag{9}$$

(b) *The median minimizes the gross error sensitivity among all location M-estimates.*

(c) *If* $K > 1/[2\varphi(0)]$, *then there exists a unique constant* $c > 0$ *such that* $\psi_c \in \mathcal{C}_K$, *that is* $\gamma(\psi_c, \Phi) = K$.

Proof: To prove (a) write

$$\begin{aligned}
E_\Phi\left\{\psi(y)y\right\} &= \int_{-\infty}^{\infty} \psi(y)y\varphi(y)dy \\
&= 2\int_0^{\infty} \psi(y)y\varphi(y)dy \\
&\leq 2\psi(\infty)\int_0^{\infty} y\varphi(y)dy = -2\psi(\infty)\int_0^{\infty} \varphi'(y)dy \\
&= -2\psi(\infty)[\varphi(\infty) - \varphi(0)] = 2\psi(\infty)\varphi(0) \\
&= 2\psi(\infty)/\sqrt{2\pi}.
\end{aligned}$$

Therefore,

$$\gamma(\psi, \Phi) = \frac{\psi(\infty)}{E_\Phi\left\{\psi(y)y\right\}} \geq \frac{\sqrt{2\pi}}{2} = \sqrt{\frac{\pi}{2}}.$$

Part (b) follows directly from (a) and the fact that in the case of the median $\psi(y) = \text{sign}(y)$

$$\begin{aligned}
\gamma(\text{sign}, \Phi) &= \frac{\text{sign}(\infty)}{E_\Phi\left\{\text{sign}(y)\, y\varphi(y)\right\}} \\
&= \frac{1}{\int_{-\infty}^{\infty} \text{sign}(y)\, y\varphi(y)dy} \\
&= \left[\int_{-\infty}^{\infty} |y|\,\varphi(y)dy\right]^{-1} \\
&= \left[2\int_0^{\infty} y\varphi(y)dy\right]^{-1} \\
&= \left[-2\int_0^{\infty} \varphi'(y)dy\right]^{-1} \\
&= \left[-2\varphi(\infty) + 2\varphi(0)\right]^{-1} \\
&= 1/\left[2\varphi(0)\right] \\
&= \sqrt{\frac{\pi}{2}}.
\end{aligned}$$

To prove (c) notice that

$$E_\Phi\left\{\psi_c(y)y\right\} = 1$$

and therefore

$$\gamma(\psi_c, \Phi) = \frac{c}{[2\Phi(c) - 1]} \equiv g(c)$$

The function $g(c)$ is continuous, with

$$\lim_{c \to \infty} g\,(c) = \infty$$

and

$$\lim_{c \to 0} g\,(c) \quad = \quad \lim_{c \to 0} \frac{c}{[2\Phi(c) - 1]} = \lim_{c \to 0} \frac{1}{2\varphi(c)} = \sqrt{\pi/2}.$$

This proves the *existence* part of (c). The *uniqueness* part of (c) follows from the fact

$$g'(c) \quad = \quad \frac{2\Phi(c) - 1 - 2c\varphi(c)}{[2\Phi(c) - 1]^2} \geq 0, \qquad \text{for all} \quad c \geq 0 \quad \text{(see Problem 3).}$$

Theorem 2 *Given* $K > 1/[2\varphi(0)]$, *there exists a unique* $c > 0$ *such that* $\gamma(\psi_c, \Phi) = K$ *(see (8)) and*

$$AV(\psi, \Phi) \quad \geq \quad AV(\psi_c, \Phi), \qquad \textit{for all } \psi \in \mathcal{C}_K.$$

Proof: Let $K > 1/[2\varphi(0)]$. By the previous theorem we know there is a unique

$$c \quad = \quad c(K)$$

such

$$\gamma(\psi_c, \Phi) \quad = \quad K.$$

Set

$$\beta \quad = \quad \beta(c) = \frac{1}{[2\Phi(c) - 1]}.$$

Now write

$$\mathcal{J}(\psi) \quad = \quad E_{\Phi}\left\{\psi^2(y)\right\}$$

and

$$\mathcal{I}(\psi) \quad = \quad \int_{-\infty}^{\infty} [\psi(y) - \beta y]^2 \varphi(y) dy.$$

For all $\psi \in \mathcal{C}_K$,

$$\mathcal{I}(\psi) \quad = \quad \int_{-\infty}^{\infty} [\psi^2(y) + \beta^2 y^2 - 2\beta y \psi(y)]\varphi(y) dy = \mathcal{J}(\psi) + \beta^2 - 2\beta. \quad (10)$$

Therefore, ψ^* minimizes $\mathcal{J}(\psi)$ over $\mathcal{C}_K$ if and only if ψ^* minimizes $\mathcal{I}(\psi)$ over $\mathcal{C}_K$.

By symmetry,

$$\mathcal{I}(\psi) \quad = \quad 2\int_{0}^{\infty} [\psi(y) - \beta y]^2 \varphi(y) dy$$

and so it suffices to study the squared difference $[\psi(y) - \beta y]^2$ for $y \geq 0$.

Clearly, since $\psi \in \mathcal{C}_K$, for $y \leq c$ we have

$$[\psi(y) - \beta y]^2 \quad \geq \quad 0 = [\psi_c(y) - \beta y]^2 \quad \text{for} \quad y \leq c. \quad (11)$$

On the other hand, since $\psi \in \mathcal{C}_K$, for $y > c$ we have

$$0 \quad \leq \quad \psi(y) \leq K = \psi_c(y), \quad \text{for} \quad y > c$$

and so

$$[\psi(y) - \beta y]^2 \quad \geq \quad [K - \beta y]^2 = [\psi_c(y) - \beta y]^2, \quad \text{for} \quad y > c. \quad (12)$$

From (11) and (12) we have (see Figure 1)

$$\begin{aligned}\mathcal{I}(\psi) &= 2\int_0^\infty [\psi(y) - \beta y]^2 \varphi(y)dy \\ &\geq 2\int_0^\infty [\psi_c(y) - \beta y]^2 \varphi(y)dy = \mathcal{I}(\psi_c) \end{aligned} \tag{13}$$

Finally the theorem follows (i.e. $\psi^*(y) = \psi_c(y)$) from (10) and (13).

By Theorem 2, for each

$$K > \sqrt{\pi/2}$$

there is a unique $c > 0$ such that $\psi_c(y)$ (see Equation (8)) minimizes the asymptotic variance among all ψ functions with

$$\gamma(\psi_c, \Phi) = K.$$

The specific value of c,

$$c = c(K),$$

can be determined from the equation

$$\frac{c}{[2\Phi(c) - 1]} = K. \tag{14}$$

Minimax Variance (Huber's Problem)

Consider the *symmetrical* contamination neighborhood

$$\mathcal{F}_\epsilon^* \quad = \quad \{H: \quad H(y) = (1-\epsilon)\Phi + \epsilon G, \quad G \textit{is symmetric}\}. \tag{15}$$

The technical advantage of restricting attention to $\mathcal{F}_\epsilon^*$ is that for *any* M-estimate $\hat{\mu}$ we have

$$\hat{\mu}(H) \quad = \quad 0, \quad \text{for all} \quad H \in \mathcal{F}_\epsilon^*.$$

Therefore, we do not need to worry about the problem of *asymptotic bias* when we restricted to $\mathcal{F}_\epsilon^*$. On the other hand, the serious practical limitations imposed by this restriction are obvious.

The original problem considered by Huber was solved assuming that the *dispersion parameter σ is known.* The solution of Huber's problem *when σ is unknown (* and must be robustly estimated) was obtained by Li and Zamar (1991).

Huber's goal was to find the location M-estimate score function ψ^* that minimizes the maximum asymptotic variance over the symmetric contamination family (15). Huber's mathematical approach was very elegant and goes along the following lines.

Let

$$\overline{AV}(\psi) \quad = \quad \sup_{F \in \mathcal{F}_\epsilon^*} AV(\psi, H),$$

be the maximum asymptotic variance over $\mathcal{F}_\epsilon^*$ for the location M-estimate with score function ψ. Robust M-estimates are then expected to have a relatively small and stable AV over the contamination neighborhood. Therefore, an estimate with smaller $\overline{AV}(\psi)$ should be preferred, from a robustness point of view (other things being equal). Then, we wish to consider the minimization problem

$$\inf_\psi \overline{AV}(\psi) \quad = \quad \inf_\psi \sup_{F \in \mathcal{F}_\epsilon^*} AV(\psi, H). \tag{16}$$

Suppose that one can prove (as Huber did) that ψ^* solves (16) if and only if ψ^* solves the *dual problem*

$$\sup_{F \in \mathcal{F}_\epsilon^*} \inf_{\psi} AV(\psi, H). \tag{17}$$

Then, we may wish to solve the "dual" optimality problem (17) instead of the (harder) optimality problem (16).

Let

$$\psi_H(y) = -\frac{h'(y)}{h(y)}$$

be the maximum likelihood score function under H and let $I(H)$ be the corresponding Fisher Information index

$$I(H) = E_H^2 \{\psi_H(y)\} = \int_{-\infty}^{\infty} \left(-\frac{h'(y)}{h(y)}\right)^2 h(y)\, dy.$$

It is not difficult to verify that

$$\begin{aligned}
\sup_{F \in \mathcal{F}_\epsilon^*} \inf_{\psi} AV(\psi, H) &= \sup_{F \in \mathcal{F}_\epsilon^*} AV(\psi_H, H) \\
&= \sup_{F \in \mathcal{F}_\epsilon^*} \frac{1}{I(H)} \\
&= \frac{1}{\inf_{F \in \mathcal{F}_\epsilon^*} I(H)}
\end{aligned}$$

Therefore, the original min-max problem reduces to that of finding the *least favorable distribution,* that is, the distribution with smallest Fisher information in the family (15). Huber did precisely that and found its famous truncated-linear score function when H_0 is Normal.

An Alternative Derivation of Huber's Optimality Result

We will see that in the case of "contamination neighborhoods" Huber's optimality result can be directly obtained using Hampel's result, and vice-versa.

That is, the two problems are mathematically equivalent. This is not too surprising if we notice that the two problems yield the same family of optimal score functions.

The elegant approach of solving the min-max variance problem by first finding the least favorable distribution is not easy to apply to models other than the simple *pure location model* treated by Huber. On the other hand, the approach of attacking the min-max problem directly (brute force approach, so to speak) shows better promise of possible extensions as it has already been successfully applied by Li and Zamar (1991) to the location-dispersion model.

It is easy to show (using a simplified version of proof of Theorem ?? in Section ??) that in the case of *known dispersion* - taken equal to one, without loss of generality - the asymptotic distribution of location M-estimates is normal with mean 0 and variance

$$AV(\psi, H) = \frac{E_H\left\{\psi^2(x)\right\}}{\left[E_H\left\{\psi'(x)\right\}\right]^2}. \tag{18}$$

We have the following result.

Theorem 3 *Suppose that the dispersion parameter σ is known (and then taken equal to one without loss of generality). If ψ is non-decreasing then*

$$\sup_{H\in\mathcal{F}_\epsilon} AV(\psi, H) = \frac{1}{(1-\epsilon)} AV(\psi, \Phi) + \frac{\epsilon}{(1-\epsilon)^2}\gamma^2(\psi, \Phi). \tag{19}$$

where $\gamma(\psi, \Phi)$ is the gross-error sensitivity at the Normal model. Therefore the maximum asymptotic variance under the normal ϵ-contamination family $\mathcal{F}_\epsilon$ is a linear combination of the normal asymptotic variance and gross-error-sensitivity, with coefficients $1/(1-\epsilon)$ and $\epsilon/(1-\epsilon)^2$, respectively.

Proof: Let

$$H = (1-\epsilon) N(0,1) + \epsilon H_1 \tag{20}$$

From (18) and (20) we have

$$AV(\psi, H) = \frac{(1-\epsilon) E_\Phi\left\{\psi^2(x)\right\} + \epsilon E_{H_1}\left\{\psi^2(x)\right\}}{\left[(1-\epsilon) E_\Phi\left\{\psi'(x)\right\} + \epsilon E_{H_1}\psi'(x)\right]^2}.$$

Since ψ is non-decreasing, it is clear that

$$(1-\epsilon) E_{\Phi}\left\{\psi^2(x)\right\} + \epsilon E_{H_1}\left\{\psi^2(x)\right\} \leq (1-\epsilon) E_{\Phi}\left\{\psi^2(x)\right\} + \epsilon\psi^2(\infty) \quad (21)$$

Moreover, since $\psi'(x) \geq 0$ for all x, we can write

$$(1-\epsilon) E_{\Phi}\left\{\psi'(x)\right\} + \epsilon E_{H_1}\left\{\psi'(x)\right\} \geq (1-\epsilon) E_{\Phi}\left\{\psi'(x)\right\} \quad (22)$$

Using (21) and (22) we obtain

$$\begin{aligned}
AV(\psi, H) &= \frac{(1-\epsilon) E_{\Phi}\left\{\psi^2(x)\right\} + \epsilon E_{H_1}\left\{\psi^2(x)\right\}}{\left[(1-\epsilon) E_{\Phi}\left\{\psi'(x)\right\} + \epsilon E_{H_1}\psi'(x)\right]^2} \\
&\leq \frac{(1-\epsilon) E_{\Phi}\left\{\psi^2(x)\right\} + \epsilon\psi^2(\infty)}{\left[(1-\epsilon) E_{\Phi}\left\{\psi'(x)\right\}\right]^2} \\
&= \frac{1}{(1-\epsilon)} \frac{E_{\Phi}\left\{\psi^2(x)\right\}}{\left[E_{\Phi}\left\{\psi'(x)\right\}\right]^2} + \frac{\epsilon}{(1-\epsilon)^2} \frac{\psi^2(\infty)}{\left[E_{\Phi}\left\{\psi'(x)\right\}\right]^2} \\
&= \frac{1}{(1-\epsilon)} AV(\psi, \Phi) + \frac{\epsilon}{(1-\epsilon)^2}\gamma^2(\psi, \Phi), \quad \text{for all } H \in \mathcal{F}_{\epsilon}.
\end{aligned}$$

Therefore,

$$\sup_{H \in \mathcal{F}_{\epsilon}} AV(\psi, H) \leq \frac{1}{(1-\epsilon)} AV(\psi, \Phi) + \frac{\epsilon}{(1-\epsilon)^2}\gamma^2(\psi, \Phi). \quad (23)$$

where $\gamma(\psi, \Phi)$ is the gross-error-sensitivity of the M-estimate at the Normal model.

On the other hand, taking

$$H_y = (1-\epsilon) N(0,1) + \epsilon \; Uniform \; (y-\delta, y+\delta; -y-\delta, -y+\delta)$$

we have that

$$\begin{aligned}\sup_{H\in\mathcal{F}_\epsilon} AV(\psi,H) &\geq \sup_y AV(\psi,H_y) \geq \lim_{y\to\infty} AV(\psi,H_y) \\ &= \frac{1}{(1-\epsilon)} AV(\psi,\Phi) + \frac{\epsilon}{(1-\epsilon)^2}\gamma^2(\psi,\Phi). \qquad (24)\end{aligned}$$

Therefore, (19) follows from (23) and (24).

Theorem 3 shows that the maximum asymptotic variance minimized by Huber is the Lagrangian representation of Hampel's optimality problem. Hampel sets an explicit bound on the gross-error-sensitivity while Huber (in a rather fortuitous way) had the gross-error-sensitivity included as a "penalty" term in the maximum asymptotic variance he minimized.

The following result formally shows that Hampel's and Huber's' problems are equivalent from a mathematical point of view. However, it is very important to understand that the mathematical equivalence of the two problems doesn't imply their statistical equivalence. More precisely, Hampel's problem has a deep statistical meaning while Huber's problem is just a nice mathematical elaboration.

Theorem 4 *(Equivalence between Huber's and Hampel's optimal location results)*

(a) Suppose that ψ^* solves Huber's optimality problem for some $0<\epsilon<1/2$. Then ψ^* solves Hampel's optimality problem with gross-error-sensitivity bound equal to $K=\gamma(\psi^*,\Phi)$.

(b) For each $0<\epsilon<1/2$ the solution ψ^* to Huber's min-max variance problem on $\mathcal{F}_\epsilon$ belongs to the set of score functions ψ_c given by (8). That is,

$$\min_\psi \max_{H\in\mathcal{F}_\epsilon} AV(\psi,H) = \min_{c>0} \max_{H\in\mathcal{F}_\epsilon} AV(\psi_c,H)$$

where the min *on the left side is taken over the family of all monotone score functions ψ.*

Proof: To prove Part (a) notice that, by hypothesis,

$$\max_{H\in\mathcal{F}_\epsilon} AV(\psi^*,H) \quad\leq\quad \max_{H\in\mathcal{F}_\epsilon} AV(\psi,H), \quad \text{for all } \psi$$

Let

$$\begin{aligned}\alpha &= 1/(1-\epsilon) \\ \beta &= \epsilon/(1-\epsilon)^2 .\end{aligned} \tag{25}$$

By Theorem 3 we have

$$\alpha AV(\psi^*, \Phi) + \beta\gamma^2(\psi^*, \Phi) \leq \alpha AV(\psi, \Phi) + \beta\gamma^2(\psi, \Phi), \quad \text{for all } \psi.$$

Hence,

$$\alpha AV(\psi^*, \Phi) + \beta K(c)^2 \leq \alpha AV(\psi, \Phi) + \beta\gamma^2(\psi, \Phi), \quad \text{for all } \psi$$

or equivalently

$$AV(\psi^*, \Phi) \leq AV(\psi, \Phi) + \left[\frac{\beta}{\alpha}\gamma^2(\psi, \Phi) - \frac{\beta}{\alpha}K(c)^2\right], \quad \text{for all } \psi.$$

Therefore

$$AV(\psi^*, \Phi) \leq AV(\psi, \Phi), \quad \text{for all } \psi \text{ with } \gamma(\psi, \Phi) \leq K$$

and Part (a) follows.

To prove part (b), let $0 < \epsilon < 1/2$ be fixed and set (as in (25)) $\alpha = 1/(1-\epsilon)$ and $\beta = \epsilon/(1-\epsilon)^2$.

Denote by ψ_c the solution to Hampel's problem with bound $K(c)$ on the gross-error-sensitivity. Then

$$AV(\psi_c, \Phi) \leq AV(\psi, \Phi) \quad \text{for all } \gamma(\psi, \Phi) \leq K(c).$$

This implies the weaker statement

$$AV(\psi_c, \Phi) \leq AV(\psi, \Phi) \quad \text{for all } \gamma(\psi, \Phi) = K(c). \tag{26}$$

Using (26) we can write,

$$\alpha AV\left(\psi_c,\Phi\right)+\beta K \quad \leq \quad \alpha AV\left(\psi,\Phi\right)+\beta K\left(c\right)^2, \quad \text{for all } \gamma\left(\psi,\Phi\right)=K$$

Hence, by Theorem 3

$$\max_{H\in\mathcal{F}_\epsilon} AV\left(\psi_c,\Phi\right) \quad = \quad \max_{H\in\mathcal{F}_\epsilon} AV\left(\psi,H\right) \qquad \text{for all } \gamma\left(\psi,\Phi\right)=K\left(c\right) \tag{27}$$

Let

$$\begin{aligned} g\left(c\right) &= \alpha AV\left(\psi_c,\Phi\right)+\beta K\left(c\right) \\ &= \frac{1}{1-\epsilon} AV\left(\psi_K^*,\Phi\right)+\frac{\epsilon K}{\left(1-\epsilon\right)^2} \end{aligned} \tag{28}$$

and denote by c^* the minimizer of $g\left(c\right)$, that is,

$$\min_{c>0} g\left(c\right) \quad = \quad g\left(c^*\right) \qquad \text{(see Problem 2).} \tag{29}$$

Finally by (27), (29) and Theorem 1 we can write

$$\begin{aligned} \max_{H\in\mathcal{F}_\epsilon} AV\left(\psi_{c^*},\Phi\right) &= \max_{H\in\mathcal{F}_\epsilon} AV\left(\psi_c,\Phi\right), \quad \text{for all } c>0 \\ &\leq \max_{H\in\mathcal{F}_\epsilon} AV\left(\psi,H\right), \quad \text{for all } \gamma\left(\psi,\Phi\right)=K\left(c\right) \quad \text{and for all } c>0 \\ &\leq \max_{H\in\mathcal{F}_\epsilon} AV\left(\psi,H\right), \quad \text{for all } \ \psi. \end{aligned}$$

Problems

Problem 1 *Show that if* ψ^* *satisfies*

$$\psi^* = \arg \min_{\gamma(\psi, H_0) \leq K_1} AV(\psi, \Phi)$$

for some $K_1 > \sqrt{2/\pi}$*, then* ψ^* *also satisfies*

$$\psi^* = \arg \min_{AV(\psi, H_0) \leq K_2} \gamma(\psi, H_0)$$

for some $K_2 > 1$.

Problem 2 *(a) Show that Equation (14) has a unique solution and write a computer program to solve it for each* $K > \sqrt{\pi/2}$.

(b) Let

$$g(c) = \alpha AV(\psi_c, \Phi) + \beta K(c)$$

and recall that

$$\alpha = 1/(1-\epsilon)$$

$$\beta = \epsilon/(1-\epsilon)^2$$

(see (25)). Find (numerically if necessary) the minimizing value

$$c^* = c^*(\epsilon) = \arg \min_{c>0} g(c), \tag{30}$$

for given $0 < \epsilon < 1/2$.

(c) Find

$$\lim_{\epsilon \to 0} c^*(\epsilon) = c_0$$

and

$$\lim_{\epsilon \to 0.5} c^*(\epsilon) \quad = \quad c_{1/2}.$$

(d) Explain the statistical meaning of c_0 *and* $c_{1/2}$.

Problem 3 *Verify that*

$$\frac{2\Phi(c) - 1 - 2c\varphi(c)}{[2\Phi(c) - 1]^2} \geq 0$$

for all $c \geq 0$.

Problem 4 *Find a sharp lower bound for the gross error sensitivity of dispersion M-estimates at the normal central model. Recall that*

$$GES\,(s, \Phi) = s_0 \frac{\max\{\chi(\infty) - b, b\}}{E_{H_0}\left\{\chi'\left(\frac{y - m_0}{s_0}\right)\left(\frac{x - m_0}{s(H_0)}\right)\right\}}$$

and argue that, without loss of generality we can work with the simplified expression

$$GES\,(s, \Phi) = \frac{\max\{\chi(\infty) - b, b\}}{E_\Phi\{\chi'(y)\, y\}}.$$

Problem 5 *Solve Hampel's optimality problem of minimizing the asymptotic variance at the central normal model subject to a bound on the GES for dispersion M-estimates at the normal central model.*

Problem 6 *Suppose that there exists* (x_0, y_0) *such that*

$$\min_x g(x, y_0) = \max_y g(x_0, y) = g\,(x_0, y_0)\,. \tag{31}$$

A point (x_0, y_0) *with this property is called a saddle point for the function* g. *Prove that if* (x_0, y_0) *is a saddle point for* g *then*

$$\min_x \max_y g(x, y) = \max_y \min_x g(x, y) = g\,(x_0, y_0)\,.$$

Problem 7 *Find the least favorable distribution* H^* *for the family of symmetric distributions* $\mathcal{F}_1$ *with finite variance. Show that in this case (31) with* $g(x, y) = AV\,(\psi, H)\,,\ x_0 = \psi_{H^*}$ *and* $y_0 = H^*$, *holds. Discuss the convenience and practical relevance issues in relation with* $\mathcal{F}_1$ *and the strengths and weaknesses of the corresponding minimax estimate.*

Problem 8 *Find the least favorable distribution* H^* *for the family* $\mathcal{F}_2$ *of symmetric and strongly unimodal distributions* H *with* $h\,(\mu) = k$, *where* μ *is the common center of symmetry for the distributions in* $\mathcal{F}_2$. *Show that in this case (31) also holds and therefore the min-max and max-min problem are equivalent. As in Problem 7 discuss the convenience and practical relevance of the family* $\mathcal{F}_2$ *and the strengths and weaknesses of the corresponding minimax estimate.*

References

Huber, P. J. (1964). Robust estimation of a location parameter. *Ann. Math. Statist.* **35**, 73–101.

Hampel, F.R. (1974). The influence curve and its role in robust estimation. *J. Am. Stat. Assoc.*, **69**, 383–393.

Li, B. and Zamar, R.H. (1991). Min-max asymptotic variance when scale is unknown. *Statistics and Probability Letters*, **11**, 139-145, (1991).