Ann. Inst. Statist. Math.
35 (1983), Part A, 235-241

A MINIMAX RESULT RELATED TO STEIN'S EFFECT

SAM GUTMANN

(Received Jan. 19, 1982)

1. Introduction

Stein [2] showed that the usual estimator for k normal means is inadmissible with squared error loss, $k \geqq 3$. But of course this estimator is minimax for all k. The purpose of this note is to explain the difference in a general context.

Imagine two statisticians, each to play a game against nature. The statisticians are sequestered and each is told what strategy nature is using in the *other's* problem. One interpretation [1] of Stein's effect is that the information is sometimes useful, in this sense: there may exist admissible procedures for the individual statisticians, which taken together become inadmissible, using summed risk, after the information is given. In Stein's result, one statistician must estimate θ after observing $X \sim N(\theta, 1)$; the other must estimate (ϕ_1, ϕ_2) after observing $(Y_1, Y_2) \sim N(\phi_1, 1) \times N(\phi_2, 1)$. The procedures $\hat{\theta}(X) = X$ and $\hat{\phi}(Y) = Y$ are admissible, but the joint procedure $\hat{\theta}$, $\hat{\phi}$ is inadmissible if the first statistician can use Y and second can use X also. A fortiori the joint procedure $\hat{\theta}$, $\hat{\phi}$ is inadmissible if the first statistician is told ϕ and the second θ.

But what if the statisticians believe in the minimax criterion only? The conclusion of the Theorem below is that then this information (each told nature's strategy in the other's problem) is *not* useful. That is, using a minimax strategy in each problem remains jointly minimax even after the information is given. Essentially the only hypothesis is that (at least) one of the games satisfy the minimax theorem.

In Section 3 this result is applied to the usual, more structured decision theoretic setting in which each statistician is told the data in the other's problem. Again, procedures minimax in the individual problems remain minimax when used in the joint problem.

2. Main result

As in [1], we start with the approach that a statistical decision problem is just a collection of risk functions. In the sequel, "measure" always means "probability measure", and "$\int f d\tau$ defined" means f is measurable for τ, and the integral is not of the form $\infty-\infty$.

Let Θ and Φ be sets, and let A (resp. B) be a collection of real valued functions on Θ (resp. Φ). Define a set of functions on Θ by

$$A^*=\left\{r^*:\ r^*(\theta)=\int r(\theta)\tau(dr) \text{ where } \tau \text{ is a measure on } A \text{ such that the integral is defined } \forall\,\theta\right\}$$

and define a set of functions on Φ by

$$B^*=\left\{R^*:\ R^*(\phi)=\int R(\phi)\tau(dR) \text{ where } \tau \text{ is a measure on } B \text{ such that the integral is defined } \forall\,\phi\right\} .$$

Here A^* and B^* play the role of randomized risk functions. (Of course randomized risk functions could already have been included in A, but it is possible that $A^{**}\neq A^*$.) Define nature's set of mixed strategies in the first problem by

$$\Pi=\left\{\pi:\ \pi \text{ is a measure on } \Theta \text{ s.t. } \int r(\theta)\pi(d\theta) \text{ is defined } \forall\, r\in A\right\} .$$

The minimax values of the games are given by

$$mM(A)=\inf_{r^*\in A^*}\sup_{\theta\in\Theta} r^*(\theta) \qquad \text{and} \qquad mM(B)=\inf_{R^*\in B^*}\sup_{\phi\in\Phi} R^*(\phi) .$$

THEOREM. *Assume*

(1) $$mM(A)+mM(B) \textit{ is defined, and}$$

(2) $$\sup_{\pi\in\Pi}\inf_{r\in A}\int r(\theta)\pi(d\theta)=\inf_{r^*\in A^*}\sup_{\theta\in\Theta} r^*(\theta) .$$

Let $\theta\to R_\theta$ *and* $\phi\to r_\phi$ *map* $\Theta\to B$ *and* $\Phi\to A$, *respectively. Assume*

(3) $$\int R_\theta(\phi)\pi(d\theta) \textit{ is defined } \forall\,\phi\in\Phi,\ \forall\,\pi\in\Pi .$$

Then $\sup_{\phi,\theta}[r_\phi(\theta)+R_\theta(\phi)]\geqq mM(A)+mM(B)$.

Note that R_θ is to be interpreted as the strategy used by the player in problem B when told that θ is the true state of nature in problem A. Note also that assumption (2) is the minimax theorem for problem A.

PROOF. First assume that $mM(A)+mM(B)$ is finite. Without loss of generality, assume $mM(A)+mM(B)=0$.

To derive a contradiction, assume $\sup_{\theta,\phi}[R_\theta(\phi)+r_\phi(\theta)]$ is negative, which implies that for some $\varepsilon>0$,

$$R_\theta(\phi)\leqq -r_\phi(\theta)-\varepsilon \ . \tag{4}$$

Now $mM(B)=\inf_{R^*}\sup_{\phi} R^*(\phi)$

$$\leqq \inf_{\pi\in\Pi}\sup_{\phi}\int R_\theta(\phi)\pi(d\theta)$$

since the second inf is taken over (possibly) fewer measures.

Now $\inf_{\pi\in\Pi}\sup_{\phi}\int R_\theta(\phi)\pi(d\theta)$

$$\begin{aligned}
&\leqq \inf_{\pi\in\Pi}\sup_{\phi}\int [-r_\phi(\theta)-\varepsilon]\pi(d\theta) && \text{by (4)}\\
&= -\sup_{\pi\in\Pi}\inf_{\phi}\left[\int r_\phi(\theta)\pi(d\theta)+\varepsilon\right]\\
&= -\sup_{\pi\in\Pi}\inf_{\phi}\int r_\phi(\theta)\pi(d\theta)-\varepsilon\\
&\leqq -\sup_{\pi\in\Pi}\inf_{r\in A}\int r(\theta)\pi(d\theta)-\varepsilon\\
&= -\inf_{r^*\in A^*}\sup_{\theta\in\Theta} r^*(\theta)-\varepsilon && \text{by (2)}\\
&= -mM(A)-\varepsilon \ .
\end{aligned}$$

So $mM(B)+mM(A)\leqq -\varepsilon$, the desired contradiction.

If $mM(A)+mM(B)=-\infty$, the conclusion of the theorem is vacuous. If $mM(A)+mM(B)=+\infty$, assume $\sup_{\theta,\phi}[R_\theta(\phi)+r_\phi(\theta)]<+\infty$, say $<K$. Proceed as above to deduce the contradiction $mM(B)\leqq -mM(A)+K$.

Assumption (3) can be replaced by the following condition involving upper integrals.

(3′) Let I be any set, $\alpha\to R_\alpha$ any map $I\to B$, and μ any measure on I. Then

$$\sup_{\phi}\overline{\int} R_\alpha(\phi)\mu(d\alpha)\geqq mM(B)\ .$$

Presumably, the standard statistical problems satisfy (3′).

As a converse to the theorem, assume that a problem A does not satisfy the minimax theorem, and

$$\inf_{r^*\in A^*}\sup_{\theta\in\Theta} r^*(\theta)>\sup_{\pi\in\Pi}\inf_{r\in A}\int r(\theta)\pi(d\theta)\ . \tag{5}$$

Define $\Phi = A$ and let $\phi \to r_\phi$ be the identity map. Now if

$$B = \{R_\theta :\ R_\theta(\phi) = -r_\phi(\theta),\ \theta \in \Theta\}$$

then $[R_\theta(\phi) + r_\phi(\theta)] \equiv 0$, but $mM(A) + mM(B)$

$$= \inf_{r^* \in A^*} \sup_{\theta \in \Theta} r^*(\theta) + \inf_{R^* \in B^*} \sup_{\phi \in \Phi} R^*(\phi)$$

$$= \inf_{r^* \in A^*} \sup_{\theta \in \Theta} r^*(\theta) - \inf_{\pi \in \Pi} \sup_{r \in A} \int r(\theta)\pi(d\theta) > 0 \qquad \text{by (5)}.$$

3. Application

If the decision problems A and B have the usual structure, then the above theorem can be interpreted as follows: using a minimax procedure in each problem is a minimax procedure in the joint problem, as defined below.

Let Θ, X, $\mathcal{A}$ be spaces (parameter, sample, and decision, resp.). For any space U let $\mathcal{S}(U)$ denote a given σ-algebra on U. For each $\theta \in \Theta$, let F_θ be a measure on $\mathcal{S}(X)$, and for each fixed $S \in \mathcal{S}(X)$, assume $F_\theta(S)$ is $\mathcal{S}(\Theta)$-measurable. Let $L_1 : \Theta \times \mathcal{A} \to [0, \infty]$ be an $\mathcal{S}(\Theta) \times \mathcal{S}(\mathcal{A})$-measurable (loss) function. The risk functions in this problem are of the form

$$\int_X F_\theta(dx) \int_{\mathcal{A}} L_1(\theta, a)\delta(x, da)$$

where $\delta(x, \cdot)$ is a measure on $\mathcal{S}(\mathcal{A})$ for each x, and for fixed $S \in \mathcal{S}(\mathcal{A})$, $\delta(\cdot, S)$ is $\mathcal{S}(X)$-measurable.

Assume a second problem is given with spaces Φ, Y, $\mathcal{B}$, corresponding σ-algebras $\mathcal{S}(\Phi)$, $\mathcal{S}(Y)$, $\mathcal{S}(\mathcal{B})$, measures G_ϕ and loss function L_2 satisfying the analogous conditions.

The joint problem is now defined by spaces $\Theta \times \Phi$, $X \times Y$, $A \times B$, corresponding σ-algebras $\mathcal{S}(\Theta) \times \mathcal{S}(\Phi)$, $\mathcal{S}(X) \times \mathcal{S}(Y)$, $\mathcal{S}(\mathcal{A}) \times \mathcal{S}(\mathcal{B})$ and measures $F_\theta \times G_\phi$ (to make the data in the two problems independent). The loss function in the joint problem is given by

$$L((\theta, \phi), (a, b)) = L_1(\theta, a) + L_2(\phi, b)\ .$$

If A (resp. B) denotes the set of finite risk functions in the first (resp. second) individual problem we define

$$A^u = A \cup \{f :\ f \text{ unbounded maps } \Theta \to [0, \infty]\} \quad \text{and}$$

$$B^u = B \cup \{f :\ f \text{ unbounded maps } \Phi \to [0, \infty]\}$$

and we assume

$$(6) \qquad \forall r^* \in A^*, \quad \exists f \in A^u \quad \text{with } f \leqq r^*, \text{ and}$$

$$(7) \qquad \forall R^* \in B^*, \quad \exists f \in B^u \quad \text{with } f \leqq R^* .$$

Recall that A and B were defined to include risk functions arising from randomized decisions.

COROLLARY. *Let the individual problems* A *and* B *satisfy* (6), (7) *and let* A *satisfy the minimax theorem*

$$(8) \qquad \sup_{\mu} \inf_{r \in A} \int r(\theta)\mu(d\theta) = \inf_{r \in A} \sup_{\mu} \int r(\theta)\mu(d\theta)$$

where the sup *is taken over all measures* μ *on* $\mathcal{S}(\Theta)$.

Further, assume that the smallest σ*-algebra making every* $r \in A$ *measurable is* $\mathcal{S}(\Theta)$.

Let $\rho(\theta, \phi)$ *be an arbitrary risk function in the joint problem. Then*

$$\sup_{\theta, \phi} \rho(\theta, \phi) \geqq \inf_{r} \sup_{\mu} \int r(\theta)\mu(d\theta) + \inf_{R} \sup_{\gamma} \int R(\phi)\gamma(d\phi) .$$

PROOF. The first step consists of showing that $\rho(\theta, \phi)$ can be written in the form $r_\phi(\theta) + R_\theta(\phi)$. A risk function in the joint problem is of the form

$$(9) \qquad \int_X \int_Y F_\theta(dx) G_\phi(dy) \int_{\mathcal{A} \times \mathcal{B}} L((\theta, \phi), (a, b)) \delta((x, y), d(a, b)) .$$

Define $\delta_i((x, y), \cdot)$, $i = 1, 2$ to be the marginal measures (projections) defined by $\delta((x, y), \cdot)$ on $\mathcal{A}$ and $\mathcal{B}$ respectively. Then the risk function (9) can be written as

$$(10) \qquad \int_X \int_Y F_\theta(dx) G_\phi(dy) \int_{\mathcal{A}} L_1(\theta, a) \delta_1((x, y), da)$$
$$+ \int_{\mathcal{B}} L_2(\phi, b) \delta_2((x, y), db) .$$

Define measures $\delta_\phi(x, \cdot)$ and $\delta_\theta(y, \cdot)$ by

$$\delta_\phi(x, S) = \int_Y \delta_1((x, y), S) G_\phi(dy) \quad \text{and} \quad \delta_\theta(y, S) = \int_X \delta_2((x, y), S) F_\theta(dx) .$$

Interchanging integrals, the risk function (10) can be written as

$$(11) \qquad \int_X F_\theta(dx) \int_{\mathcal{A}} L_1(\theta, a) \delta_\phi(x, da) + \int_Y G_\phi(dy) \int_{\mathcal{B}} L_2(\phi, b) \delta_\theta(y, db)$$

which is also $r_\phi(\theta) + R_\theta(\phi)$.

The measurability imposed on the problems ensures that the interchange of integrals is valid, and that δ_ϕ and δ_θ are procedures in the respective problems. Every $R \in A$ is $\mathcal{S}(\Theta)$ measurable, and since $\mathcal{S}(\Theta)$ was assumed to coincide with the smallest σ-algebra making every $r \in A$ measurable, assumption (3) of the Theorem must also hold.

Next, we use (6) and (8) to ensure that assumption (2) of the Theorem holds, and that

$$mM(A) = \inf_r \sup_\mu \int r(\theta)\mu(d\theta) .$$

Since $\mathcal{S}(\Theta)$ is large enough to make every $r(\theta)$ measurable, for fixed $r \in A$ it follows that

$$\sup_\mu \int r(\theta)\mu(d\theta) = \sup_\theta r(\theta)$$

and hence that

$$\inf_r \sup_\mu \int r(\theta)\mu(d\theta) = \inf_r \sup_\theta r(\theta) . \tag{12}$$

Since every measure on $\mathcal{S}(\Theta)$ is in Π, we have

$$\begin{aligned}\sup_{\pi \in \Pi} \inf_r \int r(\theta)\pi(d\theta) &\geqq \sup_\mu \inf_r \int r(\theta)\mu(d\theta) \\ &= \inf_r \sup_\mu \int r(\theta)\mu(d\theta) && \text{by (8)} \\ &= \inf_r \sup_\theta r(\theta) && \text{by (12)} \\ &\geqq \sup_\pi \inf_r \int r(\theta)\pi(d\theta) && \text{by an inequality}\end{aligned}$$

(one half of the minimax theorem) that always holds. We have shown

$$\inf_r \sup_\mu \int r(\theta)\mu(d\theta) = \inf_r \sup_\theta r(\theta) = \sup_\pi \inf_r \int r(\theta)\pi(d\theta) . \tag{13}$$

In order to conclude that assumption (2) holds and that the left-most expression in (13) equals $mM(A)$ we must only show

$$\inf_r \sup_\theta r(\theta) = \inf_{r^* \in A^*} \sup_\theta r^*(\theta) .$$

This follows quickly from (6). Likewise

$$mM(B) = \inf_R \sup_\gamma \int R(\phi)\gamma(d\phi)$$

follows from (7) and the proof is complete.

Applying the Corollary repeatedly provides an analogous result for k problems joined together.

NORTHEASTERN UNIVERSITY

REFERENCES

[1] Gutmann, S. (1983). Decisions immune to Stein's effect, *Sankhyā*, A (to appear).
[2] Stein, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution, *Proc. 3rd Berkeley Symp. Math. Statist. Prob.*, Univ. of Calif. Press, Berkeley, 1, 197-206.