Sankhyā : The Indian Journal of Statistics
Special Issue on Bayesian Analysis
1998, Volume 60, Series A, Pt. 3, pp. 383-404

BAYES DECISION PROBLEMS AND STABILITY*

By JOSEPH B. KADANE
Carnegie Mellon University, Pittsburgh
and
C. SRINIVASAN
University of Kentucky, Lexington

SUMMARY. Stability of Bayes decision problems under uniform convergence of losses is revisited and sufficient conditions for stability are obtained. The results generalize and complement the earlier works of Kadane and Chuang, Chuang, and Salinetti. General conditions are also given for the equivalence of two definitions of stability.

1. Introduction

Recent years have seen excellent progress in the ability of Bayesians to compute posterior and predictive quantities of interest. This progress adds to the importance that must be placed on what the inputs to those computations should be, that is, what likelihood, prior and loss function to use. While there are various standpoints on these questions, the one we find most satisfying in principle is the subjective view of Savage (1954, 1962). This requires subjective elicitation of the inputs.

Getting serious about elicitation means to admit that elicited quantities can not be held to be exact representations of the opinions, in the case of priors and likelihoods, or desires, in case of loss functions, of the person being elicited. One would like the problem to be "forgiving" (or "robust") in the sense that small errors in the inputs should not cause decisions to appear optimal that are very much worse than could have been made were the inputs correct.

The theory of stability of decision problems, as formulated by Kadane and Chuang (1978), is a way to formalize whether an elicitation problem is nonrobust. If strong stability obtains, close enough elicitation leads to decisions with

AMS (1991) *subject classification.* 62C10
Key words and phrases. Bayes decisions, stability, prior distributions and epiconvergence.
* Research supported in part by NSF Grants ATM-9108177, SES-9123370, DMS-9303557 and ONR Contract N00014-89-J-1851.

expected loss, evaluated correctly, nearly as small as is achievable. If weak stability obtains, a particular stabilizing decision will achieve the benefit described above. But if neither is the case, even a very small elicitation error can lead to discontinuously much additional expected loss.

The main focus of this paper is a detailed study of stability under uniform convergence of losses with the purpose of deriving sufficient conditions for stability and the equivalence of the two definitions of stability. The results obtained here generalize the results of Kadane and Chuang (1978) and Chuang (1984) in different directions. An important earlier work related to this paper is Salinetti (1994). Indeed, the results presented here (see Section 4) heavily rely on some of the tools and results from the above article.

The literature on the sensitivity analysis of Bayes decision problems with respect to the prior and the loss is by no means limited to the papers listed above. Early work in this direction is due to Edwards, Lindeman and Savage (1963). Other important contributions include Stone (1963), Fishburn *et al.* (1967) and Brittney and Winkler (1974). The last paper, in particular, contains an illuminating discussion on the sensitivity of Bayes estimates to misspecification of the loss.

As a strong form of continuity, even if stability obtains, it does not specify the rate at which additional loss approaches zero as the extent of error approaches zero. This and the role of stability in quantitative robustness analysis are the subjects of subsequent papers.

Finally, the remainder of the paper is organized as follows. The two definitions of stability and an example showing their non-equivalence are in section 2. In section 3, the conditions under which the definitions are equivalent are discussed. Sections 4 and 5 contain sufficient conditions for strong stability I and II, respectively.

2. **Stability**

To formulate the definitions of stability of a decision problem, suppose that the parameter space is $\Theta \subset R^m$, the decision space is $D \subset R^m$ and the likelihood is a bounded continuous function $\ell_0(\theta)$. Let $L_0(\theta, d)$ be a loss function and P_0 be a prior distribution on Θ. Also, let $L_n,\ n = 1, 2, \ldots$ denote a sequence of loss functions converging (in some topology) to L_0 and $P_n, n = 1, 2 \ldots$ a sequence of priors converging weakly to P_0. We denote weak convergence by "$P_n \Rightarrow P_0$".

Definition I. The decision problem (L_0, ℓ_0, P_0) is Strongly Stable I (SSI) if for every sequence $P_n \Rightarrow P_0$ and $L_n \to L_0$

$$\lim_{\epsilon\downarrow 0} \limsup_{n\to\infty} \left[\int L_n(\theta, d_0(\epsilon))\ell_0(\theta)dP_n - \inf_{d\epsilon D} \int L_n(\theta, d)\ell_0(\theta)dP_n \right] = 0 \quad \ldots (2.1)$$

for every $d_0(\epsilon)$ such that

$$\int L_0(\theta, d_0(\epsilon))\ell_0(\theta)dP_0 \leq \inf_{d\in D} \int L_0(\theta, d)\ell_0(\theta)dP_0 + \epsilon. \qquad \ldots (2.2)$$

The triple (L_0, ℓ_0, P_0) is said to be Weakly Stable I (WSI) if (2.1) holds for a particular choice $d_0(\epsilon)$.

Viewing (L_0, ℓ_0, P_0) as the approximate specification by the statistician and assuming the "truth" lies along the sequence (L_n, ℓ_0, P_n), the Strong Stability I of (L_0, ℓ_0, P_0) implies that "small" errors in the specification of the loss and the prior will not result in substantially worse (in terms of the risk) decisions. If, on the other hand, (L_0, ℓ_0, P_0) is unstable (i.e. not even weakly stable), even small errors in the specification of the loss and the prior may result in worse decisions. This essentially motivates the above definition.

A more general and stringent definition of stability is possible and is as follows.

Definition II. The decision problem (L_0, ℓ_0, P_0) is Strongly Stable II if for all sequences $P_n \Rightarrow P_0$, $Q_n \Rightarrow P_0$, $L_n \to L_0$ and $W_n \to L_0$

$$\lim_{\epsilon\downarrow 0} \limsup_{n\to\infty} \left[\int L_n(\theta, d_{Q_n}(\epsilon))\ell_0(\theta)dP_n - \inf_{d\in D} \int L_n(\theta, d)\ell_0(\theta)dP_n\right] = 0 \quad \ldots (2.3)$$

for every $d_{Q_n}(\epsilon)$ satisfying

$$\int W_n(\theta, d_{Q_n}(\epsilon))\ell_0(\theta)dQ_n \leq \inf_{d\in D} \int W_n(\theta, d)\ell_0(\theta)dQ_n + \epsilon. \qquad \ldots (2.4)$$

The problem (L_0, ℓ_0, P_0) is Weakly Stable II if (2.3) holds for a particular choice of $d_{Q_n}(\epsilon)$.

The reader is referred to Kadane and Chuang (1978) for a motivation of the second definition of stability. In the same paper, the authors studied the stability of a decision problem under the topology of uniform convergence (i.e. $L_n(\theta, d)$ converges to $L_0(\theta, d)$ uniformly in θ and d) for the losses and obtained sufficient conditions for stability. In particular, they noted that under uniform convergence (2.1) is equivalent to

$$\lim_{\epsilon\downarrow 0} \limsup_{n\to\infty} \left[\int L_0(\theta, d_0(\epsilon))\ell_0(\theta)dP_n - \inf_{d} \int L_0(\theta, d)\ell_0(\theta)dP_n\right] = 0 \quad \ldots (2.5)$$

and, (2.3) and (2.4) are equivalent to

$$\lim_{\epsilon\downarrow 0} \limsup_{n\to\infty} \left[\int L_0(\theta, d_{Q_n}(\epsilon))\ell_0(\theta)dP_n - \inf_{d} \int L_0(\theta, d)\ell_0(\theta)dP_n\right] = 0 \quad \ldots (2.6)$$

and

$$\int L_0(\theta, d_{Q_n}(\epsilon))\ell_0(\theta)dQ_n \leq \inf_d \int L_0(\theta, d)\ell_0(\theta)dQ_n + \epsilon \qquad \ldots(2.7)$$

respectively. This fact considerably simplifies the task of verifying stability.

As indicated earlier, SSII implies SSI. However, the converse is not in general true as the following simple example shows. Chuang (1984) has also given a similar, but more involved and somewhat artificial, example.

EXAMPLE 2.1. Let $\Theta = D = [-1, 1]$ and the loss L_0 be given by

$$L_0(\theta, d) = \begin{cases} 1 & \text{if } d \neq 0,\ \theta d \leq 0 \\ (\theta - d)^2 & \text{otherwise.} \end{cases}$$

Consider the decision problem (L_0, ℓ_0, P_0) where $\ell_0(\theta) \equiv 1$ and $P_0 = \delta_{\{0\}}$, the degenerate probability distribution at $\theta = 0$. It is easy to see that $\int L(\theta, d)dP_0(\theta) = 0$ if $d = 0$ and 1 if $d \neq 0$. Consequently, for any $1 > \epsilon > 0$, the optimal decision at ϵ of (L_0, ℓ_0, P_0) is $d_0(\epsilon) \equiv 0$. Now, let $\{P_n\}$ be any sequence of prior distributions on Θ such that $P_n \Rightarrow P_0$. Then

$$\inf_d \int L_0(\theta, d)dP_n \leq \int L_0(\theta, 0)dP_n = \int_{-1}^{1} \theta^2 dP_n \to 0$$

as $n \to \infty$. Hence

$$\int L_0(\theta, d_0(\epsilon))dP_n - \inf_d \int L_0(\theta, d)dP_n$$

$$\leq \int L_0(\theta, 0)dP_n = \int_{-1}^{1} \theta^2 dP_n \to 0$$

as $n \to \infty$. This proves (L_0, ℓ_0, P_0) is Strongly Stable by Definition I. Finally, to show that (L_0, ℓ_0, P_0) is not Strongly Stable by Definition II, consider the two sequences of prior distributions $P_n \equiv \delta_{\{-\frac{1}{n}\}}$ and $Q_n \equiv \delta_{\{\frac{1}{n}\}}$. The expected loss of any decision d under the prior distribution Q_n is

$$\int L_0(\theta, d)dQ_n = \begin{cases} (\frac{1}{n} - d)^2 & \text{if } d \geq 0 \\ 1 & \text{if } d < 0 \end{cases}$$

and an ϵ-optimal decision of (L_0, ℓ_0, Q_n) is $d_{Q_n} = \frac{1}{n}$. A similar calculation shows

$$\inf_d \int L_0(\theta, d)dP_n = 0$$

Combining these facts,

$$\int L_0(\theta, d_{Q_n})dP_n - \inf_d \int L_0(\theta, d)dP_n$$
$$= \int L_0(\theta, \frac{1}{n})dP_n = L_0(-\frac{1}{n}, \frac{1}{n}) = 1.$$

Thus (2.3) is not satisfied and, hence, (L_0, ℓ_0, P_0) is not Strongly Stable by Definition II. □

3. **Equivalence of Definitions**

In view of the example given in the preceding section, it is natural to seek sufficient conditions for the equivalence of the two definitions. This section contains two results which give sufficient conditions for the equivalence of the definitions for strong stability.

The first result is applicable for general loss functions but involves the following differentiability assumption. For simplicity,the result is stated for the case when the dimension $m = 1$. Extension to the multidimensional case is straightforward.

Let $D_\epsilon^P = \{d : \int L_0(\theta, d)\ell_0(\theta)dP \leq \inf_t \int L_0(\theta, t)\ell_0(\theta)dP + \epsilon\}$, i.e., ${D_\epsilon}^P$ is the set of all ϵ-optimal decisions of (L_0, ℓ_0, P).

Assumption A : (I) For some $\epsilon_0 > 0$, there exist a weak neighborhood $N(P_0)$ of P_0 and a compact convex set $K \subset D$ such that for every $P \in N(P_0)$, $D_{\epsilon_0}^P \subset K$.

(II) For every $P \in N(P_0)$, $\int L_0(\theta, t)\ell_0(\theta)dP$ is twice continuously differentiable and there exist positive constants B and δ such that for all $t \in K$

$$B \geq \frac{d^2}{dt^2} \int L_0(\theta, t)\ell_0(\theta)dP \geq \delta.$$

A remark about the assumption is in order. The assumption (A.I) holds if there exists a function g(t) with bounded level sets such that for every P in some weak neighborhood $N(P_0)$,$\int L_0(\theta, t)\ell_0(\theta)dP \geq g(t)$ and, for each t, $L_0(\theta, t)\ell_0(\theta)$ is bounded.

Theorem 3.1. *Suppose (A) holds.Then (L_0, ℓ_0, P_0) is Strongly Stable I if and only if it is Strongly Stable II.*

The following lemma plays a crucial role in the proof of the theorem. Though the lemma is more generally true, the version given below is for functions with real domain.

Lemma 3.2. *Let $g : R \to R$ be a differentiable function. Suppose, for a given $\epsilon > 0$, X_ϵ is such that $g(X_\epsilon) \leq \inf_{X \in R} g(X) + \epsilon$. Then there exists Y_ϵ satisfying* *(i)* $g(Y_\epsilon) \leq g(X_\epsilon)$

(ii) $| Y_\epsilon - X_\epsilon | < \sqrt{\epsilon}$

(iii) $| g'(Y_\epsilon) | < \sqrt{\epsilon}$.

where g' is the derivative of g.

For a proof of this lemma see Ekeland and Temam (1979). The lemma essentially asserts the existence of 'near stationary' ϵ-minimizer of g in the vicinity of any ϵ-minimizer.

PROOF OF THEOREM 3.1. It suffices to prove the "only if" part. So, assume (L_0, ℓ_0, P_0) is Strongly Stable I. Let $\{P_n\}$ and $\{Q_n\}$ be two arbitrary sequences of prior distributions converging weakly to P_0. Without loss of any generality, assume that $P_n \in N(P_0)$ and $Q_n \in N(P_0)$ for all n where $N(P_0)$ is the neighborhood given by the assumption (A). Now let $d_{Qn}(\epsilon)$ be an arbitrary sequence of ϵ-optimal solutions of (L_0, ℓ_0, Q_n) decision problems *i.e.* for $n = 1, 2, \ldots$

$$\int L_0(\theta, d_{Q_n}(\epsilon)) dQ_n \leq \inf_d \int L_0(\theta, d) dQ_n + \epsilon. \qquad \ldots (3.1)$$

Below it is shown that

$$\lim_{\epsilon \downarrow 0} \limsup_{n \to \infty} \left[\int L_0(\theta, d_{Q_n}(\epsilon)) dP_n - \inf_d \int L_0(\theta, d) dP_n \right] = 0, \qquad \ldots (3.2)$$

proving (L_0, ℓ_0, P_0) is Strongly Stable II.

Towards this, observe that by Lemma 3.2 there exists a sequence $\{\tilde{d}_{Q_n}(\epsilon)\}$ such that

(i) $\quad | d_{Qn}(\epsilon) - \tilde{d}_{Qn}(\epsilon) | < \sqrt{\epsilon}$

(ii) $\quad \int L_0(\theta, \tilde{d}_{Q_n}(\epsilon)) \ell_0(\theta) dQ_n \leq \inf_d \int L_0(\theta, d) \ell_0(\theta) dQ_n + \epsilon \qquad \ldots (3.3)$

(iii) $\quad | \frac{d}{dt} \int L_0(\theta, t) \ell_0(\theta) dQ_n |_{t = \tilde{d}_{Q_n}(\epsilon)} | < \sqrt{\epsilon}$

Moreover, since (L_0, ℓ_0, P_0) is Strongly Stable I,

$$\lim_{\epsilon \downarrow 0} \limsup_{n \to \infty} \left[\int L_0(\theta, d_0(\epsilon)) \ell_0(\theta) dQ_n - \int L_0(\theta, \tilde{d}_{Q_n}(\epsilon)) \ell_0(\theta) dQ_n \right] = 0 \quad \ldots (3.4)$$

for every $d_0(\epsilon)$ given by

$$\int L_0(\theta, d_0(\epsilon)) \ell_0(\theta) dP_0 \leq \inf_d \int L_0(\theta, d) \ell_0(\theta) dP_0 + \epsilon. \qquad \ldots (3.5)$$

Expanding $\int L_0(\theta, d_0(\epsilon))\ell_0(\theta)dQ_n$ about $\tilde{d}_{Q_n}(\epsilon)$ by Taylor Series (3.4) leads to

$$\lim_{\epsilon\downarrow 0} \limsup_{n\to\infty} \Big[(d_0(\epsilon) - \tilde{d}_{Q_n}(\epsilon))\tfrac{d}{dt}\int L_0(\theta,t)\ell_0(\theta)dQ_n \mid_{t=\tilde{d}_{Q_n}(\epsilon)}$$
$$+ \tfrac{(d_0(\epsilon)-\tilde{d}_{Q_n}(\epsilon))^2}{2}\tfrac{d^2}{dt^2}\int L_0(\theta,t)\ell_0(\theta)dQ_n \mid t = d_n^*\Big] = 0 \qquad \ldots (3.6)$$

where $d_n^* \in K$. Now, invoking assumption $(A.II)$, it follows from (3.3 ii) and (3.6) that

$$\lim_{\epsilon\downarrow 0}\limsup_{n\to\infty}\left[-|d_0(\epsilon) - \tilde{d}_{Q_n}(\epsilon)|\epsilon + \frac{(d_0(\epsilon) - \tilde{d}_{Q_n}(\epsilon))^2}{2}\delta\right] \leq 0. \qquad \ldots (3.7)$$

Since $d_0(\epsilon) \in K, \tilde{d}_{Q_n}(\epsilon) \in K$,and $\delta > 0$ this implies

$$\lim_{\epsilon\downarrow 0}\limsup_{n\to\infty} \mid d_0(\epsilon) - \tilde{d}_{Q_n}(\epsilon) \mid = 0 \qquad \ldots (3.8)$$

Therefore, by assumption (A) and (3.8),

$$\lim_{\epsilon\downarrow 0}\limsup_{n\to\infty} \mid \int L_0(\theta, \tilde{d}_{Q_n}(\epsilon))\ell_0(\theta)dP_n - \int L_0(\theta, d_0(\epsilon))\ell_0(\theta)dP_n \mid$$
$$\leq B_2 \lim_{\epsilon\downarrow 0}\limsup_{n\to\infty} \mid \tilde{d}_{Qn}(\epsilon) - d_0(\epsilon) \mid = 0 \qquad \ldots (3.9)$$

where B_2 is a constant.

Finally, since (L_0, ℓ_0, P_0) is Strongly Stable I and $P_n \Rightarrow P_0$, by (3.9)

$$\lim_{\epsilon\downarrow 0}\limsup_{n\to\infty}\Big[\int L_0(\theta, \tilde{d}_{Qn}(\epsilon))\ell_0(\theta)dP_n - \inf_d \int L_0(\theta,d)\ell_0(\theta)dP_n\Big]$$
$$= \lim_{\epsilon\downarrow 0}\limsup_{n\to\infty}\Big[\int L_0(\theta, \tilde{d}_{Qn}(\epsilon))\ell_0(\theta)dP_n - \int L_0(\theta, d_0(\epsilon))\ell_0(\theta)dP_n\Big] = 0 \qquad \ldots (3.10)$$

The proof is now completed by observing that (3.10) along with 3.3(ii) implies (3.2). □

The next result establishes the equivalence of the two definitions for location invariant loss functions $L_0(\theta - d) : R^m \to R_+$ and satisfying

(i) $L(0) = 0$

(ii) $L(t)$ is continuous in t

(iii) $\{t : L(t) \leq c\}$ is convex and bounded for every $c > 0$.

Loss functions satisfying the conditions (i) – (iii) are known as bowlshaped loss functions and are often used in prediction and estimation problems. The condition (iii) deserves a comment. It implies the loss function is non-decreasing at $t \to \infty$ in any direction and that the loss $L(t) \to \infty$ as $t \to \infty$.

THEOREM 3.3. *Suppose $L_0(\theta - d)$ satisfies the conditions (i) — (iii) and $L_0(\theta - d)\ell_0(\theta)$ is bounded continuous on θ for every d. Then (L_0, ℓ_0, P_0) is Strongly Stable I if and only if it is Strongly Stable II.*

The proof of the theorem is an immediate consequence of the following lemmas. In each of these lemmas, the conditions stated in Theorem 3.3 are assumed.

LEMMA 3.4. *Let $\epsilon > 0$ and $d_0(\epsilon)$ be an ϵ-optimal solution of (L_0, ℓ_0, P_0). Then $\mid d_0(\epsilon) \mid < B_3$ for some constant B_3 depending only on ϵ.*

PROOF. It suffices to show that for any $B > 0$, the set $S = \{d : \int L_0(\theta - d)\ell_0(\theta)dP_0 < B\}$ is bounded. Towards this, let K be a compact set such that $\int_K \ell_0(\theta)dP_0 > \frac{1}{2}$ and C be the bounded set $\{t : L_0(t) \leq 4B\}$. Define $\mid K \mid = \sup_{\theta \in K} \parallel \theta \parallel$ and $\mid C \mid = \sup_{t \in C} \parallel t \parallel$. Clearly, $\mid K \mid < \infty$, $\mid C \mid < \infty$ since K and C are bounded.

Now we claim that for any $d \in S$, $\parallel d \parallel < \mid K \mid + \mid C \mid$. To see this, assume the contrary. Then, by triangle inequality $\parallel \theta - d \parallel > \mid C \mid$ for all $\theta \in K$ *i.e.* $\theta - d \in R^m - C$ for all $\theta \in K$ and, therefore leading to the contradiction

$$B > \int_K L_0(\theta - d)\ell_0(\theta)dP_0 > 4B \int_K \ell_0(\theta)dP_0 > 2B.$$

□

LEMMA 3.5. *Suppose (L_0, ℓ_0, P_0) is Strongly Stable I and $Q_n \Rightarrow P_0$. Then for any $\epsilon > 0$ there exists a constant B_4 and an integer n_0 such that for all $n \geq n_0, \parallel d_{Q_n}(\epsilon) \parallel < B_4$ for every ϵ-optimal decision $d_{Q_n}(\epsilon)$ of (L_0, ℓ_0, Q_n).*

PROOF. The strong stability I of (L_0, ℓ_0, P_0) and the fact $L_0(\theta - d)\ell_0(\theta)$ is bounded continuous in θ for a fixed d together imply that there exists n_0 such that for all $n \geq n_0$

$$\int L_0(\theta - d_{Q_n}(\epsilon))\ell_0(\theta)dQ_n \leq \inf_d \int L_0(\theta - d)\ell_0(\theta)dP_0 + \epsilon$$

Thus the sequence $\{\int L_0(\theta - d_{Qn}(\epsilon))\ell_0(\theta)dQ_n(\theta)\}$ is uniformly bounded. Now using the argument given in Lemma 3.4, with K such that $Q_n(K) > \frac{1}{2}$ for all n, the result follows. □

LEMMA 3.6. *Let $P_n \Rightarrow P_0$ and $K_1 \subset R^m$ be a compact set. Then for any $\epsilon > 0$ there exists n_0 such that for all $n \geq n_0$ and all $d_1, d_2 \in K_1$*

$$\mid \int (L_0(\theta - d_1) - L_0(\theta - d_2))\ell_0(\theta)dP_n - \int (L_0(\theta - d_1) - L_0(\theta - d_2))\ell_0(\theta)dP_0 \mid < \epsilon$$

PROOF. Since $P_n \Rightarrow P_0$, there exists a compact set K such that $P_n(K) > 1 - \frac{\epsilon}{4B}$ where $B = 2 \sup_{d \in K_1} \sup_\theta L_0(\theta - d)\ell_0(\theta)$. Clearly,

$$\int_{R^m - K} (L_0(\theta - d_1) - L_0(\theta - d_2))\ell_0(\theta) dP_n < \frac{\epsilon}{4} \qquad \ldots (3.11)$$

for all $d_1, d_2, \in K$, and for all n.

Let K^γ be a compact γ-neighborhood of K for some $\gamma > 0$. Since $L_0(t)$ is uniformly continuous on compact sets, for given $\epsilon_1 > 0$ there exists $\delta_1 > 0$ such that for all $\theta \in K^\gamma$, $\mid L_0(\theta, x) - L_0(\theta - y) \mid < \epsilon_1$ whenever $x, y \in K_1$ and $\| x - y \| < \delta_1$. By the compactness of K_1, there exist finitely many, say k, spheres of radius δ_1, covering K_1. Denote the centers of these spheres by $t_1, \ldots, t_k$. Then, for any $d_1, d_2 \in K_1$ there exist t_i, t_j such that

$$\mid \int_K (L_0(\theta - d_1) - L_0(\theta - d_2)) - (L_0(\theta - t_i) - L_0(\theta - t_j))\ell_0(\theta) dP_n \mid < 2\epsilon_1 \quad \ldots (3.12)$$

for all n.

Also, by Urysohn's Lemma (Kelly 1955) there exists a continuous function $g(\theta)$ with the property $g(\theta) \equiv 1$ on K and $g(\theta) \equiv 0$ on $R^m - K^\gamma$ and, therefore, for all t_i, t_j and all $n \geq 1$,

$$\begin{aligned} &\mid \textstyle\int_K (L_0(\theta - t_i) - L_0(\theta - t_j))\ell_0(\theta) dP_n - \int (L_0(\theta - t_i) - L_0(\theta - t_j))\ell_0(\theta) g(\theta) dP_n \mid \\ &< \textstyle\int_{K^\gamma - K} (L_0(\theta - t_i) + L_0(\theta - t_j))\ell_0(\theta) dP_n < B\frac{\epsilon}{4B} = \frac{\epsilon}{4} \end{aligned} \qquad \ldots (3.13)$$

Finally, by the bounded continuity of $L_0(\theta - t_i)\ell_0(\theta)g(\theta)$, $i = 1, 2 \ldots, k$, there exists n_o such that for all $n \geq n_0$

$$\begin{aligned} &\mid \int (L_0(\theta - t_i) - L_0(\theta - t_j))\ell_0(\theta) g(\theta) dP_n \\ &- \int (L_0(\theta - t_i) - L_0(\theta - t_j))\ell_0(\theta) g(\theta) dP_0 \mid \\ &< \frac{\epsilon}{8} \end{aligned} \qquad \ldots (3.14)$$

for all $i, j = 1, 2, \ldots, k$. Combining this with 3.10 - 3.12 the result follows. □

PROOF OF THEOREM 3.3. By Lemmas 3.4 and 3.5, for a given $\epsilon > 0$ there exists a compact set K_1 such that for all n, $d_{Q_n}(\epsilon) \in K_1$ and $d_0(\epsilon) \in K_1$ where $Q_n \Rightarrow P_0$. Now, since $P_n \Rightarrow P_0$, by Lemma 3.6 and the triangle inequality, with $d_0(\epsilon)$ and $d_{Q_n}(\epsilon)$ substituted for d_2 and d_1 respectively,

$$\begin{aligned} &\mid \int (L_0(\theta - d_{Q_n}(\epsilon)) - L_0(\theta - d_0(\epsilon))\ell_0(\theta) dP_n \\ &- \int (L_0(\theta - d_{Qn}(\epsilon)) - L_0(\theta - d_0(\epsilon))\ell_0(\theta) dQ_n \mid \leq \epsilon \end{aligned}$$

for all $n \geq N_2(\epsilon)$ for some $N_2(\epsilon)$. The second term goes to zero as $n \to \infty$ and $\epsilon \to 0$ because (L_0, ℓ_0, P_0) is Strongly Stable I and $d_{Q_n}(\epsilon)$ is (L_0, ℓ_0, Q_n) ϵ-optimal. The proof is now completed by observing

$$\lim_{\epsilon \downarrow 0} \left(\limsup_{n \to \infty} \int L_0(\theta - d_0(\epsilon))\ell_0(\theta) dP_n - \inf_d \int L_0(\theta - d)\ell_0(\theta) dP_n \right) = 0$$

by the SSI of (L_0, ℓ_0, P_0). □

The final result of this section treats convex loss functions. The next theorem shows that,under very mild conditions, the two definitions of stability are equivalent for convex losses.The following Proposition summarizes some important properties of convex functions used in the proof of the theorem. For details see Rockafellar (1970).

PROPOSITION 3.1. *Suppose $\Phi_n, n \geq 1$, and Φ_0 are finite convex functions defined on R^m. Then the following hold.*

(i) If $\Phi_n \to \Phi_0$ pointwise, then $\Phi_n \to \Phi_0$ uniformly on compact sets.

(ii) If Φ_0 has a non-empty bounded level set then all its level sets are bounded. In addition, they are closed and convex.

(iii) The minimum set of Φ_0 is non-empty and bounded if, and only if, for some x and every $y \neq 0$

$$\lim_{\lambda \to \infty} \frac{\Phi_0(x + \lambda y) - \Phi_0(x)}{\lambda} > 0.$$

LEMMA 3.7. *Let $\Phi_n, n \geq 1$, and Φ_0 be finite non-negative convex functions defined on R^m. Assume*

(a) the minimum set of Φ_0 is non-empty and bounded.

(b) $\Phi_n \to \Phi_0$ pointwise.

(c) for every $x_\epsilon \in M_\epsilon^0$ $=\{x : \Phi_0(x) \leq \inf \Phi_0(y) + \epsilon\}$

$$\lim_{\epsilon \downarrow 0} \limsup_{n \to \infty} [\Phi_n(x_\epsilon) - \inf_x \Phi_n(x)] = 0$$

Then for every $\beta > 0$ there exist ϵ_0, n_0 and r_0 such that

(i) for all $\epsilon \leq \epsilon_0$ and $n \geq n_0$

$$M_\epsilon^0 \subset M_\beta^n = \{x : \Phi_n(x) \leq \inf_y \Phi_n(y) + \epsilon\}.$$

(ii) $\limsup_{n \to \infty} M_\beta^n \subset B(r_0)$ *where $B(r)$ is the sphere of radius r centered at the origin.*

PROOF. Part (i) is an immediate consequence of the assumption (c).

The proof of (ii) is by contradiction. Towards this, let $\beta > 0$ and suppose $\limsup_{n \to \infty} M_\beta^n$ is unbounded. Then there exists a subsequence $\{m\} \subset \{n\}$ and $x_m \in M_\beta^m$ such that $\|x_m\| \uparrow \infty$.

Next, by assumption (a) and the preceding Proposition, the set $M^0_{2\beta}$ is bounded and, therefore, contained in B(r) for some $r > 0$. Now, let x_0 be a point in the minimum set of Φ_0. Since the level sets are nested, it follows from (i) that $x_0 \in M^m_\beta$ for all sufficiently large m. Moreover, due to the convexity of Φ_m,the line segment $[x_0, x_m]$ is contained in M^m_β for all sufficiently large m. Therefore, for every m such that $\|x_m\| \geq 3r$, there exists $y_m \in [x_0, x_m]$ with the property $2r < \|y_m\| < 3r$. Let $\{y_k\}$ be a convergent subsequence of $\{y_m\}$ with limit y_0. Note that $\|y_0\| > r$ and therefore $\Phi_0(y_0) \geq 2\beta$. Also, for all sufficiently large k, $\Phi_k(y_k) \leq \beta$.

But $\Phi_k(y_k) \to \Phi_0(y_0)$ because, by the assumption (b) and the Proposition, $\Phi_n \to \Phi_0$ uniformly on compact sets. This contradiction completes the proof. □

Remark. The conclusions of the Lemma 3.7 continue to hold if the domain R^m is replaced by an open convex subset of R^m.

Theory 3.8. *Let $(L_0(\theta, d), \ell_0(\theta), P_0)$ be a decision problem with $\theta \in \Theta$ and $d \in D$, an open convex set. Suppose*

(i) $L_0(\theta, d)\ell_0(\theta)$ is bounded continuous in θ for each d and convex in d for each θ.

(ii) The minimum set of $\int L_0(\theta, d)\ell_0(\theta)dP_0$ is nonempty and bounded.

Then (L_0, ℓ_0, P_0) is Strongly Stable I if and only if it is Strongly Stable II.

Proof. Enough to prove the "only if" part. Assume (L_0, ℓ_0, P_0) is SSI. Let $P_n \implies P_0$ and $Q_n \implies P_0$. The condition (i) implies that, for every n, $\int L_0(\theta, d)\ell_0(\theta)dP_n$ and $\int L_0(\theta, d)\ell_0(\theta)dQ_n$ are finite, continuous convex functions in d. Moreover, both integrals converge pointwise, and hence uniformly on compact sets, to $\int L_0(\theta, d)\ell_0(\theta)dP_0$.

Since (L_0, ℓ_0, P_0) is SSI it follows from (ii) that the conditions of the lemma are met by the convex functions $\int L_0(\theta, d)\ell_0(\theta)dQ_n$, $n \geq 1$, and $\int L_0(\theta, d)\ell_0(\theta)dP_0$. Therefore, for all sufficiently large n, the sets $D^{Q_n}_\epsilon$ consisting of the ϵ-optimal decisions of (L_0, ℓ_0, Q_n) are uniformly bounded.

Now, to establish the Strong Stability II of (L_0, ℓ_0, P_0), let $d_{Q_n}(\epsilon) \in D^{Q_n}_\epsilon$ be an ϵ-optimal decision of (L_0, ℓ_0, Q_n).Then, for each ϵ, $d_{Q_n}(\epsilon), n \geq 1$ are bounded and, by the triangle inequality,

$$\limsup_{n\to\infty} \left| \int L_0(\theta, d_{Q_n}(\epsilon))\ell_0(\theta)dP_n - \int L_0(\theta, d_{Q_n}(\epsilon))\ell_0(\theta)dQ_n \right| = 0.$$

Also, since(L_0, ℓ_0, P_0) is SSI, it follows from the condition (ii) and another application of the triangle inequality

$$\limsup_{n\to\infty} \left| \inf_d \int L_0(\theta, d)\ell_0(\theta)dP_n - \inf_d \int L_0(\theta, d)\ell_0(\theta)dQ_n \right| = 0.$$

A straight-forward consequence of these is

$$\limsup_{n\to\infty}[\int L_0(\theta, d_{Q_n}(\epsilon))\ell_0(\theta)dP_n - \inf_d \int L_0(\theta, d)\ell_0(\theta)dP_n]$$
$$\leq \limsup_{n\to\infty}[\int L_0(\theta, d_{Q_n}(\epsilon))\ell_0(\theta)dQ_n - \inf_d \int L_0(\theta, d)\ell_0(\theta)dQ_n].$$
$$\leq \epsilon.$$

Finally,the result follows by letting ϵ tend to zero. □

4. Sufficient Conditions for Strong Stability I

The verification of strong stability of a given decision problem (L_0, ℓ_0, P_0) can be tedious and fairly involved. It is, therefore, natural to seek sufficient conditions on L_0, ℓ_0 and P_0 which guarantee the stability of the problem. In this section, several results are stated which give such sufficient conditions for SSI. Focusing on decision problems which are common in statistics, the finite decision problem followed by the estimation problem and a general decision problem are treated.

Clearly, the notion of strong stability requires the convergence of optimal Bayes risks of a sequence of priors converging to P_0. To derive the conditions for this in a unified manner, it is convenient to base the analysis on the following general setting and appeal to the results therein.

Suppose $\psi_n : D \to R$ is a sequence of real valued functions. The sequence $\{\psi_n\}$ is said to be epi lower semi continuous (ELSC) at $\psi_0 : D \to R$ if, for every $d \in D$ and every sequence $d_n \to d$, $\psi_0(d) \leq \liminf_{n\to\infty} \psi_n(d_n)$.

A characterization of the convergence of $\inf_d \psi_n(d)$ to $\inf_d \psi_0(d)$, suitable for the present discussion, is as follows. See Attouch (1984), Attouch and Wets (1981), Salinetti and Wets (1986), and Dupecova and Wets (1987) for more general results in this regard.

Theorem 4.1. *Let $\psi_n : D \to R$, $n = 0, 1, 2\ldots$ be a family of functions such that $\{\psi_n\}$ is ELSC at ψ_0. Suppose $\psi_n(d_0) \to \psi_0(d_0)$ for some $d_0 \in \arg\min \psi_0$. Then $\inf_d \psi_n(d) \to \inf_d \psi_0(d)$ if and only if for every $\epsilon > 0$ there exists a compact set $K \subset D$ and a sequence $d_n(\epsilon) \in K$ for all large n such that*

$$\psi_n(d_n(\epsilon)) \leq \inf_d \psi_n(d) + \epsilon \qquad \ldots (4.1)$$

For a proof of this theorem see Salinetti (1994).

From now on, the $d_n(\epsilon)$ satisfying (4.1) will be called the ϵ-optimal solution of ψ_n.

In all the applications of this theorem that follow, the roles of $\psi_n(d)$ and $\psi_0(d)$ will be played by $\int L_0(\theta, d)\ell_0(\theta)dP_n$ and $\int L_0(\theta, d)\ell_0(\theta)dP_0$ respectively. It is worth, therefore, noting that the theorem requires the existence of the optimal solution d_o for the problem (L_0, ℓ_0, P_0). Most statistically interesting decision problems meet this condition. There are, however, situations where the optimal solutions do not exist and, to tackle such problems, one needs more general versions of the preceding theorem. But, for the sake of simplicity, this is not discussed here.

Consider the standard finite decision statistical problem where Θ is an open subset of R^m and $D = \{d_1, \ldots, d_k\}$. Assume the loss $L_0(\theta, d)$ is bounded. This framework includes the standard finite action statistical problems like the test of hypotheses and the monotone decision problem.

Let $D_{L_0} = \{\theta : \theta \text{ is a discontinuity point of } L_0(\cdot, d_i) \text{ for some } d_i \in D\}$. The following result relates to the SSI of (L_0, ℓ_0, P_0) for finite D.

THEOREM 4.2. *Let (L_0, ℓ_0, P_0) be a finite decision problem. Suppose $L_0(\theta, d)\ell_0(\theta)$ is bounded in θ for each d and $P_0(D_{L_0}) = 0$. Under the uniform convergence of losses, (L_0, ℓ_0, P_0) is strongly stable I.*

PROOF. Let $P_n \Rightarrow P_0$. Set $\psi_n(d) = \int L_0(\theta, d)\ \ell_0(\theta) dP_n$ and $\psi_0(d) = \int L_0(\theta, d)\ell_0(\theta) dP_0$. Let d_0 be an optimal solution of ψ_0. Plainly, since $P_0(D_{L_0}) = 0$, $\psi_n(d_0) \to \psi_0(d_0)$. Moreover, since D is finite and endowed with discrete topology, for every d and every sequence $d_n \to d$, $\psi_n(d_n) \to \psi_0(d)$ (i.e. ψ_n is ELSC at ψ_0) in view of the assumption $P_0(D_{L_0}) = 0$. Therefore the conditions of Theorem 4.1 are satisfied and $\inf_d \psi_n(d) \to \inf_d \psi_0(d)$. Also, if $d_0(\epsilon)$ is an ϵ-optimal solution of ψ_0 (i.e. of (L_0, ℓ_0, P_0)). Again, by the assumption $P_0(D_{L_0}) = 0$,

$$\begin{aligned} &\lim_{n\to\infty} \int (L_0(\theta, d_0(\epsilon)) - L_0(\theta, d_0))\ell_0(\theta) dP_n \\ &= \int (L_0(\theta, d_0(\epsilon)) - L_0(\theta, d_0))\ell_0(\theta) dP_0 < \epsilon \end{aligned}$$

Putting these facts together it follows

$$\lim_{\epsilon\downarrow 0} \limsup_{n\to\infty} \int L_0(\theta, d_0(\epsilon))\ell_0(\theta) dP_n - \inf_d \int L_0(\theta, d)\ell_0(\theta) dP_n = 0.$$

This concludes the proof. □

Typically, in monotone multiple decision problems, D_{L_0} is either finite or a lower dimensional set. Consequently the condition is satisfied if P_0 is continuous. However, if P_0 happens to be discrete and $L_0(\cdot, d_0)$ is P_0-continuous for some optimal d_0 then (L_0, ℓ_0, P_0) is at least weakly stable I with stabilizing decision d_0.

The condition $P_0(D_{L_0}) = 0$ is too strong for multiple decision or ranking problems where the parameter space is also finite. The following theorem addresses such cases.

Let $D = \{d_1, \ldots, d_k\}$ and $D_0 = \{\theta : \text{a discontinuity point of } L_0(\cdot, d_0) \text{ for some } (L_0, \ell_0, P_0) \text{ optimal decision } d_0\}$.

THEOREM 4.3. *Assume $L_0(\theta, d)$ is lower semicontinuous in θ for every $d \in D$ and $P_0(D_0) = 0$. Then (L_0, ℓ_0, P_0) is strongly stable I.*

PROOF. Let $P_n \Rightarrow P_0$ and, $\psi_n(d) = \int L_0(\theta, d)\ell_0(\theta)dP_n$, $\psi_0(d) = \int L_0(\theta, d)\ell_0(\theta) dP_0$. Since $P_0(D_0) = 0$, $\psi_n(d_0) \to \psi_0(d_0)$ for any (L_0, ℓ_0, P_0) optimal decision d_0. Moreover, by the lower semicontinuity of $L_0(\theta, d)$ in θ for every $d \in D$, it follows for every $d \in D$ and $d_n \to d$

$$\liminf_{n\to\infty} \psi_n(d_n) \geq \psi_0(d) \qquad \ldots (4.2)$$

because D, being finite, is endowed with discrete topology. Hence, by Theorem 4.1, $\inf_d \psi_n(d) \to \inf_d \psi_0(d)$. Finally, for all sufficiently small $\epsilon > 0$, $d_0(\epsilon) \equiv d_0$ for some (L_0, ℓ_0, P_0) optimal decision because D is finite and therefore, in view of $P_0(D_0) = 0$,

$$\lim_{n\to\infty} \int (L_0(\theta, d_0(\epsilon)) - L_0(\theta, d_0))\ell_0(\theta)dP_n = 0$$

This completes the proof of the theorem. □

The following example involving a multiple monotone decision problem illustrates the difference between the Theorems 4.2 and 4.3.

EXAMPLE 4.1. Consider the finite decision problem with $D = \{a_1, a_2, \ldots, a_k\}$ and $\Theta = R$ with loss L_0 given by

$$L_0(\theta, a_i) = \begin{cases} 0 & \text{if } \theta_{i-1} < \theta < \theta_i \\ 1 & \text{otherwise} \end{cases}$$

for $i = 1, 2, \ldots, k$ where $-\infty = \theta_0 < \theta_1 < \ldots, < \theta_{k-1} < \theta_k = \infty$ are some distinguished points. Clearly, L_0 is not continuous in θ and it is not even lower semicontinuous because $\{\theta : L_0(\theta, a_i) > \frac{1}{2}\} = (-\infty, \theta_{i-1}] \cup [\theta_i, \infty)$, a closed set. Hence, Theorem 4.3 is not applicable. However, the problem (L_0, ℓ_0, P_0) is SSI by Theorem 4.2 provided $P_0(D_{L_0}) = P\{\theta_1, \theta_2, \ldots \theta_{k-1}\} = 0$.

On the other hand, if the loss were

$$L_1(\theta, a_i) = \begin{cases} 0 & \text{if } \theta_{i-1} \leq \theta \leq \theta_i \\ 1 & \text{otherwise} \end{cases}$$

for $i = 1, 2, \ldots, k$, then it is easy to see L_1 is lower semicontinuous in θ and, therefore, (L_1, ℓ_0, P_0) is SSI by Theorem 4.3 provided $P_0(D_0) = 0$. Here D_0 is the set of discontinuity points of $L_1(\cdot, d_0)$.

If $P_0(D_0) > 0$ then (L_1, ℓ_0, P_0) is not strongly stable. Indeed, if $P_0(D_0) > 0$, one can construct, using a routine argument, a sequence $P_n \Rightarrow P_0$ such that

$$\int L_1(\theta, d_0)\ell_0(\theta)dP_n \not\to \int L_1(\theta, d_0)\ell_0(\theta)dP_0$$

for some d_0 and this in turn implies (L_1, ℓ_0, P_0) is not SSI. Hence $P_0(D_0) = 0$ is necessary and sufficient for stability, given the other assumptions. $\square$

The next result treats loss functions which arise typically in point estimation problems. It is a straightforward consequence of a theorem due to Salinetti (1994, Proposition 3.3).

THEOREM 4.4. *Suppose $L_0(\theta, d)\ell_0(\theta)$ is jointly lower semicontinuous (in θ and d) and for each d, bounded continuous in θ. Then (L_0, ℓ_0, P_0) is SSI if for all $P_n \Rightarrow P_0$ and for every $\epsilon > 0$ there exists a bounded sequence $\{d_n(\epsilon)\}$ of ϵ-optimal solutions for (L_0, ℓ_0, P_n).*

A few comments are in order before the proof. It is not difficult to show the existence of bounded ϵ-optimal solutions for bowl-shaped loss functions. Clearly, it is enough to show that the optimal decisions $\{d_n\}$ of (L_0, ℓ_0, P_n) are bounded and this is indeed the case (see Lemma 3.5) for bowl shaped loss functions. Towards the end of the section this issue is addressed and a few sufficient conditions (on L_0) are listed for the existence of bounded ϵ-optimal decisions.

PROOF OF THEOREM 4.4. A brief sketch of the proof is given and the reader is referred to Salinetti (1994) or Attouch (1986) for various technical details. Let $d_0(\epsilon)$ be an ϵ-optimal decision of (L_0, ℓ_0, P_0). The assumptions $L_0(\theta, d)\ell_0(\theta)$ is jointly lower semicontinuous and continuous in θ for each d imply $(L_0(\theta, d) - L_0(\theta, d_0(\epsilon)))\ell_0(\theta)$ is jointly lower semicontinuous, (in θ and d) and, therefore for every d, for all $d_n \to d$ and for $P_n \Rightarrow P_0$

$$\begin{aligned} &\liminf_{n\to\infty} \textstyle\int (L_0(\theta, d_n) - L_0(\theta, d_0(\epsilon)))\ell_0(\theta) dP_n \\ &\geq \textstyle\int (L_0(\theta, d) - L_0(\theta, d_0(\epsilon)))\ell_0(\theta) dP_0 \end{aligned} \qquad \ldots (4.3)$$

i.e. $\psi_n(d) \equiv \int L_0(\theta, d) - L_0(\theta, d_0(\epsilon))\ell_0(\theta) dP_n$ is ELSC at $\psi_0(d) \equiv \int (L_0(\theta, d) - L_0(\theta, d_0(\epsilon)))\ell_0(\theta) dP_0$. Moreover, $\psi_n \to \psi_0$ pointwise in view of bounded continuity of $L_0(\theta, d)\ell_0(\theta)$. Hence, by Theorem 4.1,

$$\liminf_{n\to\infty} \inf_d \int (L_0(\theta, d) - L_0(\theta, d_0(\epsilon)))\ell_0(\theta) dP_n > -\epsilon \qquad \ldots (4.4)$$

in view of the assumption there exists a bounded sequence of ϵ-optimal decisions for (L_0, ℓ_0, P_n). Finally, to conclude the proof, note that SSI of (L_0, ℓ_0, P_0) follows from (4.4) by letting $\epsilon \to 0$. $\square$

REMARK. If the condition "$L_0(\theta, d)\ell_0(\theta)$ is continuous in θ for every d" is weakened by assuming "$L_0(\theta, d_0)\ell_0(\theta)$ is bounded continuous in θ" where d_0 is (L_0, ℓ_0, P_0) optimal, then one can still conclude (L_0, ℓ_0, P_0) is weakly stable I.

The existence of bounded ϵ-optimal solutions depends on the properties of the loss function L_0. It is possible to list a variety of sufficient conditions for this, but the following general result covers most of the commonly used loss functions.

PROPOSITION 4.1. *Suppose $L_0(\theta, d)$ satisfies the following conditions*

(i) For some $\beta > 0$, for every compact $C \subset \Theta$ there exists a compact set $K \subset D$ such that for every $d' \notin K$ there exists $d \in K$ with the property

$$L_0(\theta, d) < L_0(\theta, d') - \beta$$

for all $\theta \in C$.

(ii) For every compact set $K \subset D$, there exists $M > 0$ such that

$$\sup_{\theta} \sup_{d \in K} L_0(\theta, d)\ell_0(\theta) \ < \ M.$$

Then for every sequence $P_n \Rightarrow P_0$ there exists a bounded sequence of ϵ-optimal solutions for (L_0, ℓ_0, P_n).

PROOF. Since $P_n \Rightarrow P_0$, for $\gamma = \frac{\beta}{M+\beta+1}$ there exists a compact $C_0 \subset \Theta$. Satisfying $P_n(C_0) > 1-\gamma$, for all n. By condition (i), there is a compact $K_0 \subset D$ corresponding to C_0 and for any $d' \notin K_0$ there exists a $d \in K_0$ such that

$$\int_{C_0} L_0(\theta, d)\ell_0(\theta)dP_n < \int L_0(\theta, d')\ell_0(\theta)dP_n - \beta(1-\gamma) \qquad \ldots (4.5)$$

Hence

$$\int L_0(\theta, d)\ell_0(\theta)dP_n < \int L_0(\theta, d')\ell_0(\theta)dP_n + [\int_{\Theta - C_0} L_0(\theta, d)\ell_0(\theta)dP_n - \beta(1-\gamma)] \qquad \ldots (4.6)$$

Now, since $d \in K_0$, by condition (iii) the second term in the right side of (4.6) is bounded above by $M\gamma - \beta(1-\gamma) < 0$. Thus, for any $d' \notin K_0$ there exists a $d \in K_0$ and

$$\int L_0(\theta, d)\ell_0(\theta)dP_n < \int L_0(\theta, d')\ell_0(\theta)dP_n$$

for all n. This clearly implies the existence of a bounded sequence of ϵ-optimal solutions for (L_0, ℓ_0, P_n). □

While the condition (ii) of the proposition is easy to check, the verification of (i) may involve some work depending on the complexity of L_0. It is, however, quite straightforward to show that bowl shaped loss functions satisfy this assumption. Finally, a useful and easy to verify property which implies the condition (i) is as follows.

PROPERTY P: "For every compact $C \subset \Theta$, there exists $\delta_0 > 0$ and a strictly increasing unbounded positive function $\phi(\| d \|)$ such that

$$\inf_{\theta \in C} L_0(\theta, d) > \phi(\| d \|)$$

for all $d \in D$ with $\| d \| > \delta_0$."

Here $\| d \|$ is the norm of d. It is well known that many of the unbounded loss functions commonly used in decision theoretic estimation and prediction satisfy this property.

Reverting to Theorem 4.3, the assumption "$L_0(\theta, d)$ is continuous in θ for each d" deserves some comments. Even when this assumption is violated the theorem holds provided P_0 assigns probability zero to the set of discontinuities D_{L_0}. On the other hand, if $P_0(D_{L_0}) > 0$ the decision problem (L_0, ℓ_0, P_0) fails to be SSI. The following example illustrates these points. Another interesting feature of this example is the loss function, appropriate for fixed "width" credible region problem (i.e. the Bayesian analogue of fixed "width" confidence set problem), does not satisfy the condition (i) of Proposition 4.4 and, yet, it has a bounded sequence of ϵ-optimal decisions.

EXAMPLE 4.2. Let $\Theta = D = R^m$ and $A \subset R^m$ be a closed bounded set with nonempty interior containing the origin. Consider the decision problem (L_0, ℓ_0, P_0) where $\ell_0(\theta)$ is a bounded continuous positive likelihood function and $L_0(\theta, d) = 1 - I_A(\theta - d)$. By absorbing $\ell_0(\theta)$ into P_0 and renormalizing the resulting measure, assume without loss of generality $\ell_0(\theta) \equiv 1$. The loss L_0 is jointly lower semicontinuous because A is closed.

Now, let $P_n \Rightarrow P_0$. The tightness of $\{P_n\}$ implies that there exists a compact set C with $P_n(C) > 1 - \epsilon$ for all n, where ϵ is a fixed small positive number, and $A \subset C$. Since A has nonempty interior, it follows that there exist $\{d_1, d_2, \ldots, d_k\} \subset C$ so that $C \subset \cup_{i=1}^{k} \{A + d_i\}$. Therefore, there is a decision $d^* \epsilon \{d_1, \ldots, d_k\}$ with the property $P_n(A + d^*) > \frac{1-\epsilon}{k}$. Note that the decision d^* may depend on P_n but this is of no consequence. The important fact here is $d^* \in C$ for every P_n. Finally, let $K_1 \subset K$ be two compact sets such that $C \subset K_1$, $P_n(K_1^C) < \frac{1}{100}(\frac{1-\epsilon}{k})$ and $d' \notin K \Rightarrow A + d' \subset K_1^C$. Then, for any $d' \notin K$,

$$\begin{aligned} \int L_0(\theta, d') dP_n &= 1 - P_n(A + d') > 1 - \frac{(1-\epsilon)}{k} \\ &= \int L_0(\theta, d^*) dP_n \end{aligned}$$

for all n. Thus, there is a compact $K \subset D$ with the property "for any $d' \notin K$ there exists a $d \in k$ satisfying, $\int L_0(\theta, d) dP_n \leq \int L_0(\theta, d') dP_n$". This implies, since K does not depend on n, the existence of a bounded sequence of ϵ-optimal decisions for (L_0, ℓ_0, P_n). Incidentally, one can also conclude using the above argument that (L_0, ℓ_0, P_0) has an optimal (i.e. Bayes optimal) decision d_0.

Suppose now $P_0(\partial A + d_0) = 0$, where ∂A is the topological boundary of A. Then $L_0(\theta, d_0)$ is P_0-continuous and, by virtue of the lower semicontinuity of $L_0(\theta, d)$, it follows $\psi_n(d) = \int (L_0(\theta, d) - L_0(\theta, d_0)) dP_n$ is ELSC. This, in conjunction with $\psi_n(d_0) \equiv 0$ and (L_0, ℓ_0, P_n)'s have a bounded sequence of ϵ-optimal decisions, implies by Theorem 4.1 that

$$\inf_d \int L_0(\theta, d) dP_n \to \inf_d \int L_0(\theta, d) dP_0$$

for every sequence $P_n \Rightarrow P_0$. Furthermore, if $P_0(\partial A + d) = 0$ for all d then, in view of (4.7), it is easy to see that (L_0, ℓ_0, P_0) is SSI.

Thus, there are the following three possibilities in this example.

(i) If $P_0(\partial A + d) = 0$, $\forall d$ then (L_0, ℓ_0, P_0) is SSI.

(ii) If $P_0(\partial A + d_0) = 0$ then (L_0, ℓ_0, P_0) is Weakly Stable I with stabilizing decision d_0.

(iii) If $P_0(\partial A + d_0) > 0$ then (L_0, ℓ_0, P_0) may be Weakly Stable I but not SSI.

In conclusion, because ∂A is set of lower dimension, a credible region problem is SSI when P_0 is absolutely continuous.

5. Sufficient Conditions for Strong Stability II

The notion of Strong Stability II, as noted earlier, is fairly stringent relative to SSI and this section gives sufficient conditions for a decision problem to satisfy SSII. Following the pattern in Section 4, the stress is on decision problems relevant to statistics.

Proposition 5.1. *Suppose D is finite and $P_0(D_{L_0}) = 0$. Then (L_0, ℓ_0, P_0) is strongly stable II.*

Proof. Observe that, since $P_0(D_{L_0}) = 0$, for any sequence $Q_n \Rightarrow P_0$

$$\lim_{n\to\infty} \max_{d\in D} \left| \int L_0(\theta, d)\ell_0(\theta) dQ_n - \int L_0(\theta, d)\ell_0(\theta) dP_0 \right| = 0$$

Consequently, if $d_{Q_n}(\epsilon)$ is an ϵ-optimal decision of (L_0, ℓ_0, Q_n), and $P_n \Rightarrow P_0$, one can conclude

$$\lim_{n\to\infty} \left| \int L_0(\theta, d_{Q_n}(\epsilon))\ell_0(\theta) dQ_n - \int L_0(\theta, d_{Q_n}(\epsilon))\ell_0(\theta) dP_0 \right| = 0 \qquad \ldots (5.1)$$

and

$$\lim_{n\to\infty} \left| \int L_0(\theta, d_{Q_n}(\epsilon))\ell_0(\theta) dP_n - \int L_0(\theta, d_{Q_n}(\epsilon))\ell_0(\theta) dP_0 \right| = 0 \qquad \ldots (5.2)$$

Moreover, the conditions of the proposition imply, by Theorem 4.2, that the decision problem (L_0, ℓ_0, P_0) is SSI. Therefore

$$\begin{aligned}
&\lim_{\epsilon\downarrow 0}\limsup_{n\to\infty}\left(\int L_0(\theta,d_{Q_n}(\epsilon))\ell_0(\theta)dP_n-\inf_d\int L_0(\theta,d)\ell_0(\theta)dP_n\right)\\
&\le \lim_{\epsilon\downarrow 0}\limsup_{n\to\infty}\left[\mid\int L_0(\theta,d_{Q_n}(\epsilon))\ell_0(\theta)dP_n-\int L_0(\theta,d_0(\epsilon))\ell_0(\theta)dP_0\mid\right]\\
&\le \lim_{\epsilon\downarrow 0}\limsup_{n\to\infty}\left[\mid\int L_0(\theta,d_{Q_n}(\epsilon))\ell_0(\theta)dP_n-\int L_0(\theta,d_{Q_n}(\epsilon))\ell_0(\theta)dP_0\mid\right.\\
&\quad+\mid\int L_0(\theta,d_{Q_n}(\epsilon))\ell_0(\theta)dP_0-\int L_0(\theta,d_{Q_n}(\epsilon))\ell_0(\theta)dQ_n\mid\\
&\quad\left.+\mid\int L_0(\theta,d_{Q_n}(\epsilon))\ell_0(\theta)dQ_n-\int L_0(\theta,d_{Q_n}(\epsilon))\ell_0(\theta)dP_0\mid\right]\\
&\le \lim_{\epsilon\downarrow 0}\limsup_{n\to\infty}\left[\mid\int L_0(\theta,d_{Q_n}(\epsilon))\ell_0(\theta)dQ_n-\int L_0(\theta,d_0(\epsilon))\ell_0(\theta)dP_0\right]=0
\end{aligned}$$

by (5.1), (5.2) and the strong stability I of (L_0,ℓ_0,P_0). This proves (L_0,ℓ_0,P_0) is strongly stable II. □

REMARKS. (1) The above proposition also establishes the equivalence of the two definitions of stability for the finite-decision problem under the assumption of $P_0(D_{L_0})=0$.

(2) The condition $P_0(D_{L_0})=0$ is also nearly necessary in the sense if d_0 is (L_0,ℓ_0,P_0) optimal and $P_0(\theta : L_0(\theta,d_0)$ is discontinuous$)>0$ then the decision problem is not strongly stable according to either definition.

The extensions of Proposition 5.1 to more general loss functions require the following result due to Billingsley and Topsoe (1966), and Topsoe (1967). To state the result, define the oscillation $W_d(A)$ of $L_0(\theta,d)\ell_0(\theta)$ on a set A by

$$W_d(A)=\sup\{\mid L_0(\theta_1,d)\ell_0(\theta_1)--L_0(\theta_2,d)\ell_0(\theta_2)\mid:\theta_1,\theta_2\in A\}$$

Also, let $W_d(\theta;\delta)\equiv W_d(S(\theta;\delta))$ where $S(\theta;\delta)$ is the ball of radius δ centered at θ.

THEOREM 5.1 *(Billingsley and Topsoe). For every sequence $P_n\to P_0$,*

$$\lim_{n\to\infty}\sup_{d\in D}\mid\int L_0(\theta,d)\ell_0(\theta)dP_n-\int L_0(\theta,d)\ell_0(\theta)dP_0\mid=0$$

if, and only if

(i) $\sup_{d\in D}W_d(\Theta)<\infty$

(ii) $\lim_{\delta\downarrow 0}\sup_{d\in D}\int W_d(\theta;\delta)dP_0=0.$

See Bhattacharya and Rao (1976) for an excellent discussion about this theorem as well as its proof.

The following result, a generalization of Proposition 5.1 to bounded loss functions, is an immediate consequence of Theorem 5.1.

THEOREM 5.2. *Suppose $L_0(\theta,d)\ell_0(\theta)$ is bounded in θ and d, and $\lim_{\delta\downarrow 0}\sup_d\int W_d(\theta;\delta)dP_0=0$. Then (L_0,ℓ_0,P_0) is strongly stable II.*

PROOF. Let $P_n \Rightarrow P_0$, $Q_n \Rightarrow P_0$ and $d_{Q_n}(\epsilon)$ be ϵ-optimal decisions of (L_0, ℓ_0, Q_n). Then

$$\begin{aligned}
\int & \quad L_0(\theta, d_{Q_n}(\epsilon))\ell_0(\theta)dP_n - \inf_d \int L_0(\theta, d)\ell_0(\theta)dP_n \\
= & \quad \sup_{d\in D}\left[\int L_0(\theta, d_{Q_n}(\epsilon))\ell_0(\theta)dP_n - \int L_0(\theta, d)\ell_0(\theta)dP_n\right] \\
\leq & \quad \left| \int L_0(\theta, d_{Q_n}(\epsilon))\ell_0(\theta)dP_n - \int L_0(\theta, d_{Q_n}(\epsilon))\ell_0(\theta)dQ_n \right| \qquad \ldots(5.3) \\
+ & \quad \left[\int L_0(\theta, d_{Q_n}(\epsilon))\ell_0(\theta)dQ_n - \inf_d \int L_0(\theta, d)\ell_0(\theta)dQ_n\right] \\
+ & \quad \sup_d \left| \int L_0(\theta, d)\ell_0(\theta)dP_n - \int L_0(\theta, d)\ell_0(\theta)dQ_n \right|
\end{aligned}$$

Now taking lim sup over n it follows from Theorem 5.1 that the first and the third terms are zero. Moreover, by the definition of $d_{Q_n}(\epsilon)$, the second term is bounded above by ϵ. Finally, let $\epsilon \to 0$ to complete the proof. □

As a corollary to the theorem we have the following result of Kadane and Chuang (1978; Theorem 2.7).

COROLLARY 5.1. *Suppose* $L_0(\theta, d)\ell_0(\theta)$ *is bounded in* (θ, d) *and* $\{L_0(\theta, d)\ell_0(\theta) : d \in D\}$ *is equicontinuous in* θ *at every* $\theta \in \Theta$. *Then* (L_0, ℓ_0, P_0) *is strongly stable II for every* P_0.

PROOF. This is an immediate consequence of the fact equicontinuity of $\{L_0(\theta, d)\ell_0(\theta) : d \in D\}$ implies $\lim_{\delta\downarrow 0}\sup_{d\in D}\int W_d(\theta;\delta)dP_0 = 0$ for every P_0. For details, see Bhattacharya and Rao (1978; Corollary 2.7, Ch. 1). □

Kadane and Chuang (1978) proved this result under the assumption $L_0(\theta, d)\ell_0(\theta)$ is continuous in θ uniformly in d. Clearly, this assumption implies equicontinuity.

A few remarks are in order about the condition $(*) : \lim_{\delta\downarrow 0}\sup_d \int W_d(\theta;\delta)dP_0 = 0$ in the above theorem. Simple continuity of $L_0(\theta, d)\ell_0(\theta)$ in θ will not in general guarantee this condition. As mentioned earlier, this condition follows from equicontinuity of $\{L_0(\theta, d)\ell_0(\theta) : d \in D\}$.

Another sufficient condition for $(*)$ is a local Lipschitz continuity in the following sense: there exist a function $\zeta(\theta, d)$ and $\alpha > 0$ such that for every $\theta' \in B(\theta;\epsilon)$

$$| L_0(\theta, d)\ell_0(\theta) - L_0(\theta', d)\ell(\theta') | \leq \zeta(\theta, d) \parallel \theta - \theta' \parallel^{\alpha} \qquad \ldots(5.4)$$

and $\sup_d \int \zeta(\theta, d)dP_0 < \infty$.

Many of the smooth bowl shaped bounded loss functions commonly used in estimation and prediction possess this type of Lipschitz continuity.

On the other hand, $(*)$ may hold for some P_0 even when the loss is discontinuous. For example, if $L_0(\theta, d) = 1 - I_C(\theta - d)$ where C is a symmetric, bounded and closed convex set centered at the origin, $(*)$ holds provided P_0 is absolutely continuous.

The next result gives sufficient conditions for SSII in the case where the loss is not bounded.

THEOREM 5.3. *Suppose $L_0(\theta, d)$ satisfies the following conditions:*

1. $L_0(\theta, d)$ is jointly continuous in θ and d.

2. For each $d, L_0(\theta, d)l_0(\theta)$ is bounded.

3. For every compact set $C \subset \Theta$ and every $M > 0$ there exists a compact set $K \subset D$ such that

$$\inf_{\theta \in C} L_0(\theta, d) > M$$

for all $d \notin K$.

Then the decision problem (L_o, l_0, P_0) is strongly stable II.

PROOF. Let $Q_n \Rightarrow P_0$ and $\{d_{Q_n}(\epsilon)\}$ be a sequence of ϵ-optimal solutions of (L_0, l_0, Q_n). The assumption (iii) implies that there exists a compact set $K_0 \subset D$ such that $d_{Q_n}(\epsilon) \in K_0$ for all large n. Hence, the conditions of Theorem 4.4 are satisfied and (L_0, l_0, P_0) is strongly stable I.

Now let $P_n \Rightarrow P_0$. Then, for $\gamma > 0$ (to be chosen later) there exists a compact set $C \subset \Theta$ satisfying $P_n(C) > 1 - \gamma, Q_n(C) > 1 - \gamma$ for all n. Set $B = \sup_{d \in K} \sup_{\theta} L_0(\theta, d)l_0(\theta)$. Choose a compact set C_1 such that $C \subset C_1 \subset \Theta$ and $P_n(C_1) > 1 - \gamma/B$ and $Q_n(C_1) > 1 - \gamma/B$ for all n. Then, by the compactness of C_1 and $K_0, L_0(\theta, d)$ is uniformly continuous on $C_1 \times K_0$ and

$$\sup_{d \in K_0} |\int_{C_1} L_0(\theta, d)l_0(\theta)dP_n - \int_{C_1} L_0(\theta, d)l_0(\theta)dP_0| \to 0 \qquad \ldots (5.5)$$

as $n \to \infty$ by Theorem 5.1. Therefore, since $d_{Q_n}(\epsilon) \in K_0$ it follows

$$\begin{aligned} & \limsup_{n\to\infty} |\int L_0(\theta, d_{Q_n}(\epsilon)l_0(\theta)dP_n - \int L_0(\theta, d_{Q_n}(\epsilon))l_0(\theta)dP_0| \\ \leq \quad & \limsup_{n\to\infty} |\int_{C_1} L_0(\theta, d_{Q_n}(\epsilon))l_0(\theta)dP_n - \int_{C_1} L_0(\theta, d_{Q_n}(\epsilon))l_0(\theta)dP_0| \\ + \quad & B[1 - P_n(C_1) + 1 - P_0(C_1)] \leq 2\gamma \end{aligned} \qquad \ldots (5.6)$$

Also note the above statements hold for the sequence $\{Q_n\}$. Therefore by triangle inequality

$$\limsup_{n\to\infty} |\int L_0(\theta, d_{Q_n}(\epsilon))l_0(\theta)dP_n - \int L_0(\theta, d_{Q_n}(\epsilon))l_0(\theta)dQ_n| \leq 4\gamma \quad \ldots (5.7)$$

Finally, to complete the proof, observe that since $\gamma > 0$ is arbitrary, (2.3) follows from (5.7) and Strong Stability I of (L_0, l_0, P_0) $\square$

It is possible to state and prove a variant of the above theorem for the case where $L_0(\theta, d)$ is jointly lower semi continuous. Such a result, however, will involve some additional conditions.

References

Attouch, H. (1984). *Variational Convergence for Functions and Operators.* Pitman Applicable Mathematics Series.

Attouch, H. and Wets, R. (1981). Approximation and convergence in nonlinear optimization. *Nonlinear Programming* (Mangesarian, Meyer and Robinson, eds.), **4**, 367-394. Academic Press, New York.

Berger, J. (1984). The robust Bayesian viewpoint. In *Robustness of Bayesian Analysis* (J. Kadane, ed.). North Holland, New York.

Bhattacharya, R.N. and Ranga Rao, R. (1976). *Normal Approximation and Asymptotic Expansions.* John Wiley & Sons.

Billingsley, P. and Topsoe, F. (1967). Uniformity in weak convergence. *Zeit. Wahr. Ver. Geb.* **7** 1-16.

Britney, R. and Winkler, R. (1974). Bayesian point estimation and prediction. *Ann. Inst. Stat. Math.* **26** 15-34.

Chuang, D.T. (1984). Further theory of stable decisions. In *Robustness of Bayesian Analysis* (J.Kadane, ed.). North Holland, New York.

Dupecova, J. and Wets, R. (1988). Asymptotic behavior of statistical estimators and optimal solutions of stochastic optimization problems. *Ann. Stat.* **16** 1517-1549.

Edwards, W., Lindemen, H., and Savage, L.J. (1963). Bayesian statistical inference for psychological research. *Psychological Review.* **70** 193-242.

Ekeland, I. and Temam, R. (1979). *Convex Analysis and Variational Problems.* North Holland, Amsterdam.

Fishburn, P., Murphy, A. and Isaacs, H. (1967). Sensitivity of decision to probability estimation errors: a reexamination. *Op. Res.* **15** 254-267.

Kadane, J. and Chuang, D.T. (1978). Stable decision problems. *Ann. Stat.* **6** 1095-1110.

Kelly, J.L. (1955). *General Topology.* D. Van Nostrand, Princeton.

Rockafellar, R.T. (1970). *Convex Analysis.* Princeton University Press.

Stone, M. (1963). Robustness of nonideal decision procedures. *J. Amer. Statist. Assoc.* **58** 480-486.

Salinetti, G. and Wets, R. (1986). Convergence of infima, especially stochastic infima. *Technical Report*, Univ. Rome "La Sapienza".

Salinetti, G. (1994). Stability of Bayesian decisions. *J. Statist. Plan. Inf.* **40** 313-329.

Stone, M. (1963). Robustness of nonideal decision procedures. *J. Amer. Statist. Assoc.* **58** 480-486.

Topsoe, F. (1968). Preservation of weak convergence under mappings. *Ann. Math. Stat.* **39** 1661-1665.

Joseph B. Kadane
Carnegie Mellon University
Department of Statistics
Baker Hall 232E
Pittsburgh, PA 15213
e-mail: kadane@stat.cmu.edu

C. Srinivasan
University of Kentucky
Department of Statistics
Room 869 Patterson Office Tower
Lexington, KY 40506-0027
e-mail: srini@ms.uky.edu