The Annals of Statistics
1978, Vol. 6, No. 5, 1095–1110

STABLE DECISION PROBLEMS[1]

BY JOSEPH B. KADANE AND DAVID T. CHUANG

Carnegie-Mellon University

A decision problem is characterized by a loss function V and opinion H. The pair (V, H) is said to be strongly stable iff for every sequence $F_n \to_\omega H$, $G_n \to_\omega H$ and $L_n \to V$, $W_n \to V$ uniformly,

$$\lim_{\varepsilon \downarrow 0} \limsup_{n\to\infty} [\int L_n(\theta, D_n(\varepsilon))\, dF_n(\theta) - \inf_D \int L_n(\theta, D)\, dF_n(\theta)] = 0$$

for every sequence $D_n(\varepsilon)$ satisfying

$$\int W_n(\theta, D_n(\varepsilon))\, dG_n(\theta) \leqq \inf_D \int W_n(\theta, D)\, dG_n(\theta) + \varepsilon\,.$$

We show that squared error loss is unstable with any opinion if the parameter space is the real line and that any bounded loss function $V(\theta, D)$ that is continuous in θ uniformly in D is stable with any opinion H. Finally we examine the estimation or prediction case $V(\theta, D) = h(\theta - D)$, where h is continuous, nondecreasing in $(0, \infty)$ and nonincreasing in $(-\infty, 0)$ and has bounded growth. While these conditions are not enough to assure strong stability, various conditions are given that are sufficient. We believe that stability offers the beginning of a Bayesian theory of robustness.

1. Introduction. "Subjectivists should feel obligated to recognize that any opinion (so much more the initial one) is only vaguely acceptable. (I feel that objectivists should have the same attitude.) So it is important not only to know the exact answer for an exactly specified initial position, but what happens changing in a reasonable neighborhood the assumed initial opinion." De Finetti, as quoted by Dempster (1975).

A well-known principle of personalistic Bayesian theory is that no one can tell someone else what loss function to have or what opinion to hold. Having said that, the reasons for looking into properties of particular choices of loss functions and opinions might be obscure.

The standard of personalistic Bayesian theory may be too severe for many of us. Generally when a personalistic Bayesian tells you his loss function and opinion, he means them only approximately. He hopes that his approximation is good, and that whatever errors he may have made will not lead to decisions with loss substantially greater than he would have obtained had he been able to write down his true loss function and opinion.

There are two special cases that have been considered. In the first, one cannot (or need not) obtain one's exact prior probability. Stone (1963) studied decision

---

Received June 1976; revised June 1977.

[1] This research was supported in part by the Office of Naval Research under Contract No. N00014-75-C-0516 and Task No. NR042-309 and in part by the National Science Foundation. Discussions on the subject with A. P. Dawid, M. H. DeGroot and W. F. Rogers, Jr. were particularly helpful.

AMS 1970 *subject classifications*. Primary 62C10; Secondary 62G35.

Key words and phrases. Decision theory, robustness, stable estimation, stable decisions.

procedures with respect to the use of wrong prior distributions. He emphasized the possible usefulness of nonideal procedures that do not require full specification of the prior probability distribution. Fishburn, Murphy and Isaacs (1967) and Pierce and Folks (1969) also discussed decision making under uncertainty when the decision maker has difficulty in assigning prior probabilities. They outlined six approaches that may be used to assign probabilities. In the second case, one cannot obtain one's exact utility function. Britney and Winkler (1974) have investigated the properties of Bayesian point estimates under loss functions other than the simple linear and quadratic loss functions. They also discussed the sensitivity of Bayesian point estimates to misspecification in the loss function. Schlaifer (1959) and Antelman (1965) discuss relating the utility of the optimal decision to the utility of suboptimal decisions in certain contexts.

The closest related work, however, is the material on stable estimation in Edwards, Lindeman and Savage (1963). They propose that there is data such that the likelihood function will be sufficiently peaked as to dominate the prior distribution. The criterion for robustness is that the densities of various possible posterior distributions are close.

Another important line of comparison is the work on robustness in the classical context, as exemplified for instance, in Andrews et al. (1972), Bickel and Lehmann (1975a, b) and Huber (1972, 1973). While they study how estimates change as a consequence of outliers, we study here how the *worth* of the estimates change.

To give an initial formalization of our question, suppose that the parameter space is $\Theta \subset \mathbb{R}^k$ for some k, and the decision space is $\mathscr{D} \subset \mathbb{R}^l$ for some l. If $F_\infty(\theta)$ is my (approximate) opinion over $\theta \in \Theta$, and $L_\infty(\theta, D)$ my (approximate) loss function, the (approximate) loss of the decision problem to me is

$$W_\infty = \inf_{D \in \mathscr{D}} \int L_\infty(\theta, D)\, dF_\infty(\theta)\,, \tag{1}$$

which is here assumed to be finite. Then for every $\varepsilon > 0$, there is a decision $D_\infty(\varepsilon)$ which is ε-optimal, that is

$$\int L_\infty(\theta, D_\infty(\varepsilon))\, dF_\infty(\theta) \leqq W_\infty + \varepsilon\,. \tag{2}$$

Suppose, however, that my "true" opinion over Θ is on a sequence $F_n(\theta)$ which converges to $F_\infty(\theta)$ in a sense to be specified later. Also suppose that my "true" loss function over Θ is $L_n(\theta, D)$ which converges to $L_\infty(\theta, D)$ again in a sense to be specified later. Then there is a sequence of "true" losses generated by

$$\omega_n = \inf_{D \in \mathscr{D}} \int L_n(\theta, D)\, dF_n(\theta)$$

and a sequence of losses generated by behaving according to the approximate opinion and loss function:

$$\omega_n' = \int L_n(\theta, D_\infty(\varepsilon))\, dF_n(\theta)\,.$$

The worth of knowing the truth is then

$$B_n = \omega_n' - \omega_n$$

which is always nonnegative. Note that B_n is a function of ε, $D_\infty(\varepsilon)$, n, L_n and F_n. Suppose that

$$\lim_{\varepsilon \downarrow 0} \limsup_{n\to\infty} B_n = 0 \tag{3}$$

for every choice of $L_n \to L_\infty$, $F_n \to F_\infty$, and every choice of $D_\infty(\varepsilon)$ satisfying (2). In this case, the pair (L_∞, F_∞) is called strongly stable (by Definition 1). The above definition makes sense since the nonnegativity of B_n implies that, for each ε,

$$\limsup_{n\to\infty} B_n \geqq 0 .$$

Further, as ε decreases to zero, the set of possible choices $D_\infty(\varepsilon)$ is nonincreasing. Thus the possible values of $\limsup_{n\to\infty} B_n$ is monotone and bounded below by zero. Hence the limit in (3) exists.

There are situations in which (3) holds for every choice of $L_n \to L_\infty$ and $F_n \to F_\infty$, but only for some particular choice $D_\infty(\varepsilon)$. In this case, $D_\infty(\varepsilon)$ is called the stabilizing decision, and the pair (L_∞, F_∞) is called weakly stable (by Definition 1). If (L_∞, F_∞) is not stable (either strongly or weakly), it is called unstable.

The motivation for these definitions is that if an opinion and loss function are strongly stable, then small errors in either will not result in substantially worse decisions. If, on the other hand, a Bayesian finds that the loss function and opinion he has written down are unstable, then he may wish to reassess his loss function and opinion to be certain that no errors have been made. When he finds he has written down a loss function and opinion which is weakly but not strongly stable, a Bayesian may choose to make the stabilizing decision to have protection against errors in either the loss function or opinion.

There are a number of interesting and potentially enlightening choices that might be made for the sense of convergence of F_n to F_∞ and L_n to L_∞. In this paper we chose to start with weak convergence in the distribution and uniform convergence in both arguments in the losses. Another choice worthy of study is to take the likelihood function as known and agreed upon, a weakly convergent sequence of priors, and study the resultant sense of convergence in the posterior opinions. The sense of convérgence studied here is the special case in which that agreed-upon likelihood function is flat, which is equivalent to considering fuzziness in the likelihood function on the same footing as fuzziness in the prior. Perhaps the more general sense of convergence is closer yet in spirit to the work of Edwards, Lindeman and Savage (1963).

We also note that uniform convergence in the loss sequence is a very strong assumption. For example, if $L_\infty(\theta, D) = |\theta - D|^p$ for some p, $0 < p \leqq 1$, the sequence $L_n(\theta, D) = |\theta - D|^{p+c/n}$ for some nonzero constant c is a reasonable sequence of loss functions that do not converge uniformly in θ and D to L_∞.

From a more general point of view we can formulate our problem as follows: for every sequence (L_n, F_n) of truths, and every sequence (W_n, G_n) of approximations satisfying

$$L_n \to V\,, \qquad W_n \to V \quad \text{uniformly}$$

and

$$F_n \to_\omega H\,, \qquad G_n \to_\omega H\,,$$

act as if (W_n, G_n) were true and evaluate at L_n, F_n.

Let $D_n(\varepsilon)$ be defined by

$$\int W_n(\theta, D_n(\varepsilon))\, dG_n(\theta) \leqq \inf_D \int W_n(\theta, D)\, dG_n(\theta) + \varepsilon\,. \tag{4}$$

If for every such choice of $D_n(\varepsilon)$,

$$\lim_{\varepsilon\downarrow 0} \limsup_{n\to\infty} \left[\int L_n(\theta, D_n(\varepsilon))\, dF_n(\theta) - \inf_D \int L_n(\theta, D)\, dF_n(\theta)\right] = 0 \tag{5}$$

then (V, H) is strongly stable (by Definition 2). If there is some choice of $D_n(\varepsilon)$ which makes (5) hold, then (V, H) is weakly stable and $D_n(\varepsilon)$ is the stabilizing decision (by Definition 2).

The second definition has the attractive feature that it permits the reader another interpretation: the apparent truth can be on a sequence (L_n, F_n) approaching the fixed truth (V, H). Definition 2 allows both the apparent truth (L_n, F_n) and the actual truth (W_n, G_n) to be sequences, and is thus more general in the sense that any pair (H, V) that is stable by Definition 2 is clearly stable by Definition 1. All theorems in this paper proving stability have been proved for Definition 2 so they apply to Definition 1 as well. However, all counterexamples to stability have been counterexamples by Definition 1. Hence all statements about the stability or instability of pairs (H, V) in this paper apply to both definitions. This observation leads us to conjecture that Definitions 1 and 2 might be equivalent.

Section 2 introduces Definitions 3 and 4 which are apparently simpler than Definition 2, and shows their equivalence to Definitions 1 and 2. Then some simple examples are given. In Section 3, bounded loss functions that are continuous in the right way are examined, and shown to be strongly stable when paired with any opinion. Finally Section 4 takes up estimation (or, equivalently, prediction) loss functions subject to a Lipschitz-condition restraint on growth, and finds some of them strongly stable, and some unstable. To simplify matters, assume the one-dimensional case ($k = l = 1$).

2. A general structure theorem and some examples. In the first part of this section we introduce two more definitions of strong (weak) stability, Definitions 3 and 4, and show their equivalence to Definitions 1 and 2, respectively. The greater simplicity of the new definitions helps to simplify the rest of the paper. Define, for every $\varepsilon > 0$, the decision $D_\infty(\varepsilon)$ as in (2). Then (L_∞, F_∞) is strongly (weakly) stable (by Definition 3) iff for every sequence $F_n \to_\omega F$ and for every (for some) such $D_\infty(\varepsilon)$,

$$\lim_{\varepsilon\downarrow 0} \limsup_{n\to\infty} \left[\int L_\infty(\theta, D_\infty(\varepsilon))\, dF_n(\theta) - \inf_D \int L_\infty(\theta, D)\, dF_n(\theta)\right] = 0\,. \tag{6}$$

Similarly define, for every $\varepsilon > 0$, the decision $D_n(\varepsilon)$ as in (4) but with W_n taken to be V. Then (V, H) is strongly (weakly) stable by Definition 4 iff for every sequence $F_n \to_\omega H$ and $G_n \to_\omega H$ and for every (for some) such $D_n(\varepsilon)$, (5) holds with V substituted for L_n. Thus Definitions 3 and 4 differ from Definitions 1 and 2 in that, for the latter, only the opinions move, while the loss functions stay constant.

THEOREM 1. (a) *(V, H) is strongly (weakly) stable by Definition* 1 iff *(V, H) is strongly (weaky) stable by Definition* 3.

(b) *(V, H) is strongly (weakly) stable by Definition* 2 iff *(V, H) is strongly (weakly) stable by Definition* 4.

PROOF. The proofs of parts (a) and (b) are similar, so only the proof of (b) is discussed in detail. If (V, H) is strongly (weakly) stable by Definition 2, one of the allowable choices for L_n and W_n is $L_n = W_n = V$ for all n. Strong (weak) stability by Definition 4 then follows trivially.

Suppose, then, that (V, H) is strongly (weakly) stable by Definition 4, and suppose that L_n and W_n are arbitrary sequences of loss functions converging uniformly in θ and D to V. Choose $\varepsilon > 0$, and let $D_n(\varepsilon)$ be defined by equation (4). Choose N_1 such that $\forall\, n \geqq N_1$, $|W_n(\theta, D) - V(\theta, D)| < \varepsilon$ for every θ and D, using the uniform convergence of W_n to V. Then

$$\begin{aligned}
&\inf_D \textstyle\int W_n(\theta, D)\, dG_n(\theta) - \inf_D \int V(\theta, D)\, dG_n(\theta)\\
&\quad = \inf_D \textstyle\int W_n(\theta, D)\, dG_n(\theta) - \inf_D \int (V(\theta, D) - W_n(\theta, D) + W_n(\theta, D))\, dG_n(\theta)\\
&\quad \leqq -\inf_D \textstyle\int (V(\theta, D) - W_n(\theta, D))\, dG_n(\theta)\\
&\quad \leqq \sup_D \textstyle\int (W_n(\theta, D) - V(\theta, D))\, dG_n(\theta) < \varepsilon\,.
\end{aligned}$$

Also $\inf_D \int W_n(\theta, D)\, dG_n(\theta) - \inf_D \int V(\theta, D)\, dG_n(\theta) > -\varepsilon$. Then

$$|\inf_D \textstyle\int W_n(\theta, D_n(\varepsilon))\, dG_n(\theta) - \inf_D \int V(\theta, D)\, dG_n(\theta)| < \varepsilon\,.$$

Also $|\int W_n(\theta, D_n(\varepsilon))\, dG_n(\theta) - \int V(\theta, D_n(\varepsilon))\, dG_n(\theta)| < \varepsilon$. By the triangle inequality, we have

$$\textstyle\int V(\theta, D_n(\varepsilon))\, dG_n(\theta) - \inf_D \int V(\theta, D)\, dG_n(\theta) \leqq 3\varepsilon\,.$$

Hence if $D_n(\varepsilon)$ is ε-optimal for (W_n, G_n), it is 3ε-optimal for (V, G_n), for all $n \geqq N_1$. Choose $\delta > 0$. Then by the uniform convergence of L_n to V, $\exists\, N_2$ such that $\forall\, n \geqq N_2$

$$|L_n(\theta, D) - V(\theta, D)| < \delta\,.$$

By exactly the same argument as above, substituting L_n for W_n and F_n for G_n, we have

$$|\inf_D \textstyle\int L_n(\theta, D)\, dF_n(\theta) - \inf_D \int V(\theta, D)\, dF_n(\theta)| < \delta\,.$$

Also $|\int L_n(\theta, D_n(\varepsilon))\, dF_n(\theta) - \int V(\theta, D_n(\varepsilon))\, dF_n(\theta)| < \delta$. Hence $\forall\, \delta > 0\ \exists\, N_2$ such that $\forall\, n \geqq N_2$,

$$|[\int L_n(\theta, D_n(\varepsilon))\, dF_n(\theta) - \inf_D \int L_n(\theta, D)\, dF_n(\theta)] \\ - [\int V(\theta, D_n(\varepsilon))\, dF_n(\theta) - \inf_D \int V(\theta, D)\, dF_n(\theta)]| < 2\delta\,.$$

Thus

$$\limsup_{n\to\infty} [\int L_n(\theta, D_n(\varepsilon))\, dF_n(\theta) - \inf_D \int L_n(\theta, D)\, dF_n(\theta)] \\ = \limsup_{n\to\infty} [\int V(\theta, D_n(\varepsilon))\, dF_n(\theta) - \inf_D \int V(\theta, D)\, dF_n(\theta)]\,.$$

Finally, taking $D_n(\varepsilon)$ defined by (4), $\varepsilon' = 3\varepsilon$ and $D_n'(\varepsilon')$ defined by (4), with V substituted for L_n,

$$\lim_{\varepsilon\downarrow 0} \limsup_{n\to\infty} [\int L_n(\theta, D_n(\varepsilon))\, dF_n(\theta) - \inf_D \int L_n(\theta, D)\, dF_n(\theta)] \\ = \lim_{\varepsilon'\downarrow 0} \limsup_{n\to\infty} [\int V(\theta, D_n'(\varepsilon'))\, dF_n(\theta) - \inf_D \int V(\theta, D)\, dF_n(\theta)] = 0\,.$$

Thus if (V, H) is strongly (weakly) stable by Definition 4, it is strongly (weakly) stable by Definition 2. □

EXAMPLE 1. *Composite hypothesis, composite alternative.* Suppose that there are only two available decisions $\{1, 2\}$, and suppose that V is defined as follows:

$$\begin{aligned} V(\theta, 1) &= 0 \quad \text{and} \quad V(\theta, 2) = b \quad \text{if} \quad \theta \leqq a\,; \\ V(\theta, 1) &= c \quad \text{and} \quad V(\theta, 2) = 0 \quad \text{if} \quad \theta > a\,, \end{aligned}$$

where b and c are assumed to be positive.

Since our purpose is to show a counterexample to stability, we temporarily adopt Definition 3.

Then

$$\begin{aligned} D_\infty(\varepsilon) &= 1 && \text{if} \quad bH(a) > c(1 - H(a)) + \varepsilon\,; \\ &= 2 && \text{if} \quad bH(a) < c(1 - H(a)) - \varepsilon\,; \\ &= \text{either of above} && \text{otherwise.} \end{aligned}$$

$$\begin{aligned} \int V(\theta, D_\infty(\varepsilon))\, dF_n(\theta) &= c(1 - F_n(a)) && \text{if} \quad bH(a) > c(1 - H(a)) + \varepsilon\,; \\ &= bF_n(a) && \text{if} \quad bH(a) < c(1 - H(a)) - \varepsilon\,; \\ &= \text{either (depends on } D_n(\varepsilon)) && \text{otherwise.} \end{aligned}$$

Then

$$\begin{aligned} &\int V(\theta, D_\infty(\varepsilon))\, dF_n(\theta) - \inf_D \int V(\theta, D)\, dF_n(\theta) \\ &\quad = \max\{0, c(1 - F_n(a)) - bF_n(a)\}\,, && \text{if} \quad H(a) > \frac{c + \varepsilon}{b + c}\,; \\ &\quad = \max\{0, bF_n(a) - c(1 - F_n(a))\}\,, && \text{if} \quad H(a) < \frac{c - \varepsilon}{b + c}\,; \\ &\quad = \text{either of above (depends on } D_n(\varepsilon)) && \text{otherwise.} \end{aligned}$$

Suppose $H(a-) < c/(b + c) < H(a)$. Then $\exists\, \varepsilon > 0$ such that $H(a-) <$

$(c - \varepsilon)/(b + c) < H(a)$. Take F_n to be a sequence such that

$$F_n(a) \to \theta^{**} \quad \text{where} \quad H(a-) < \theta^{**} < \frac{c - \varepsilon}{b + c}, \qquad \text{and} \qquad F_n \to_\omega H .$$

Then $D_\infty(\varepsilon) = 1$ and

$$\int V(\theta, D_\infty(\varepsilon))\, dF_n(\theta) - \inf_D \int V(\theta, D)\, dF_n(\theta) = \max\{0, c(1 - F_n(a)) - bF_n(a)\} .$$

As $n \to \infty$, $c(1 - F_n(a)) - bF_n(a) \to c - (b + c)\theta^{**} > c - (b+c)((c-\varepsilon)/(b+c)) = \varepsilon > 0$. Hence

$$\lim_{\varepsilon \downarrow 0} \limsup_{n\to\infty} \left[\int V(\theta, D_\infty(\varepsilon))\, dF_n(\theta) - \inf_D \int V(\theta, D)\, dF_n(\theta)\right] = c - (b + c)\theta^{**} > 0 ,$$

so (V, H) is unstable in this case by Definition 3, and hence by Definition 4.

Similarly we can show if $H(a-) < c/(b + c) = H(a)$ then (V, H) is weakly stable and the stabilizing decision is 2 by Definition 4, and hence by Definition 3. In all other cases (V, H) is strongly stable by Definition 4, and hence by Definition 3. So we can see (V, H) is unstable iff $H(a-) < c/(b + c) < H(a)$, weakly stable iff $H(a-) < c/(b + c) = H(a)$, with 2 being the stabilizing decision, and strongly stable otherwise. In particular, if H is continuous at a then (V, H) is strongly stable. All the above holds for both definitions. This example is important because it shows that all three phenomena, strong stability, weak stability and instability, exist.

EXAMPLE 2. *Simple hypothesis, composite alternative.* An alternative two decision problem can be defined as follows: let

$$\begin{aligned} &V(\theta, 1) = 0 \quad \text{and} \quad V(\theta, 2) = b \quad \text{if} \quad \theta = a \\ &V(\theta, 1) = c \quad \text{and} \quad V(\theta, 2) = 0 \quad \text{if} \quad \theta \neq a , \end{aligned}$$

where b and c are positive.

Let $J_n(a) = F_n(a) - F_n(a-)$ and $K(a) = H(a) - H(a-)$. Then the calculation of B_n, formula (6), is exactly as Example 1 with J_n replacing F_n and K replacing G.

Thus

$$\begin{aligned} B_n &= \max\{0, c(1 - J_n(a)) - bJ_n(a)\} && \text{if} \quad K(a) > \frac{c + \varepsilon}{b + c}; \\ &= \max\{0, bJ_n(a) - c(1 - J_n(a))\} && \text{if} \quad K(a) < \frac{c - \varepsilon}{b + c}; \\ &= \text{either of the above} && \text{otherwise.} \end{aligned}$$

From this it is easy to see that (V, H) is

(i) strongly stable if $H(a) - H(a-) < c/(b + c)$;

(ii) weakly stable if $H(a) - H(a-) = c/(b + c)$ (the stabilizing decision is 2); and

(iii) unstable if $H(a) - H(a-) > c/(b + c)$.

This analysis again holds for both definitions.

EXAMPLE 3. *Squared error loss.* Consider $\mathscr{D} = \Theta = \mathbb{R}$, the real line, and the pair $((\theta - D)^2, H)$ for any opinion $H(\theta)$ with finite variance. Since we are looking for a counterexample, we use Definition 3. Thus let $G_n = H \ \forall\, n$, and let μ_∞ and σ_∞^2 be the mean and variance of $H(\theta)$, which we assume exists. Then

$$\int V(\theta, D)\, dH(\theta) = \sigma_\infty^2 + (\mu_\infty - D)^2 .$$

When $D = \mu_\infty$ we achieve the infimum σ_∞^2 and for every $\varepsilon > 0$, and every $D_n(\varepsilon)$,

$$\mu_\infty - \varepsilon^{\frac{1}{2}} \leqq D_n(\varepsilon) \leqq \mu_\infty + \varepsilon^{\frac{1}{2}} .$$

By finiteness of σ_∞^2, the infimum value is finite. Let $F_n(\theta)$ be a convex combination of $H(\theta)$ and $J_n(\theta)$ with weights $(1 - 1/n)$ and $1/n$, where $J_n(\theta)$ is the distribution function of the random variable sure to take the value $\theta = n$. Also let μ_n be the mean of F_n. Then $\mu_n = (1 - 1/n)\mu_\infty + 1$, and

$$\begin{aligned}
\limsup_{n\to\infty} & \left[\int V(\theta, D_n(\varepsilon))\, dF_n - \inf_D \int V(\theta, D)\, dF_n(\theta)\right] \\
&= \limsup_{n\to\infty} \left[\int \{(\theta - D_n(\varepsilon))^2 - (\theta - \mu_n)^2\}\, dF_n(\theta)\right] \\
&= (\mu_n - D_n(\varepsilon))^2 \\
&\geqq \left(1 - \frac{\mu_\infty}{n} - \varepsilon^{\frac{1}{2}}\right)^2 \\
&\to (1 - \varepsilon^{\frac{1}{2}})^2 .
\end{aligned}$$

Thus, for any opinion $H(\theta)$ with finite variance, the pair $((\theta - D)^2, H)$ is unstable by Definition 3, and hence by Definition 4.

3. Bounded continuous loss functions. The distinction between two concepts of uniform continuity of a function $f(x, y)$ of two variables is important in the sequel: f is called continuous in x uniformly in y iff

$$\forall\, \varepsilon > 0 , \quad \forall\, x , \quad \exists\, \delta > 0 \quad \text{such that} \quad \forall\, y ,$$
$$|x - x_0| < \delta \Rightarrow |f(x, y) - f(x_0, y)| < \varepsilon ;$$

f is called uniformly continuous in x uniformly in y iff

$$\forall\, \varepsilon > 0 , \quad \exists\, \delta > 0 \quad \text{such that} \quad \forall\, x , \quad \forall\, y ,$$
$$|x - x_0| < \delta \Rightarrow |f(x, y) - f(x_0, y)| < \varepsilon .$$

The following lemma shows that these concepts are related in the same way that continuity and uniform continuity are.

LEMMA 1. *Suppose $f(x, y)$ is continuous in x uniformly in y on a compact set $x \in S$. Then f is uniformly continuous in x uniformly in y.*

The proof is a simple extension of the proof that a continuous function on a compact set is uniformly continous, and is therefore left to the reader.

LEMMA 2. *Suppose that*

(i) $|V(\theta, D)| \leqq B$ *for all* θ *and* D;
(ii) $V(\theta, D)$ *is continuous in* θ *uniformly in* D;
(iii) $F_n \rightarrow_\omega H$;

then

$$\forall\, \varepsilon > 0\,, \quad \exists\, N \quad \text{such that} \quad \forall\, n \geqq N\,, \quad \text{and} \quad \forall\, D$$
$$|\textstyle\int V(\theta, D)\, d(H(\theta) - F_n(\theta))| < \varepsilon\,.$$

PROOF. Choose $\varepsilon > 0$. Choose a and b, points of continuity of $H(x)$, so that $H(a) \leqq \varepsilon$, $1 - H(b) \leqq \varepsilon$. In the closed interval $[a, b]$ the function $V(\theta, D)$ is uniformly continuous in θ uniformly in D, by Lemma 1 and Assumption (ii). Then there exist points of continuity of $H(\theta)$ in $[a, b]$ $a = a_0 < a_1 < \cdots < a_s = b$ such that

$$|V(\theta, D) - V(a_k, D)| < \varepsilon$$

for all D and for $a_k \leqq \theta \leqq a_{k+1} k = 0, \cdots, s-1$.
Let

$$\begin{aligned} V_\varepsilon(\theta, D) &= V(a_k, D) \qquad a_k \leqq \theta \leqq a_{k+1} \quad k = 0, \cdots, s-1 \\ &= 0 \qquad\qquad\quad \text{otherwise.} \end{aligned}$$

Then for any distribution function $G(\theta)$,

$$\textstyle\int V_\varepsilon(\theta, D)\, dG(\theta) = \sum_{k=0}^{s-1} V(a_k, D)[G(a_{k+1}) - G(a_k)]\,.$$

Since $F_n(\theta) \rightarrow H(\theta)$ as $n \rightarrow \infty$ at $\theta = a_k$

$$\textstyle\int V_\varepsilon(\theta, D)\, dF_n(\theta) \rightarrow \int V_\varepsilon(\theta, D)\, dH(\theta) \quad \forall\, D$$

and since s is finite, the above occurs uniformly in D. Thus

$$\forall\, \varepsilon > 0\,, \quad \exists\, N\,, \quad \forall\, n \geqq N\,, \quad \forall\, D\,, \qquad |\textstyle\int V_\varepsilon(\theta, D)(d(F_n(\theta) - H(\theta)))| < \varepsilon\,.$$

For any distribution function $G(\theta)$

$$\begin{aligned} &\textstyle\int |V(\theta, D) - V_\varepsilon(\theta, D)|\, dG(\theta) \\ &\quad = \textstyle\int_{-\infty}^{a} |V(\theta, D) - V_\varepsilon(\theta, D)|\, dG(\theta) \\ &\qquad + \textstyle\int_a^b |V(\theta, D) - V_\varepsilon(\theta, D)|\, dG(\theta) + \int_b^\infty |V(\theta, D) - V_\varepsilon(\theta, D)|\, dG(\theta) \\ &\quad \leqq BG(a) + \varepsilon[G(b) - G(a)] + B[1 - G(b)] \quad \forall\, D\,. \end{aligned}$$

Applying this to $H(\theta)$ yields

$$\textstyle\int |V(\theta, D) - V_\varepsilon(\theta, D)|\, dH(\theta) \leqq (2B + 1)\varepsilon\,.$$

Applying it to $F_n(\theta)$ and noting that $F_n(a) \rightarrow H(a)$, $F_n(b) \rightarrow H(b)$, yields that, for large enough n,

$$\textstyle\int |V(\theta, D) - V_\varepsilon(\theta, D)|\, dF_n(\theta) \leqq (2B + 2)\varepsilon\,.$$

Then $\exists N$ such that $\forall n \geqq N \ \forall D$

$$\begin{aligned}
|\textstyle\int V(\theta, D)\, dF_n(\theta) - \int V(\theta, D)\, dH(\theta)| \\
&\leqq |\textstyle\int [V(\theta, D) - V_\varepsilon(\theta, D)]\, dF_n| + |\int V_\varepsilon(\theta, D)[dF_n(\theta) - dH(\theta)]| \\
&\quad + |\textstyle\int (V(\theta, D) - V_\varepsilon(\theta, D))\, dH(\theta)| \\
&\leqq (2B + 2)\varepsilon + \varepsilon + (2B + 1)\varepsilon = (4B + 4)\varepsilon\,.
\end{aligned}$$

Since ε is arbitrary, Lemma 2 is proved. □

THEOREM 2. *Suppose* (i) $|V(\theta, D)| \leqq B$ *for all* θ *and* D *and* (ii) $V(\theta, D)$ *is continuous in* θ *uniformly in* D. *Then* (V, H) *is strongly stable by Definition* 4.

PROOF. By Lemma 2, $\forall \varepsilon > 0$, $\exists N_1$ such that $\forall n > N_1$, $\forall D$

$$|\textstyle\int V(\theta, D)\, d(H(\theta) - F_n(\theta))| < \varepsilon\,, \qquad \text{and}$$

$$\exists N_2 \quad \text{such that} \quad \forall n > N_2\,, \quad \forall D\,, \qquad |\textstyle\int V(\theta, D)\, d(H(\theta) - G_n(\theta))| < \varepsilon\,.$$

Then $\forall n > \max(N_1, N_2)$, $\forall D$

$$\begin{aligned}
\textstyle\int V(\theta, D)\, dF_n(\theta) - \int V(\theta, D_n(\varepsilon))\, dF_n(\theta) \\
&\geqq (\textstyle\int V(\theta, D)\, dH - \varepsilon) - (\int V(\theta, D_n(\varepsilon))\, dH(\theta) + \varepsilon) \\
&\geqq (\textstyle\int V(\theta, D)\, dG_n - \varepsilon) - (\int V(\theta, D_n(\varepsilon))\, dG_n + \varepsilon) - 2\varepsilon \\
&\geqq -5\varepsilon\,.
\end{aligned}$$

So

$$\lim_{\varepsilon \downarrow 0} \limsup_{n\to\infty} [\textstyle\int V(\theta, D_n(\varepsilon))\, dF_n(\theta) - \inf_D \int V(\theta, D)\, dF_n(\theta)] = 0\,. \qquad \square$$

EXAMPLE 4. Take the same example as Example 3, only restrict the domain, so that $\mathscr{D} = \Theta = C$ where C is some compact subset of $\mathbb{R}$. Then squared error satisfies the condition of Theorem 2, and is therefore strongly stable when paired with any opinion H by both Definitions 3 and 4.

4. Estimation or prediction loss functions with bounded growth. In this section, the following assumptions are frequently used:

(i) $V(\theta, D) = h(\theta - D)$, where h is continuous, nondecreasing in $(0, \infty)$, nonincreasing in $(-\infty, 0)$ and $h(0) = 0$.

(ii) h satisfies the following Lipschitz condition in the tail: $|h(x) - h(y)| \leqq B|x - y|$ for all $|y| > y_0$, and x, and for some constant $B > 0$.

Note that in this section B represents a bound on the growth of h. However h itself may be unbounded. The following example shows that Assumptions (i) and (ii) are not sufficient to ensure stability.

EXAMPLE 5. Let

$$\begin{aligned}
h(x) &= |x| \qquad \text{if} \quad -1 < x \\
&= 1 \qquad \text{otherwise,}
\end{aligned}$$

and let $H(\theta)$ be the distribution function of any random variable that has a finite mean. Again since we are looking for a counterexample, we use Definition 3.

Hence let $G_n \equiv H(\theta)$. Then $D_\infty(\varepsilon)$ is defined as any decision D satisfying

$$\int h(\theta - D_\infty(\varepsilon))\, dH(\theta) \leqq \inf_D \int h(\theta - D)\, dH(\theta) + \varepsilon .$$

Let $F_n(\theta)$ be a convex combination of $H(\theta)$ and $J_n(\theta)$ with weights $(1 - 1/n)$ and $1/n$ respectively, where $J_n(\theta)$ is the distribution function of the random variable sure to take the value $\theta = 3n$. Then

$$\begin{aligned}\int V(\theta, D_\infty(\varepsilon))\, dF_n(\theta) &- \inf_D \int V(\theta, D)\, dF_n(\theta) \\ &\geqq \int h(\theta - D_\infty(\varepsilon))\, dF_n(\theta) - \int h(\theta - 3n)\, dF_n(\theta) \\ &\geqq 2 - \int_{-\infty}^{3n} h(\theta - 3n)\, dH(\theta) - \int_{3n}^{\infty} h(\theta - 3n)\, dH(\theta) \\ &\geqq 2 - 1 - \int_{3n}^{\infty} \theta\, dH(\theta) .\end{aligned}$$

The existence of the mean of H implies that

$$\lim_{a \to \infty} \int_a^\infty \theta\, dH(\theta) = 0 ,$$

so

$$\lim_{\varepsilon \downarrow 0} \limsup_{n \to \infty} \left[\int V(\theta, D_\infty(\varepsilon))\, dF_n(\theta) - \inf_D \int V(\theta, D)\, dF_n(\theta)\right] \geqq 1 ,$$

so (V, H) is unstable by Definition 3, and therefore by Definition 4.

LEMMA 3. *The pair* (V, H) *is strongly stable by Definition* 4 *if, in addition to conditions* (i) *and* (ii), *the following condition* (iii) *obtains*:

(iii) *There is a compact interval* $[a, b]$ *and an* $\varepsilon_0 > 0$ *such that for every* ε, $0 < \varepsilon < \varepsilon_0$, *every sequence* $G_n \to_w H$, *and every sequence of* ε*-optimal decisions* $D_1, D_2, \cdots$ *for* (V, G_n), *there is an* N *such that for all* $n > N$, $D_n \in [a, b]$.

PROOF. Without loss of generality we may assume $b > y_0$, and $a < -y_0$. Since h is continuous in $[a, b]$, h is uniformly continuous in $[a, b]$. Thus given $\varepsilon > 0$, there exists a $\delta > 0$ such that for every $x, y \in [a, b]$, $|h(x) - h(y)| < \varepsilon$ if $|x - y| < \delta$. Choose $\delta < (b - a)/2$. Now there is a finite open covering of $[a, b]$ $\{(c_i, d_i) | i = 1, 2, \cdots, k\}$ such that $d_i - c_i < \min\{\delta, \varepsilon\}$ for all $i = 1, 2, \cdots, k$. Let $e_i \in (c_i, d_i)$. We now proceed to show that $|h(\theta - e_i) - h(\theta - e_j)|$ is bounded.

$$\begin{aligned}|h(\theta - e_i) - h(\theta - e_j)| &\leqq h(y_0) + h(-y_0) + B|e_i - e_j| \\ &\leqq h(y_0) + B(b - a + 2\varepsilon) + h(-y_0) \\ &\leqq h(y_0) + 2B(b - a) + h(-y_0) .\end{aligned}$$

Thus $|h(\theta - e_i) - h(\theta - e_j)|$ is bounded. By the Helly-Bray theorem there exist N_{ij} and M_{ij} such that $\forall\, n > N_{ij}$

$$\left|\int (V(e_i, \theta) - V(e_j, \theta))\, dF_n(\theta) - \int (V(e_i, \theta) - V(e_j, \theta))\, dH(\theta)\right| < \varepsilon ,$$

and $\forall\, n > M_{ij}$,

$$\left|\int (V(e_i, \theta) - V(e_j, \theta))\, dG_n(\theta) - \int (V(e_i, \theta) - V(e_j, \theta))\, dH(\theta)\right| < \varepsilon .$$

Let $N_0 = \max(N_{12}, N_{13}, \cdots, N_{k-1\,k}, M_{12}, M_{13}, \cdots, M_{k-1\,k})$.

Now suppose $t_1 \in (c_i, d_i)$ and $t_2 \in (c_i, d_i)$ for some i. Our purpose is to bound $|h(\theta - t_1) - h(\theta - t_2)|$. Without loss of generality, assume $t_1 > t_2$.

(a) If $\theta \geqq t_1 + b$, then

$$|h(\theta - t_1) - h(\theta - t_2)| \leqq B|t_1 - t_2| \leqq B\varepsilon \leqq (B + 1)\varepsilon ;$$

(b) If $t_1 + b > \theta \geqq t_2 + b$, then

$$\begin{aligned}|h(\theta - t_1) - h(\theta - t_2)| &\leqq |h(\theta - t_1) - h(b)| + |h(b) - h(\theta - t_2)| \\ &\leqq \varepsilon + B|\theta - t_2 - b| \leqq \varepsilon + B\varepsilon = (B + 1)\varepsilon .\end{aligned}$$

By similar arguments, it can be shown that when $t_2 + b > \theta \geqq a + t_1$, $a + t_1 > \theta \geqq a + t_2$ or $a + t_2 > \theta$ the inequality $|h(\theta - t_1) - h(\theta - t_2)| \leqq (B + 1)\varepsilon$ still holds. Let $D \in [a, b]$. Then there is an l such that $D \in (c_l, d_l)$. Let $n > N_0$. There is an m such that $D_n(\varepsilon) \in (c_m, d_m)$. Then

$$\begin{aligned}&\int (V(d, \theta) - V(D_n(\varepsilon), \theta))\, dF_n(\theta) \\ &\quad = \int [V(D, \theta) - V(e_l, \theta) + V(e_l, \theta) - V(D_n(\varepsilon), \theta) + V(e_m, \theta) \\ &\qquad - V(e_m, \theta)]\, dF_n(\theta) \\ &\quad \geqq -2(B + 1)\varepsilon + \int (V(e_l, \theta) - V(e_m, \theta))\, dF_n(\theta) \\ &\quad \geqq -2(B + 1)\varepsilon + \int (V(e_l, \theta) - V(e_m, \theta))\, dH(\theta) - \varepsilon \\ &\quad \geqq -2(B + 1)\varepsilon + \int (V(e_l, \theta) - V(e_m, \theta))\, dG_n(\theta) - 2\varepsilon \\ &\quad \geqq -2(B + 2)\varepsilon + \int (V(d, \theta) - V(D_n(\varepsilon), \theta))\, dG_n(\theta) - 2(B + 1)\varepsilon \\ &\quad \geqq -(4B + 7)\varepsilon .\end{aligned}$$

Then $\forall\, n \geqq N_0$

$$\inf_{D \in [a,b]} \int V(D, \theta)\, dF_n(\theta) - \int V(D_n(\varepsilon), \theta)\, dF_n(\theta) \geqq -(4B + 7)\varepsilon .$$

Now $F_n \to_\omega H$, so if $D_n{}^*(\varepsilon)$ is a sequence of ε-optimal decisions for (F_n, V) then $\exists\, N$ such that $\forall\, n \geqq N$, $D_n{}^*(\varepsilon) \in [a, b]$. Thus $\forall\, n > \max(N, N_0)$,

$$\inf_{D \in [a,b]} \int V(D, \theta)\, dF_n(\theta) = \inf_D \int V(D, \theta)\, dF_n(\theta) .$$

Hence $\forall\, n > \max(N, N_0)$,

$$\inf_D \int V(D, \theta)\, dF_n(\theta) - \int V(D_n(\varepsilon), \theta)\, dF_n(\theta) \geqq -(4B + 7)\varepsilon .$$

Now we conclude

$$\lim_{\varepsilon \downarrow 0} \limsup_{n\to\infty} [\int V(D_n(\varepsilon), \theta)\, dF_n(\theta) - \inf_D \int V(D, \theta)\, dF_n(\theta)] = 0 .$$

Thus (V, H) is stable by Definition 4. □

THEOREM 3. *The pair* (V, H) *is strongly stable by Definition* 4 *if, in addition to conditions* (i) *and* (ii), *the following condition* (iv) *obtains*:

(iv) *There exist* $r > 0$ *such that* $h(x) \geqq r|x|$, $\forall\, x$.

PROOF. Since H is a distribution function, we can find b large enough such that $b > y_0$, both b and $-b$ are continuity points of H, and $(H(b) - H(-b))/(1 - H(b)) > 2B/r$. The strategy below is to prove that any decision D either must lie in a certain interval or $D^* = 0$ beats D by at least ε. Hence all ε-optimal strategies lie in the interval, which permits use of Lemma 3. Let

$$D > \frac{h(-y_0) + h(y_0) + rb + \varepsilon_0}{H(b) - H(-b)} \cdot \frac{2}{r} .$$

It is straightforward to show $D/2 > b > y_0$.

$$\begin{aligned}
&\textstyle\int_{-\infty}^{\infty} (h(\theta - D^*) - h(\theta - D))\, dH(\theta) \\
&\quad = \left(\int_{-\infty}^{-b} + \int_{-b}^{-y_0} + \int_{-y_0}^{y_0} + \int_{y_0}^{b} + \int_{b}^{D-y_0} + \int_{D-y_0}^{D+y_0} + \int_{D+y_0}^{\infty}\right)(h(\theta) - h(\theta - D))\, dH(\theta) \\
&\quad \leqq [B(b - y_0) + h(-y_0) - r(y_0 + D)] \textstyle\int_{-b}^{-y_0} dH(\theta) \\
&\qquad + [h(y_0) + h(-y_0) - r(D - y_0)] \textstyle\int_{-y_0}^{y_0} dH(\theta) \\
&\qquad + [h(y_0) + Bb - r(D - b)] \textstyle\int_{y_0}^{b} dH(\theta) + [h(y_0) + BD - ry_0] \int_{b}^{D-y_0} dH(\theta) \\
&\qquad + [h(y_0) + BD] \textstyle\int_{D-y_0}^{D+y_0} dH(\theta) + BD \int_{D+y_0}^{\infty} dH(\theta) \\
&\quad \leqq h(-y_0) + h(y_0) + BD(1 - H(b)) - rD(H(b) - H(-b)) + rb \\
&\quad \leqq h(-y_0) + h(y_0) + rb - \frac{rD}{2}(H(b) - H(-b)) - rD(H(b) - H(-b)) \\
&\quad < -\varepsilon .
\end{aligned}$$

Similarly if

$$D < -\frac{h(-y_0) + h(y_0) + rb + \varepsilon_0}{H(b) - H(-b)} \cdot \frac{2}{r} ,$$

then

$$\textstyle\int_{-\infty}^{\infty} (h(\theta - D^*) - h(\theta - D))\, dH(\theta) < -\varepsilon .$$

So any ε-optimal decision D_ε for H must satisfy

$$|D_\varepsilon| < \frac{h(-y_0) + h(y_0) + rb + \varepsilon_0}{H(b) - H(-b)} \cdot \frac{2}{r} .$$

Let b_1 be a continuity point of H chosen so that $b_1 > y_0$ and

$$(H(b_1) - H(-b_1))/(1 - H(b_1)) > 1 + \frac{2B}{r} .$$

Let $J_n \to_\omega H$. Then $\exists\, N$ such that $\forall\, n \geqq N$,

$$J_n(b_1) - J_n(-b_1)/(1 - J_n(b_1)) > 2B/r$$

and

$$J_n(b_1) - J_n(-b_1) > \tfrac{1}{2}(H(b_1) - H(-b_1)) .$$

Let $m = 2(h(-y_0) + h(y_0) + rb_1 + \varepsilon_0)/r$. The ε-optimal decisions for (J_n, V) for all $n > N$ is within

$$(-m/(J_n(b_1) - J_n(-b_1)) , \qquad m/(J_n(b_1) - J_n(-b_1))) ,$$

and hence within

$$(-2m/(H(b_1) - H(-b_1)) , \qquad 2m/(H(b_1) - H(-b_1))) .$$

Thus condition (iii) obtains, and hence (V, H) is strongly stable by Definition 4 using Lemma 3. □

COROLLARY 1. *Let $I(\cdot)$ be the usual indicator function. Then $V(\theta, D) = a(\theta - D)I(\theta \geqq D) + b(D - \theta)I(\theta < D)$, where $a > 0$, $b > 0$, is strongly stable by Definition* 4 *with any H. When $a = b$, V in Corollary* 1 *specializes to absolute error.*

The following example shows that conditions (i) and (ii) and symmetry of h around zero ($h(x) = h(-x)$) are not sufficient to assure strong stability of (V, H).

EXAMPLE 6.

$$
\begin{aligned}
h(x) &= x && \text{if } 0 \leqq x \leqq 1 \\
&= 1 && \text{if } 1 < x \leqq 2 \cdot (2)^3 \\
&= (x - 2j^{j+1}) \cdot \frac{1}{j} + (j-1)^{j-1} && \text{if } 2j^{j+1} < x \leqq 3j^{j+1} - j(j-1)^{j-1} \\
&= j^j && \text{if } 3j^{j+1} - j(j-1)^{j-1} < x \leqq 2(j+1)^{j+2} \\
& && \text{for } j = 2, 3, \cdots,
\end{aligned}
$$

and let $h(-x) = h(x)$.

Then h is continuous, symmetric, piece-wise linear, nondecreasing in $(0, \infty)$, nonincreasing in $(-\infty, 0)$, and satisfies $h(0) = 0$ and the Lipschitz condition. Now let H be the distribution function of the random variable sure to take the value $\theta = 0$, and since we are looking for a counterexample, we take Definition 3 and let $G_n \equiv H$. Let $F_n(\theta)$ be a convex combination of $H(\theta)$ and $J_n(\theta)$ with weights $(1 - (1/n))$ and $1/n$, where $J_n(\theta)$ is the distribution function of the random variable sure to take the value $3(n^{n+1}) - n(n-1)^{n-1}$. Then $F_n \to_\omega H$, and $D_n(\varepsilon) \in (-\varepsilon, \varepsilon)$ where $\varepsilon < 1$. By comparing the expected losses of the decisions, $D_n(\varepsilon)$ and $2(n)^{n+1}$, it can be shown that (V, H) is unstable by Definition 3, and hence by Definition 4.

THEOREM 4. *(V, H) is strongly stable by Definition* 4 *if, in addition to Assumptions* (i) *and* (ii), *the following condition* (v) *is satisfied*:

(v) *$h(x) = h(-x)$, h is unbounded, and $h(x + y) \leqq h(x) + h(y)$, for $x, y > 0$.*

PROOF. Our strategy is to apply Lemma 3 by proving condition (iii). Choose $\varepsilon_0 > 0$, and ε such that $0 < \varepsilon < \varepsilon_0$.

Since H is a distribution, there exists a positive number b such that $H(-b) \leqq \frac{1}{4}$, $H(b) \geqq \frac{3}{4}$, $b > y_0$ and b and $-b$ are continuity points of H. Since $h(x)$ is unbounded, there is a $D_0 > 0$ such that

$$h(D_0) > 2h(b) + Bb + 4\varepsilon_0 .$$

Now we will show that $D^* = 0$ is better, by at least ε, than any $D > b + D_0$ or any $D < -b - D_0$. Suppose first that $D > b + D_0$. Then

$$
\begin{aligned}
I &= \int V(\theta, D^*)\, dH(\theta) - \int V(\theta, D)\, dH(\theta) \\
&= \int_{-\infty}^{-b} (h(\theta - D^*) - h(\theta - D))\, dH(\theta) + \int_{-b}^{b} (h(\theta - D^*) - h(\theta - D))\, dH(\theta) \\
&\quad + \int_{b}^{\infty} (h(\theta - D^*) - h(\theta - D))\, dH(\theta) = I_1 + I_2 + I_3 .
\end{aligned}
$$

It follows:

(1) $I_1 \leqq 0$; and

(2) in I_2, $h(\theta - D^*) \leqq h(b)$ and $\theta - D < 0$.

Also applying condition (v) to I_3, we have

$$\begin{aligned} I &\leqq \textstyle\int_{-b}^{b} (h(b) - h(b - D))\, dH(\theta) + \int_b^\infty h(D - D^*)\, dH(\theta) \\ &\leqq \textstyle\int \tfrac{1}{2}h(b) - \tfrac{1}{4}h(D - b) + \tfrac{1}{4}Bb \\ &\leqq -\varepsilon_0 . \end{aligned}$$

Thus the ε-optimal decision for H cannot be greater than $b + D_0$. Similarly it cannot be smaller than $-b - D_0$. Consider now the sequence $G_n \to_\omega H$. There is a point b_1 such that both b_1 and $-b_1$ are continuity points of H satisfying $b_1 > b$, $H(b_1) \geqq \frac{7}{8}$, and $H(-b_1) \leqq \frac{1}{8}$. Let D_1 satisfy

$$h(D_1) > 2h(b_1) + Bb_1 + 4\varepsilon_0 .$$

Since $G_n \to_\omega H$, there is an N such that $\forall\, n > N$, $G_n(-b_1) \leqq \frac{1}{4}$ and $G_n(b_1) \geqq \frac{3}{4}$. Then for all such n, $D_n(\varepsilon) \in (-b_1 - D_1, b_1 + D_1)$. Lemma 3 now applies, so (V, H) is stable by Definition 4. □

COROLLARY 2. *If* $V(\theta, D) = |\theta - D|^p 0 < p \leqq 1$ *then* (V, H) *is strongly stable by Definition* 4.

The next example shows the effect of asymmetry.

EXAMPLE 7. Let $V(\theta, D) = h(\theta - D)$, where

$$\begin{aligned} h(x) &= x^{\frac{1}{2}} \quad & x \geqq 0 \\ &= |x|^{\frac{1}{3}} \quad & x < 0 . \end{aligned}$$

Then let $H(\theta)$ and $G_n(\theta)$, $F_n(\theta)$ be the same as in Example 6 except now $J_n(\theta)$ is the distribution function of the random variable sure to take the value $16n^4$. It can be shown that (V, H) is unstable by Definition 3 in this case.

5. Conclusion. An alternative method of presentation of our results would have been to stress that we are studying a particular kind of continuity, and that the subscript n has no real-world counterpart. We find the statement of the theory in terms of sequences to be easier to understand and, we hope, accessible to a wider audience. There are, of course, alternative topologies that might be imposed on this problem and whose consequences would be interesting to explore.

Our aim has been to study stability as an approach to a personalistic Bayesian theory of robustness. We intend for our results to be used not artificially to alter loss functions very far from the origin to achieve a theoretical advantage of no practical consequence, but rather as a way of learning more about the underlying structure of Bayesian decision theory, in much the same way that large-sample theory can be used in sampling theory—we often do not know whether the large sample theory is relevant, but it is a good guide to intuition.

Added in Proof. The dissertation of David T. Chuang (1978) shows that the conjectured equivalence of Definitions 1 and 2 is false in general but true for the estimation/prediction case. He also finds in that case that if a loss function is strongly stable with one opinion it is strongly stable with all opinions, and characterizes stable loss functions. Necessary and sufficient conditions are also given for stability in the more general case in which the likelihood function is considered known and fixed.

REFERENCES

[1] ANDREWS, D. F. et al. (1972). *Robust Estimates of Location.* Princeton Univ. Press.

[2] ANTLEMAN, G. R. (1965). Insensitivity to non-optimal design in Bayesian decision theory. *J. Amer. Statist. Assoc.* **60** 584–601.

[3] BICKEL, P. J. and LEHMANN, E. L. (1975a, b). Descriptive statistics for non-parametric models, I and II. *Ann. Statist.* **3** 1038–1044, 1045–1069.

[4] BRITNEY, R. and WINKLER, R. (1974). Bayesian point estimation and prediction. *Ann. Inst. Statist. Math.* **26** 15–34.

[5] DEMPSTER, A. (1975). *A Subjectivist Look at Robustness.* Research Report S-33, Depart. of Statist., Harvard Univ.

[6] EDWARDS, W., LINDEMAM H. and SAVAGE, L. J. (1963). Bayesian statistical inference for psychological research. *Psychological Review* **70** 193–242.

[7] FISHBURN, P., MURPHY, A. and ISAACS, H. (1967). Sensitivity of decision to probability estimation errors: a reexamination. *Operations Research* **15** 254–267.

[8] HUBER, P. J. (1972). Robust statistics, a review. *Ann. Math. Statist.* **43** 1041–1067.

[9] HUBER, P. J. (1973). Robust regression: asymptotics, conjectures and Monte Carlo. *Ann. Statist.* **1** 799–821.

[10] PIERCE, D. and FOLKS, J. L. (1969). Sensitivity of Bayes procedures to the prior distribution. *Operations Research* **17** 344–350.

[11] SCHLAIFER, R. (1959). *Probability and Statistics for Business Decisions.* McGraw-HIll, New York.

[12] STONE, M. (1963). Robustness of non-ideal decision procedures. *J. Amer. Statist. Assoc.* **58** 480–486.

DEPARTMENT OF STATISTICS
CARNEGIE-MELLON UNIVERSITY
SCHENLEY PARK
PITTSBURGH, PENNSYLVANIA 15213