AD-A046 701 CARNEGIE-MELLON UNIV PITTSBURGH PA DEPT OF STATISTICS F/G 12/2
STABLE DECISION PROBLEMS.(U)
JUN 76 J B KADANE, D T CHUANG N00014-75-C-0516
UNCLASSIFIED TR-107 NL

1 OF 1
ADA046701

END
DATE
FILMED
12-77
DDC

(12)

DEPARTMENT

OF

STATISTICS

Carnegie-Mellon University

PITTSBURGH, PENNSYLVANIA 15213

12

6

STABLE DECISION PROBLEMS.

by

10

Joseph B. Kadane and
David T. Chuang

9

Technical Report No. 107
ONR Report No. 5

11

June, 1976

14 TR-107, 5-ONR

12 42p.

Department of Statistics
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

This research was supported by the Office of Naval Research under Contract No. N00014-75-C-0516 and Task No. NR 042-309.

15

Abstract

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 $F_n \xrightarrow{w} H$, $G_n \xrightarrow{w} H$ and $L_n \to V$, $W_n \to V$ uniformly,

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

for every sequence $D_n(\varepsilon)$ satisfying

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

We show that squared error loss is unstable with any pair 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, non-decreasing in $(0,\infty)$ and non-increasing 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.

Key words: Decision Theory, Robustness, Stable Estimation, Stable Decisions

AMS Classification: Primary 62C10, Secondary 62G35

Stable Decision Problems

by

Joseph B. Kadane and

David T. Chuang

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 procedures with respect to the use of wrong prior distributions. He emphasized the possible usefulness of non-ideal 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. This paper extends that analysis by allowing a loss function, and by allowing a sequence of opinions (prior or posterior), without inquiry into whether the source of uncertainty might be the prior or the likelihood function. We see no reason why the likelihood function is known more surely than the prior distribution.

To give an initial formalization of our question, suppose that the parameter space is $\Theta \subset R^k$ for some k, and the decision space is $\mathcal{D} \subset R^\ell$ 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

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

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

$$(2) \qquad \int L_\infty(\theta, D_\infty(\epsilon))\, dF_\infty(\theta) \leq W_\infty + \epsilon$$

Suppose, however, that my "true" opinion over Θ is on a sequence $F_n(\theta)$ which converges to $F_\infty(\theta)$ in distribution. Then the probability of any subset of Θ is nearly the probability given by $F_\infty(\theta)$. Also suppose that my "true" loss function over Θ is $L_n(\theta, D)$ which converges uniformly in θ and D to $L_\infty(\theta, D)$. Then there is a sequence of "true" losses generated

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

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

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

The worth of knowing the truth is then

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

which is always non-negative. Note that B_n is a function of ϵ, $D_\infty(\epsilon)$, n, L_n and F_n. Suppose that

$$(3) \qquad \lim_{\epsilon \downarrow 0} \limsup_{n\to\infty} B_n = 0$$

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

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

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

There are situations in which (3) holds for every choice of $L_n \to L_\infty$ uniformly and $F_n \xrightarrow{w} F_\infty$, but only for some particular choice $D_\infty(\epsilon)$. In this case, $D_\infty(\epsilon)$ 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.

From a more general point of view we can formulate our problems as follows: Take (L_n, F_n) as a sequence of truths, and (W_n, G_n) as a sequence of approximations where

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

$$F_n \overset{\omega}{\to} H, \quad G_n \overset{\omega}{\to} H \; .$$

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

Let $D_n(\epsilon)$ be defined by

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

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

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

then (V,H) is strongly stable (by Definition 2).

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

In this paper, this second set of definitions is used throughout. We conjecture that the definitions are equivalent. However the second definition 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 most general.

Section 2 introduces Definition 3, which is apparently simpler than Definition 2, and shows its equivalence to Definition 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 restraining its growth, and finds some of them strongly stable, and some unstable. To simplify matters, assume the one-dimensional case $(k = \ell = 1)$.

2. A General Structure Theorem and Some Examples

In the first part of this section we introduce yet another definition of strong (weak) stability, Definition 3 and show that it is equivalent to Definition 2. The greater simplicity of Definition 3 helps to simplify the rest of the paper.

Suppose $F_n \overset{w}{\to} H$ and $G_n \overset{w}{\to} H$. Let $D_n(\epsilon)$ be defined by

$$(6) \qquad \int V(\theta, D_n(\epsilon))\,dG_n(\theta) \le \inf_D \int V(\theta, D)\,dG_n(\theta) + \epsilon.$$

If, for every such choice of F_n, G_n, and $D_n(\epsilon)$,

$$(7) \qquad \lim_{\epsilon \downarrow 0} \lim_{n \to \infty} \sup \left[\int V(\theta, D_n(\epsilon))\,dF_n(\theta) - \inf_D \int V(\theta, D)\,dF_n(\theta)\right] = 0$$

then (V,H) is strongly stable (by Definition 3). If there is some choice of $D_n(\epsilon)$ which makes (6) hold, then (V,H) is weakly stable and $D_n(\epsilon)$ is the stabilizing decision (by Definition 3).

Now we can state

<u>Theorem 1</u> (V,H) is strongly (weakly) stable by Definition 2 if and only if (V,H) is strongly (weakly) stable by Definition 3.

<u>Proof</u>: 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 3 then follows trivially.

Suppose, then, that (V,H) is strongly (weakly) stable by Definition 3, and suppose that L_n and W_n are arbitrary sequences of loss functions converging uniformly in θ and D to V. Choose $\epsilon > 0$, and let $D_n(\epsilon)$ be defined by equation (4).

Choose $N_1 \ni \forall n \geq N_1$,

$|W_n(\theta,D) - V(\theta,D)| < \epsilon$ for every θ and D, using the uniform convergence of W_n to V. Then

$$\inf_D \int W_n(\theta,D)\,dG_n(\theta) - \inf_D \int V(\theta,D)\,dG_n(\theta)$$

$$= \inf_D \int W_n(\theta,D)\,dG_n(\theta) - \inf_D \int (V(\theta,D) - W_n(\theta,D) + W_n(\theta,D))\,dG_n(\theta)$$

$$\leq - \inf_D \int (V(\theta,D) - W_n(\theta,D))\,dG_n(\theta)$$

$$\leq \sup_D \int (W_n(\theta,D) - V(\theta,D))\,dG_n(\theta)$$

$$< \epsilon$$

Also $\inf_D \int W_n(\theta,D)\,dG_n(\theta) - \inf_D \int V(\theta,D)\,dG_n(\theta)$

$$= \inf_D \int (W_n(\theta,D) - V(\theta,D) + V(\theta,D))\,dG_n(\theta) - \inf_D \int V(\theta,D)\,dG_n(\theta)$$

$$\geq \inf_D \int (W_n(\theta,D) - V(\theta,D))\,dG_n(\theta) > -\epsilon.$$

Then $\left| \inf_D \int W_n(\theta,D)\,dG_n(\theta) - \inf_D \int V(\theta,D)\,dG_n(\theta) \right| < \epsilon.$

Also $\left| \int W_n(\theta,D_n(\epsilon))\,dG_n(\theta) - \int V(\theta,D_n(\epsilon))\,dG_n(\theta) \right| < \epsilon$

$$\int V(\theta,D_n(\epsilon))\,dG_n(\theta) - \inf_D \int V(\theta,D)\,dG_n(\theta) \leq$$

$$| \int V(\theta,D_n(\epsilon))dG_n(\theta) - \int W_n(\theta,D_n(\epsilon))dG_n(\theta)|$$

$$+ | \int W_n(\theta,D_n(\epsilon))dG_n(\theta) - \inf_D \int W_n(\theta,D)dG_n(\theta)|$$

$$+ | \inf_D \int W_n(\theta,D)dG_n(\theta) - \inf_D \int V(\theta,D)dG_n(\theta)| \leq 3\epsilon \ .$$

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

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

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

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

Also $| \int L_n(\theta,D_n(\epsilon))dF_n(\theta) - \int V(\theta,D_n(\epsilon))dF_n(\theta)| < \delta.$

Hence $\forall\delta > 0 \ \exists N_2 \ni \forall n \geq N_2$,

$$|[\int L_n(\theta,D_n(\epsilon))dF_n(\theta) - \inf_D \int L_n(\theta,D)dF_n(\theta)] -$$

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

Thus

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

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

Finally, taking $D_n(\epsilon)$ defined by (4), $\epsilon' = 3\epsilon$ and $D_n'(\epsilon')$ defined by (6),

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

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

Thus if (V,H) is strongly (weakly) stable by Definition 3, it is strongly (weakly) stable by Definition 2.

Q.E.D.

When stability is referred to in the rest of this paper, Definition 2 (or equivalently, Definition 3) is to be understood, unless otherwise specified.

Example 1. Composite hypothesis, composite alternative.

Suppose that there are only two available decisions $\{1,2\}$, and suppose that V is defined as follows:

$V(\theta,1) = 0$ and $V(\theta,2) = b$ if $\theta \le a$;

$V(\theta,1) = c$ and $V(\theta,2) = 0$ if $\theta > a$,

where b and c are assumed to be positive.

Then

$$D_n(\epsilon) = \begin{cases} 1 & \text{if } bG_n(a) > c(1-G_n(a)) + \epsilon \\ 2 & \text{if } bG_n(a) < c(1-G_n(a)) - \epsilon \\ \text{either} & \text{if } c(1-G_n(a)) - \epsilon \le bG_n(a) \le c(1-G_n(a)) + \epsilon \end{cases}$$

$$\int V(\theta, D_n(\epsilon))\,dF_n(\theta) = \begin{cases} c(1 - F_n(a)) & \text{if } bG_n(a) > c(1 - G_n(a)) + \epsilon \\ bF_n(a) & \text{if } bG_n(a) < c(1 - G_n(a)) - \epsilon \\ \text{either(depends on } D_n(\epsilon)) & \text{if } c(1 - G_n(a)) - \epsilon \leq \\ & \quad bG_n(a) \leq c(1 - G_n(a)) + \epsilon \end{cases}$$

$$\int V(\theta, D_n(\epsilon))\,dF_n(\theta) - \inf_D \int V(\theta, D)\,dF_n(\theta)$$

$$= \begin{cases} \max\{0,\ c(1 - F_n(a)) - bF_n(a)\}, & \text{if } G_n(a) > \dfrac{c + \epsilon}{b + c} \\ \max\{0,\ bF_n(a) - c(1 - F_n(a))\}, & \text{if } G_n(a) < \dfrac{c - \epsilon}{b + c} \\ \text{either of above (depend on } D_n(\epsilon)) & \text{if } \dfrac{c - \epsilon}{b + c} \leq G_n(a) \leq \dfrac{c + \epsilon}{b + c} \end{cases}$$

Suppose $H(a-) < \dfrac{c}{b + c} < H(a)$.

Then $\exists\ \epsilon > 0 \ni H(a-) < \dfrac{c - \epsilon}{b + c} < \dfrac{c + \epsilon}{b + c} < H(a)$.

Take G_n to be a sequence such that

$$G_n(a) \to \theta^* \quad \text{where} \quad \frac{c + \epsilon}{b + c} < \theta^* < H(a).$$

Take F_n to be a sequence such that

$$F_n(a) \to \theta^{**} \quad \text{where} \quad H(a-) < \theta^{**} < \frac{c - \epsilon}{b + c}.$$

Then for large enough n, $D_n(\epsilon) = 1$ and

$$\int V(\theta, D_n(\epsilon))\,dF_n(\theta) - \inf_D \int V(\theta, D)\,dF_n(\theta)$$

$$= \max\{0, c(1 - F_n(a)) - bF_n(a)\}.$$

As $n \to \infty$, $c(1 - F_n(a)) - bF_n(a)$

$$\to c - (b+c)\theta^{**} > c - (b+c)\ \frac{c-\epsilon}{b+c} = \epsilon > 0.$$

Hence

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

$$= c - (b+c)\theta^{**} > 0,$$

so (V,H) is unstable in this case.

Similarly we can show if $H(a-) < \frac{c}{b+c} = H(a)$ then (V,H) is weakly stable and the stabilizing decision is 2 . So we can see (V,H) is unstable iff $H(a-) < \frac{c}{b+c} < H(a)$, weakly stable iff $H(a-) < \frac{c}{b+c} = H(a)$, with 2 being the stabilizing decision, and strongly stable otherwise. In particular, if H is continuous at a then (V,H) is strongly stable.

Example 2. Simple hypothesis, composite alternative.

An alternative two decision problem can be defined as follows:

Let

$$V(\theta, 1) = 0 \text{ and } V(\theta, 2) = b \quad \text{if} \quad \theta = a$$

$$V(\theta, 1) = c \text{ and } V(\theta, 2) = 0 \quad \text{if} \quad \theta \neq a\ .$$

$$b,\ c,\ > 0$$

Let $J_n(a) = F_n(a) - F_n(a-)$ and

$$K_n(a) = G_n(a) - G_n(a-)\ .$$

Then the calculation of B_n, formula (4), is exactly as example 1 with J replacing F and K replacing G.

$$\text{Thus} \quad B_n = \begin{cases} \max\{0, c(1-J_n(a)) - bJ_n(a)\} & \text{if } K_n(a) > \dfrac{c+\epsilon}{b+c} \\ \max\{0, bJ_n(a) - c(1-J_n(a))\} & \text{if } K_n(a) < \dfrac{c-\epsilon}{b+c} \\ \text{either the above if} & \dfrac{c-\epsilon}{b+c} \le K_n(a) \le \dfrac{c+\epsilon}{b+c} \end{cases}$$

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

(i) strongly stable if $H(a) - H(a-) < \dfrac{c}{b+c}$

(ii) weakly stable if $H(a) - H(a-) = \dfrac{c}{b+c}$

(the stabilizing decision is 2)

(iii) unstable if $H(a) - H(a-) > \dfrac{c}{b+c}$.

ample 3. Squared Error Loss.

nsider $\mathcal{D} = \Theta = R$, the real line, and the pair

$((\theta - D)^2, H)$ for any opinion $H(\theta)$ with finite variance.

Let $G_n = H \;\forall n$, and let u_∞ 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 $D = u_\infty$ we achieve the infimum σ_∞^2 and for every $\epsilon > 0$, and every $D_n(\epsilon)$,

$$u_\infty - \sqrt{\epsilon} \le D_n(\epsilon) \le u_\infty + \sqrt{\epsilon} \; .$$

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 $(1-\frac{1}{n})$ and $\frac{1}{n}$, where $J_n(\theta)$ is the distribution function of the random variable sure to take the value $\theta = n$. Also let μ_n the mean of F_n. Then $u_n = (1-\frac{1}{n})\, u_\infty + 1$, and

$$\limsup_{n\to\infty} \left[\int V(\theta, D_n(\epsilon))\,dF_n - \inf_D \int V(\theta, D)\,dF_n(\theta)\right]$$

$$= \limsup_{n\to\infty} \left[\int \{(\theta - D_n(\epsilon))^2 - (\theta - \mu_n)^2\}\,dF_n(\theta)\right]$$

$$= (u_n - D_n(\epsilon))^2$$

$$\geq \left(1 - \frac{\mu_\infty}{n} - \sqrt{\epsilon}\right)^2$$

$$\to (1 - \sqrt{\epsilon})^2.$$

Thus, for any opinion $H(\theta)$ with finite variance, the pair $((\theta - D)^2, H)$ is unstable.

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

$$\forall \epsilon > 0, \quad \forall x, \ \exists \delta > 0 \ni \ \forall y, \ |x - x_0| < \delta \Rightarrow |f(x,y) - f(x_0,y)| < \epsilon ;$$

f is called uniformly continuous in x uniformly in y iff

$$\forall \epsilon > 0 \ \exists \delta > 0 \ni \forall x \ \forall y \ |x - x_0| < \delta \Rightarrow |f(x,y) - f(x_0,y)| < \epsilon .$$

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

<u>Lemma 1</u>: 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.

Proof: Suppose the contrary. Then $\exists \epsilon > 0 \ni \forall \delta > 0 \ \exists x,y \ni |x - x_0| < \delta$ and $|f(x,y) - f(x_0,y)| > \epsilon$ choose $\delta_1, \delta_2, \ldots \ni \delta_n > 0$ and $\delta_n \to 0$. Then there exist $\alpha > 0$ and sequences u_n and v_n in S such that,

$$|u_n - v_n| < \delta \quad \text{and} \quad |f(u_n,y_n) - f(v_n,y_n)| > \alpha$$

Compactness of S implies that u_n have a limit point $\xi \in S$ and v_n must have the same limit point. Take $\delta > 0$ be arbitrary small. Then infinitely many pairs u_n, v_n lie within δ of ξ. But this contradicts the continuity of f at ξ uniformly in y.

Q.E.D.

Lemma 2: Suppose that

(i) $|V(\theta,D)| \leq B$ for all θ and D

(ii) $V(\theta,D)$ is continuous in θ uniformly in D

(iii) $F_n \xrightarrow{w} H$

then

$$\forall \epsilon > 0 \ \exists N \ni \forall n \geq N \quad \forall D \ \left| \int V(\theta,D)d(H(\theta) - F_n(\theta)) \right| < \epsilon .$$

Proo[illegible] $\epsilon > 0$. Choose a and b, points of continuity of $H(x)$, [illegible] $H(a) \leq \epsilon$, $1 - H(b) \leq \epsilon$. In the closed interval $[a,b]$ the function $V(\theta,D)$ is uniformly continuous in θ uniformly in D, by Lemma and Assumption (ii). Then there exist points of continuity of $H(\theta)$ in $[a,b]$

$$a = a_0 < a_1 < \ldots < a_s = b \quad \text{such that}$$

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

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

$$\text{Let} \quad V_\epsilon(\theta,D) = \begin{cases} V(a_k,D) & a_k \leq \theta \leq a_{k+1} \ k=0,\ldots,s-1 \\ 0 & \text{otherwise} . \end{cases}$$

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

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

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

$$\int V_\epsilon(\theta,D)dF_n(\theta) \to \int V_\epsilon(\theta,D)d\,H(\theta) \ \forall D$$

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

Thus

$$\forall \epsilon > 0 \ \exists N \ \forall n \geq N \ \forall D \ \left| \int V_\epsilon(\theta,D)(d(F_n(\theta) - H(\theta))) \right| < \epsilon.$$

For any distribution function $G(\theta)$

$$\int |V(\theta,D) - V_\epsilon(\theta,D)| dG(\theta) = \int_{-\infty}^{a} |V(\theta,D) - V_\epsilon(\theta,D)| dG(\theta)$$

$$+ \int_a^b |V(\theta,D) - V_\epsilon(\theta,D)| dG(\theta) + \int_b^\infty |V(\dot{\theta},D) - V_\epsilon(\theta,D)| dG(\theta)$$

$$\leq BG(a) + \epsilon \ [G(b) - G(a)] + B[1 - G(b)] \ \forall D.$$

Applying this to $H(\theta)$ yields

$$\int |V(\theta,D) - V_\epsilon(\theta,D)| d\,H(\theta) \leq (2B+1)\epsilon.$$

Applying it to $F_n(\theta)$ and noting that

$F_n(a) \to H(a)$, $F_n(b) \to H(b)$, yields that, for large enough n,

$$\int |V(\theta,D) - V_\epsilon(\theta,D)| \ dF_n(\theta) \leq (2B+2) \ \epsilon.$$

Then $\exists N \ni \forall n \geq N \ \ \forall D$

$$\left| \int V(\theta,D) dF_n(\theta) - \int V(\theta,D) d\,H(\theta) \right|$$

$$\leq \left| \int [V(\theta,D) - V_\epsilon(\theta,D)] dF_n \right| + \left| \int V_\epsilon(\theta,D)[dF_n(\theta) - d\,H(\theta)] \right|$$

$$+ \left| \int (V(\theta,D) - V_\epsilon(\theta,D)) d\,H(\theta) \right|$$

$$\leq \ (2B+2) \ \epsilon + \epsilon + (2B+1) \ \epsilon = (4B+4) \ \epsilon.$$

Since ϵ is arbitrary, Lemma 2 is proved.

<u>Theorem 2</u>: Suppose (i) $|V(\theta,D)| \leq B$ for all θ and D. (ii) $V(\theta,D)$ is continuous in θ uniformly in D. Then (V,H) is strongly stable.

Proof: By Lemma 2, $\forall \epsilon > 0 \ \exists \ N_1 \ni \forall n > N_1, \ \forall D$

$$\left| \int V(\theta,D)d(H(\theta) - F_n(\theta)) \right| < \epsilon, \text{ and}$$

$$\exists N_2 \ni \forall n > N_2, \quad \forall D$$

$$\left| \int V(\theta,D)d(H(\theta) - G_n(\theta)) \right| < \epsilon.$$

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

$$\int V(\theta,D)dF_n(\theta) - \int V(\theta,D_n(\epsilon))dF_n(\theta)$$

$$\geq \left(\int V(\theta,D)dH - \epsilon\right) - \left(\int V(\theta,D_n(\epsilon))dH(\theta) + \epsilon\right)$$

$$\geq \int V(\theta,D)dH - \int V(\theta,D_n(\epsilon))dH - 2\epsilon$$

$$\geq \left(\int V(\theta,D)dG_n - \epsilon\right) - \left(\int L(\theta,D_n(\epsilon))dG_n + \epsilon\right) - 2\epsilon$$

$$\geq \int V(\theta,D)dG_n - \int V(\theta,D_n(\epsilon))dG_n - 4\epsilon$$

$$\geq -5\epsilon.$$

So

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

Q.E.D.

<u>Example 4</u>: Take the same example as example 3, only restrict the domain, so that $\mathcal{D} = \Theta = C$ where C is some compact subset of R. Then squared error satisfies the condition of Theorem 2, and is therefore strongly stable when paired with any opinion H.

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, non-decreasing in $(0,\infty)$, non-increasing in $(-\infty,0)$ and $h(0)=0$.

(ii) h satisfies the following Lipschitz condition in the tail: $|h(x)-h(y)| \le B|x-y|$ for all $|y| > y_0$, all 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

$$h(x) = \begin{cases} |x| & \text{if } -1 < x \\ 1 & \text{otherwise} \end{cases}$$

and let $H(\theta)$ be the distribution function of any random variable that has a finite mean. Let $G_n(\theta) = H(\theta)$. Then $D_n(\epsilon)$ is defined as any decision D satisfying

$$\int h(\theta - D_n(\epsilon))dH(\theta) \le \inf_D \int h(\theta-D)dH(\theta) + \epsilon.$$

First we show that $D_n(\epsilon)$ is bounded below for sufficiently small $\epsilon > 0$. Let d^* be a median of H. We show that $D_n(\epsilon) < d^* - 2$ leads to a contradiction for $\epsilon < \frac{1}{2}$ as follows: Let $d < d^* - 2$

$$\int_{-\infty}^{\infty} (h(\theta-d) - h(\theta-d^*))dH(\theta) = \left[\int_{-\infty}^{d-1} + \int_{d-1}^{d} + \int_{d}^{d^*-1} + \int_{d-1}^{d^*} + \int_{d^*}^{\infty}\right] (h(\theta-d) - h(\theta-d^*))dH(\theta)$$

$$= \int_{-\infty}^{d-1} (1 - 1)\,dH(\theta) + \int_{d-1}^{d} ((d-\theta)-1)\,dH(\theta) + \int_{d}^{d^*-1} ((\theta-d)-1)\,dH(\theta)$$

$$+ \int_{d^*-1}^{d^*} ((\theta-d) - (d^*-\theta))\,dH(\theta) + \int_{d^*}^{\infty} ((\theta-d) - (\theta-d^*))\,dH(\theta)$$

$$\geq \int_{d-1}^{d} (d-d-1)\,dH(\theta) + \int_{d}^{d^*-1} ((d-d)-1)\,dH(\theta)$$

$$+ \int_{d^*-1}^{d^*} (d^*-1-d) - (d^*-d^*+1)\,dH(\theta) + \tfrac{1}{2}(d^*-d)$$

$$\geq - \int_{d-1}^{d^*} dH + \tfrac{1}{2}(d^*-d)$$

$$\geq - \tfrac{1}{2} + \tfrac{1}{2}(d^*-d) > \tfrac{1}{2}.$$

Hence we must have $D_n(\epsilon) > d^*-2$ if $\epsilon < \tfrac{1}{2}$.

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 $\theta = 2n + D_n(\epsilon)$. Then

$$\int V(\theta, D_n(\epsilon))dF_n(\theta) - \inf_D \int V(\theta, D)dF_n(\theta)$$

$$= \int h(\theta - D_n(\epsilon))dF_n(\theta) - \inf_D \int h(\theta - D)dF_n(\theta)$$

$$\geq \int h(\theta - D_n(\epsilon))dF_n(\theta) - \int h(\theta - D_n(\epsilon) - 2n)dF_n(\theta)$$

$$= \frac{1}{n}(2n + D_n(\epsilon) - D_n(\epsilon)) + (1 - 1/n)\int h(\theta - D_n(\epsilon))dH(\theta)$$

$$- \quad (1 - 1/n)\int h(\theta - D_n(\epsilon) - 2n)dH(\theta)$$

$$\geq 2 - \int_{-\infty}^{D_n(\epsilon)+2n} h(\theta - D_n(\epsilon) - 2n)dH(\theta) - \int_{D_n(\epsilon)+2n}^{\infty} h(\theta - D_n(\epsilon) - 2n)dH(\theta)$$

$$\geq 2 - 1 - \int_{D_n(\epsilon)+2n}^{\infty} \theta dH(\theta) + \int_{D_n(\epsilon)+2n}^{\infty} (2n + D_n(\epsilon))dH(\theta).$$

Now since $D_n(\epsilon)$ is bounded below by d^*-2, $D_n(\epsilon)+2n \to \infty$. The existence of the mean of H implies that

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

So the first integral above approaches zero. Similarly the existence of the mean also implies that

$$\int_{D_n(\epsilon)+2n}^{\infty} \big(2n+D_n(\epsilon)\big)dH(\theta) = (2n+D_n(\epsilon))\big(1 - H(2n+D_n(\epsilon))\big) \to 0.$$

(see Feller, 1966, p. 149, line 5). Hence

$$\lim_{\epsilon\downarrow 0} \limsup_{n\to\infty} \left[\int V\big(\theta,D_n(\epsilon)\big)dF_n(\theta) - \inf_D \int V(\theta,D)\,dF_n(\theta)\right] \geq 1,$$

so (V,H) is unstable.

<u>Lemma 3</u>: The pair (V,H) is strongly stable if, in addition to conditions (i) and (ii), the following condition (iii) obtains:

(iii) There is a compact interval $[a,b]$ and an $\epsilon_0 > 0$ such that for every ϵ, $0 < \epsilon < \epsilon_0$, every sequence $G_n \overset{w}{\Rightarrow} H$, and every sequence of ϵ-optimal decisions $D_1, D_2, \ldots$ for (V, G_n), there is an N such that for all $n > N$, $D_n \in [a,b]$.

<u>Proof</u>: 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 $\epsilon > 0$, there exists a $\delta > 0$ such that for every $x,y \in [a,b]$, $|h(x)-h(y)| < \epsilon$ if $|x-y| < \delta$. Choose $\delta < (b-a)/2$. Now there is a finite open covering of $[a,b]$ $\{(c_i,d_i) \mid i=1,2,\ldots,K\}$ such that $d_i - c_i < \min\{\delta,\epsilon\}$ for all $i=1,2,\ldots,K$. Let $e_i \in (c_i,d_i)$. We now proceed to show that $|h(\theta-e_i) - h(\theta-e_j)|$ is bounded. Without loss of generality, let $e_i > e_j$. Also let $D = \underset{i=1,\dots,K}{\mathrm{Max}}\ d_i$ and $C = \underset{i=1,\dots,K}{\mathrm{Min}}\ c_i$.

(a) If $\theta \geq e_i + b$, then

$$|h(\theta-e_i)-h(\theta-e_j)| \leq B|\ (\theta-e_i) - (\theta-e_j)| = B|e_i-e_j| \leq B(D-C).$$

(b) If $e_i + b > \theta \geq e_j + b$, then

$$\begin{aligned}|h(\theta-e_i) - h(\theta-e_j)| &\leq |h(\theta-e_i) - h(b)|+|h(b) - h(\theta-e_j)| \\ &\leq h(b) + h(b+D-C) + B|b - \theta + e_j| \\ &\leq h(b) + h(b+D-C) + B(D-C).\end{aligned}$$

(c) If $e_j + b > \theta \geq a + e_i$, then

$$|h(\theta-e_i) - h(\theta-e_j)| \leq h(b) + h(a)$$

since $(\theta-e_i) \in [a,b]$, $(\theta-e_j) \in [a,b]$.

(d) If $e_i + a > \theta \geq a + e_j$, then

$$|h(\theta-e_i) - h(\theta-e_j)| \leq |h(\theta-e_i) - h(a)| + |h(a) - h(\theta-e_j)|$$

$$\begin{aligned}&\leq B|\theta - e_i - a| + h(a) + h(a + D-C) \\ &= B(a + e_i - \theta) + h(a) + h(a+D-C) \\ &\leq B(D-C) + h(a) + h(a+D-C).\end{aligned}$$

(e) If $a + e_j > \theta$, then

$$|h(\theta - e_i) - h(\theta - e_j)| \leq B|e_i - e_j| \leq B(D-C).$$

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

$$| \int (V(e_i,\theta) - V(e_j,\theta))dF_n(\theta) - \int (V(e_i,\theta) - V(e_j,\theta))dH(\theta)| < \epsilon,$$

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

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

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

Now suppose $t_1 \epsilon (c_i d_i)$ and $t_2 \epsilon (c_i, d_i)$ 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 $\theta \geq t_1 + b$, then

$$|h(\theta-t_1) - h(\theta-t_2)| \leq B|t_1-t_2| \leq B\epsilon.$$

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

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

(c) If $t_2 + b > \theta \geq a + t_1$, then

$b > \theta - t_2 > \theta - t_1 \geq a$, and

$|(\theta - t_1) - (\theta - t_2)| = |t_1 - t_2| < \delta$.

Thus $|h(\theta - t_1) - h(\theta - t_2)| < \epsilon$.

(d) If $a + t_1 > \theta \geq a + t_2$, then

$\delta > t_1 - t_2 > \theta - a - t_2 \geq 0$. Thus

$$|h(\theta - t_1) - h(\theta - t_2)| \leq |h(\theta - t_1) - h(a)| + |h(a) - h(\theta - t_2)|$$

$$\leq B(a + t_1 - \theta) + \epsilon \leq (B+1)\epsilon.$$

(e) $a + t_2 > \theta$.

$$|h(\theta - t_1) - h(\theta - t_2)| \leq B|t_1 - t_2| \leq B\epsilon.$$

Thus for all θ,

$$|h(\theta - t_1) - h(\theta - t_2| \leq (B+1)\epsilon.$$

Let $d \in [a,b]$. Then there is an ℓ such that $d \in (c_\ell, d_\ell)$. Let $n > N_0$. There is an m such that $D_n(\epsilon) \in (c_m, d_m)$. Then

$$\int (V(d,\theta) - V(D_n(\epsilon),\theta))dF_n(\theta)$$

$$= \int [V(d,\theta) - V(e_\ell,\theta) + V(e_\ell,\theta) - V(D_n(\epsilon),\theta) + V(e_m,\theta)$$

$$- V(e_m,\theta)]dF_n(\theta)$$

$$\geq -2(B+1)\epsilon + \int (V(e_\ell,\theta) - V(e_m,\theta))dF_n(\theta)$$

$$\geq -2(B+1)\epsilon + \int (V(e_\ell,\theta) - V(e_m,\theta)) dH(\theta) - \epsilon$$

$$\geq -2(B+1)\epsilon + \int (V(e_\ell,\theta) - V(e_m,\theta))dG_n(\theta) - 2\epsilon$$

$$\geq -2(B+2)\epsilon + \int (V(d,\theta) - V(D_n(\epsilon),\theta))dG_n(\theta) - 2(B+1)\epsilon$$

$$\geq -(4B + 7)\epsilon$$

Then $\forall n \geq N_0$

$$\inf_{d\epsilon[a,b]} \int V(d,\theta)dF_n(\theta) - \int V(D_n(\epsilon),\theta)dF_n(\theta) \geq -(4B+7)\epsilon.$$

Now $F_n \xrightarrow{w} H$, so if $D_n^*(\epsilon)$ is a sequence of ϵ-optimal decisions for (F_n,V) then $\exists N \ni \forall n \geq N$, $D_n(\epsilon) \epsilon[a,b]$. Thus $\forall n > \text{Max}(N,N_0)$,

$$\inf_{d\epsilon[a,b]} \int V(d,\theta)dF_n(\theta) = \inf_d \int V(d,\theta)dF_n(\theta).$$

Thus $\forall n > \text{Max}(N,N_0)$,

$$\inf_d \int V(d,\theta)dF_n(\theta) - \int V(D_n(\epsilon),\theta)dF_n(\theta) \geq -(4B+7)\epsilon.$$

Hence

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

Thus (V,H) is stable.

Q.E.D.

Theorem 3 The pair (V,H) is strongly stable if, in addition to conditions (i) and (ii), the following condition (iv) obtains:

(iv): there exist $r > 0$ such that $h(x) \geq r|x|$, $\forall x$

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 $\dfrac{H(b) - H(-b)}{1 - H(b)} > \dfrac{2B}{r}$.

Let $D^* = 0$ and $D > \dfrac{h(-y_0) + h(y_0) + rb + \epsilon_0}{H(b) - H(-b)} \cdot \dfrac{2}{r}$.

It is straight forward to show $\frac{D}{2} > b > y_0$.

$$\int_{-\infty}^{\infty} (h(\theta - D^*) - h(\theta - D))dH(\theta)$$

$$= \left(\int_{-\infty}^{-b} + \int_{-b}^{-y_0} + \int_{-y_0}^{y_0} + \int_{y_0}^{b} + \int_{b}^{D-y_0} + \int_{D-y_0}^{D+y_0} + \int_{D+y_0}^{\infty} \right) (h(\theta) - $$

$$h(\theta - D))dH(\theta) = I_1 + I_2 + I_3 + I_4 + I_5 + I_6 + I_7.$$

$$I_1 = \int_{-\infty}^{-b} (h(\theta) - h(\theta - D))dH(\theta) \leq 0.$$

$$I_2 = \int_{-b}^{-y_0} (h(\theta) - h(\theta - D))dH(\theta) \leq [B(b - y_0) + h(-y_0) -$$

$$r(y_0 + D)] \int_{-b}^{-y_0} dH(\theta) .$$

$$I_3 = \int_{-y_0}^{y_0} (h(\theta) - h(\theta - D))dH(\theta)$$

$$\leq (h(y_0) + h(-y_0) - r(D - y_0)) \int_{-y_0}^{y_0} dH(\theta).$$

$$I_4 = \int_{y_0}^{b} (h(\theta) - h(\theta - D))dH(\theta) \leq [h(y_0) + B(b - y_0) - r(D - b)]$$

$$\int_{y_0}^{b} dH(\theta).$$

$$I_5 = \int_{b}^{D-y_0} (h(\theta) - h(\theta-D))dH(\theta) \leq (h(y_0) + B(D - 2y_0) - ry_0) \int_{b}^{D-y_0} dH(\theta).$$

$$I_6 = \int_{D-y_0}^{D+y_0} (h(\theta) - h(\theta-D))dH(\theta) \leq (h(y_0) + BD) \int_{D-y_0}^{D+y_0} dH(\theta).$$

$$I_7 = \int_{D+y_0}^{\infty} (h(\theta) - h(\theta-D))dH(\theta) \leq BD \int_{D+y_0}^{\infty} dH(\theta).$$

Putting all the pieces together, we have

$$\int_{-\infty}^{\infty} (h(\theta-D^*) - h(\theta-D))dH(\theta) \leq$$

$$[B(b-y_0) + h(-y_0) - r(y_0+D)] \int_{-b}^{-y_0} dH(\theta)$$

$$+ [h(y_0) + h(-y_0) - r(D-y_0)] \int_{-y_0}^{y_0} dH(\theta)$$

$$+ [h(y_0) + Bb - r(D-b)] \int_{y_0}^{b} dH(\theta)$$

$$+ [h(y_0) + BD - ry_0] \int_b^{D-y_0} dH(\theta)$$

$$+ [h(y_0) + BD] \int_{D-y_0}^{D+y_0} dH(\theta)$$

$$+ BD \int_{D+y_0}^{\infty} dH(\theta)$$

$$\leq \quad h(-y_0) + h(y_0) + BD(1 - H(b)) - rD(H(b) - H(-b)) + rb$$

$$\leq \quad h(-y_0) + h(y_0) + rb - \frac{rD}{2} (H(b) - H(-b))$$

$$-rD(H(b) - H(-b))$$

$$< \quad -\epsilon.$$

Similarly if $D < - \dfrac{h(-y_0) + h(y_0) + rb + \epsilon_0}{H(b) - H(-b)} \cdot \dfrac{2}{r}$,

then

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

So any ϵ-optimal decision D_ϵ for H must satisfy

$$|D_\epsilon| < \frac{h(-y_0) + h(y_0) + rb + \epsilon_0}{H(b) - H(-b)} \cdot \frac{2}{r} .$$

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 $J_n \xrightarrow{w} H$. Then $\exists$ N $\ni \forall n \geq N$,

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

and

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

Let $m = 2(h(-y_0) + h(y_0) + rb_1 + \epsilon_0)/r$. The ϵ-optimal decisions for (J_n, V) for all $n > N$ is within

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

and hence within

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

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

Q.E.D.

Corollary 1

Let $I(\cdot)$ be the usual indicator function. Then $V(\theta,D) = a(\theta-D)I(\theta \geq D) + b(D-\theta)I(\theta < D)$ is strongly stable with any H such that $\int V(\theta,D)\,dH(\theta)$ is finite for some D.

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

$$h(x) = \begin{cases} x & \text{if } 0 \le x \le 1 \\ 1 & \text{if } 1 < x \le 2\cdot(2)^3 \\ (x-2j^{j+1})\cdot \frac{1}{j} + (j-1)^{j-1} & \text{if } 2j^{j+1} < x \le 3j^{j+1} - j(j-1)^{j-1} \\ j^j & \text{if } 3j^{j+1} - j(j-1)^{j-1} < x \le 2(j+1)^{j+2} \end{cases}$$

for $j = 2, 3, \ldots$

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

Then h is continuous, symmetric, piece-wise linear, non-decreasing in $(0,\infty)$, non-increasing 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 let $G_n \equiv H$. Let $F_n(\theta)$ be a convex combination of $H(\theta)$ and $J_n(\theta)$ with weights $(1 - \frac{1}{n})$ and $\frac{1}{n}$, 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 $F_n \xrightarrow{w} H$, and $D_n(\epsilon) \epsilon (-\epsilon, \epsilon)$ where $\epsilon < 1$. Also

$$\int V(\theta, D_n(\epsilon))\, dF_n(\theta) - \inf_D \int V(\theta, D)\, dF_n(\theta)$$

$$\ge \int V(\theta, D_n(\epsilon))\, dF_n(\theta) - \int V(\theta, 2(n)^{n+1})\, dF_n(\theta)$$

$$\ge \frac{1}{n} h(3n^{n+1} - n(n-1)^{n-1}) - \epsilon - \frac{1}{n} h(3n^{n+1} - n(n-1)^{n-1} - 2n^{n+1})$$

$$- \frac{n-1}{n} h(2n^{n+1})$$

$$= \frac{1}{n} n^n - \epsilon - \frac{1}{n} h(n^{n+1} - n(n-1)^{n-1}) - \frac{n-1}{n} h(2n^{n+1})$$

$$\geq n^{n-1} - \epsilon - \frac{1}{n} h(2n^{n+1}) - \frac{n-1}{n} h(2n^{n+1})$$

$$\geq n^{n-1} - \epsilon - (n-1)^{n-1} > 1 .$$

Thus it is easy to see that (V,H) is unstable in this case.

Theorem 4 (V,H) is strongly stable if, in addition to assumptions (i) and (ii), the following condition (v) is satisfied:

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

Proof:

Our strategy is to apply Lemma 3 by proving condition (iii). Choose $\epsilon_0 > 0$, and ϵ such that $0 < \epsilon < \epsilon_0$.

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

$$h(D_0) > 2h(b) + Bb + 4\epsilon_0.$$

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

$$I = \int V(\theta,D^*)dH(\theta) - \int V(\theta,D)dH(\theta)$$

$$= \int_{-\infty}^{-b} (h(\theta-D^*) - h(\theta-D))dH(\theta) + \int_{-b}^{b} (h(\theta-D^*) - h(\theta-D))dH(\theta)$$

$$+ \int_{b}^{\infty} h(\theta - D^*) - h(\theta - \mathbf{D})) \, dH(\theta) = I_1 + I_2 + I_3.$$

$$I_1 = \int_{-\infty}^{-b} (h(\theta - D^*) - h(\theta - D)) dH(\theta) \leq 0$$

since $h(\theta - D^*) - h(\theta - D) \leq 0$ if θ if $\theta \in (-x, -b)$.

In the second region of integration, $(-b, b)$, we have $-b < \theta - D^* < b$. Then $h(\theta - D^*) \leq h(b)$. Also $\theta - D < b - (b + D_0) = -D_0 < 0$. Hence

$$I_2 = \int_{-b}^{b} (h(\theta - D^*) - h(\theta - D)) dH(\theta)$$

$$\leq \int_{-b}^{b} (h(b) - h(b - D)) dH(\theta)$$

$$= [h(b) - h(D - b)][H(b) - H(-b\ )]$$

$$\leq \tfrac{1}{2}[h(b) - h(D - b)].$$

$$I_3 = \int_{b}^{\infty} (h(\theta - D^*) - h(\theta - D)) dH(\theta)$$

$$\leq \int_{b}^{\infty} h(D - D^*) dH(\theta)$$

$$\leq \frac{1}{4} h(D)$$

Hence

$$I = I_1 + I_2 + I_3$$

$$\leq \tfrac{1}{2}[h(b) - h(D-b)] + \frac{1}{4}h(D)$$

$$= \tfrac{1}{2}h(b) - \frac{1}{4}h(D-b) + \frac{1}{4}(h(D) - h(D-b))$$

$$\leq \tfrac{1}{2}h(b) - \frac{1}{4}h(D-b) + \frac{1}{4}Bb$$

$$\leq -\epsilon_0 .$$

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 $G_n \xrightarrow{w} H$. There is a point b_1 such that both b_1 and $-b_1$ are continuity points of H satisfying $b_1 > b$, $H(b_1) \geq 7/8$, and $H(-b_1) \leq 1/8$. Let D_1 satisfy

$$h(D_1) > 2h(b_1) + Bb_1 + 4\epsilon_0.$$

Since $G_n \xrightarrow{w} H$, there is an N such that $\forall n > N$, $G_n(-b_1) \leq 1/4$ and $G_n(b_1) \geq 3/4$. Then for all such n, $D_n(\epsilon) \in (-b_1 - D_1, b_1 + D_1)$.

Lemma 3 now applies, so (V,H) is stable.

Q.E.D.

<u>Corollary 2</u>: If $V(\theta,D) = |\theta - D|^p$ $0 < p \leq 1$ then (V,H) is strongly stable. The next example shows the effect of asymmetry.

<u>Example 7</u>: Let $V(\theta,D) = h(\theta-D)$, where

$$h(x) = \begin{cases} x^{\frac{1}{2}} & x \geq 0 \\ |x|^{1/3} & x < 0 \end{cases}$$

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 in this case.

Conclusion

We are studying then a particular kind of continuity, is studied a kind we judge ~~to be~~ important especially as a prologue to attempting elicitation of prior distributions and utility functions. There are other kinds of continuity that are alternatives to those ~~we have~~ chosen, and which also deserve study.

~~We believe that~~ stability, as defined in this paper, offers the beginning of a Bayesian approach to robustness. ~~We note that~~ on the real line, squared error loss is never stable, while absolute error is strongly stable for all opinions H. While the approach in this paper is more mathematical than some other approaches to robustness, which may be a disadvantage, it has the advantage of starting from a clear philosophical foundation, namely personalistic Bayesianism.

References

[1] Antleman, G. R. 1965 "Insensitivity to Non-optimal Design in Bayesian Decision Theory" JASA 60, 584-601.

[2] Britney, R. and R. Winkler 1974 "Bayesian Point Estimation and Prediction" Annals of the Institute of Statistical Mathematics 26, 15-34.

[3] Dempster, A. 1975 "A Subjectivist Look at Robustness" Research Report S-33, Department of Statistics, Harvard University.

[4] Edwards, W., H. Lindeman and L. J. Savage 1963 "Bayesian Statistical Inference for Phychological Research" Phychological Review 70, 193-242.

[5] Feller, W. 1966 An Introduction to Probability Theory and Its Applications, Volume II, John Wiley and Sons, New York.

[6] Fishburn, P., A. Murphy and H. Isaacs 1967 "Sensitivity of Decision to Probability Estimation Errors: A Reexamination" Operations Research 15, 254-267.

[7] Pierce, D. and J. L. Folks 1969 "Sensitivity of Bayes Procedures to the Prior Distribution" Operations Research 17, 344-350.

[8] Schlaifer, R. 1959 Probability and Statistics for Business Decisions McGraw-Hill, New York.

[9] Stone, M. 1963 "Robustness of Non-ideal Decision Procedures" JASA 58, 480-486.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ONR Report No.5	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Stable Decision Problems		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER Technical Report No. 107
7. AUTHOR(s) Joseph B. Kadane and David Chuang		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0516 new
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Carnegie-Mellon University Pittsburgh, Pa. 15213		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR042-309
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Headquarters Arlington, Virginia 22217		12. REPORT DATE June , 1976
		13. NUMBER OF PAGES
14. MONITORING AGENCY NAME & ADDRESS(If different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Decision Theory, Robustness, Stable Estimation, Stable Decisions.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

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 $F_n \overset{w}{\to} H$, $G_n \overset{w}{\to} H$ and $L_n \to V$, $W_n \to V$ uniformly,

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

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

for every sequence $D_n(\varepsilon)$ satisfying

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

We show that squared error loss is unstable with any pair if the parameter space is the real line, 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, non-decreasing in $(0, \infty)$ and non-increasing 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.