The Annals of Mathematical Statistics
1972, Vol. 43, No. 1, 348–350

A NOTE ON ROBBINS' COMPOUND DECISION PROCEDURE

BY J. S. HUANG

University of Guelph

In The Second Berkeley Symposium Robbins raised the question whether or not his "bootstrap" procedure might be dominated by a Bayes procedure against the prior which puts equal mass on each "orbit" and is uniform within each orbit. We show that this is not true for the compound problem with two components.

Robbins [1] introduced the compound decision problem and investigated within the context of a particular component problem the asymptotic compound risk behavior of certain compound procedures. Let $X_1, \cdots, X_n$ represent the data from the n component problems and $p(\boldsymbol{\theta})$ denote the empirical distribution of the n unknown states $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_n)$. Robbins showed that it is possible to estimate $p(\boldsymbol{\theta})$ by $\hat{p}$ based on $X_1, \cdots, X_n$ and that the compound procedure R^* which plays Bayes versus $\hat{p} \times \cdots \times \hat{p}$ in the compound problem has compound risk $L(R^*, \boldsymbol{\theta})$ asymptotically below that of the procedure which plays minimax within each component. Robbins called the procedure R^* a "bootstrap" procedure. Since R^* is inadmissible it does raise an interesting question, what are the procedures that dominate R^*? As a possible candidate, Robbins (page 140 of [1]) suggested the procedure $\bar{R}$ which plays Bayes versus a "diffuse" prior (see (33) of [1]). The purpose of this note is to show that $\bar{R}$ does not dominate R^* for the $n = 2$ case.

For $n = 2$, the problem is the following. X_i, $i = 1, 2$, are independent random variables distributed by $N(\theta_i, 1)$, with $\theta_i \in \{-1, 1\} \equiv \Omega$. Upon observing $\boldsymbol{X} = (X_1 X_2)$ it is required to guess the means $\boldsymbol{\theta} = (\theta_1, \theta_2)$, incurring the risk $L(\cdot, \boldsymbol{\theta})$, the average number of components of $\boldsymbol{\theta}$ misguessed.

For $\frac{1}{4} \leqq \lambda \leqq \frac{1}{2}$ let β_λ be the (symmetric) prior on Ω^2 defined by

$$\begin{aligned} \beta_\lambda(\boldsymbol{\theta}) &= \lambda \qquad \text{for} \quad \boldsymbol{\theta} = (1, 1), (-1, -1); \\ &= \tfrac{1}{2} - \lambda \qquad \text{for} \quad \boldsymbol{\theta} = (1, -1), (-1, 1). \end{aligned}$$

Elementary calculation shows that β_λ leads to the following posterior marginal distributions:

$$\Pr(\theta_i = 1 \mid \boldsymbol{x}) = [\lambda A_1 A_2 + (\tfrac{1}{2} - \lambda)A_i][\lambda(A_1 A_2 + 1) + (\tfrac{1}{2} - \lambda)(A_1 + A_2)]^{-1},$$

where $A_i \equiv \exp 2x_i$, $i = 1, 2$. Let B_λ be a Bayes procedure versus β_λ. Since the loss function is symmetric, the Bayes procedure B_λ thus guesses θ_i to be a value maximizing the posterior probability of θ_i, and arbitrary if

$$\text{(1.i)} \qquad \lambda A_1 A_2 + (-1)^{i+1}(\tfrac{1}{2} - \lambda)(A_1 - A_2) - \lambda = 0.$$

Received December 3, 1970.

From (37) of [1] we find the bootstrap procedure R^* to be

$$R^*: \text{“}\theta_i = \operatorname{sgn}(x_i - x^*),\text{”}$$

where

$$\begin{aligned} x^* &= \infty && \text{if} \quad \bar{x} \leqq -1 \\ &= \frac{1}{2} \ln \frac{1-\bar{x}}{1+\bar{x}} && \text{if} \quad -1 < \bar{x} < 1 \\ &= -\infty && \text{if} \quad 1 \leqq \bar{x}\,, \end{aligned}$$

and $\bar{x} = 2^{-1}(x_1 + x_2)$.

For $-1 < \bar{x} < 1$ we see that $x_i = x^*$ iff

(2.i) $$A_i(2 + x_1 + x_2) = 2 - x_1 - x_2\,.$$

For $\lambda = \frac{1}{3}$, β_λ is called the "diffuse" prior, and $B_\lambda = \bar{R}$. To compare $\bar{R}$ with R^*, we set $\lambda = \frac{1}{3}$ and graph the four curves (1-1), (1-2), (2-1) and (2-2).

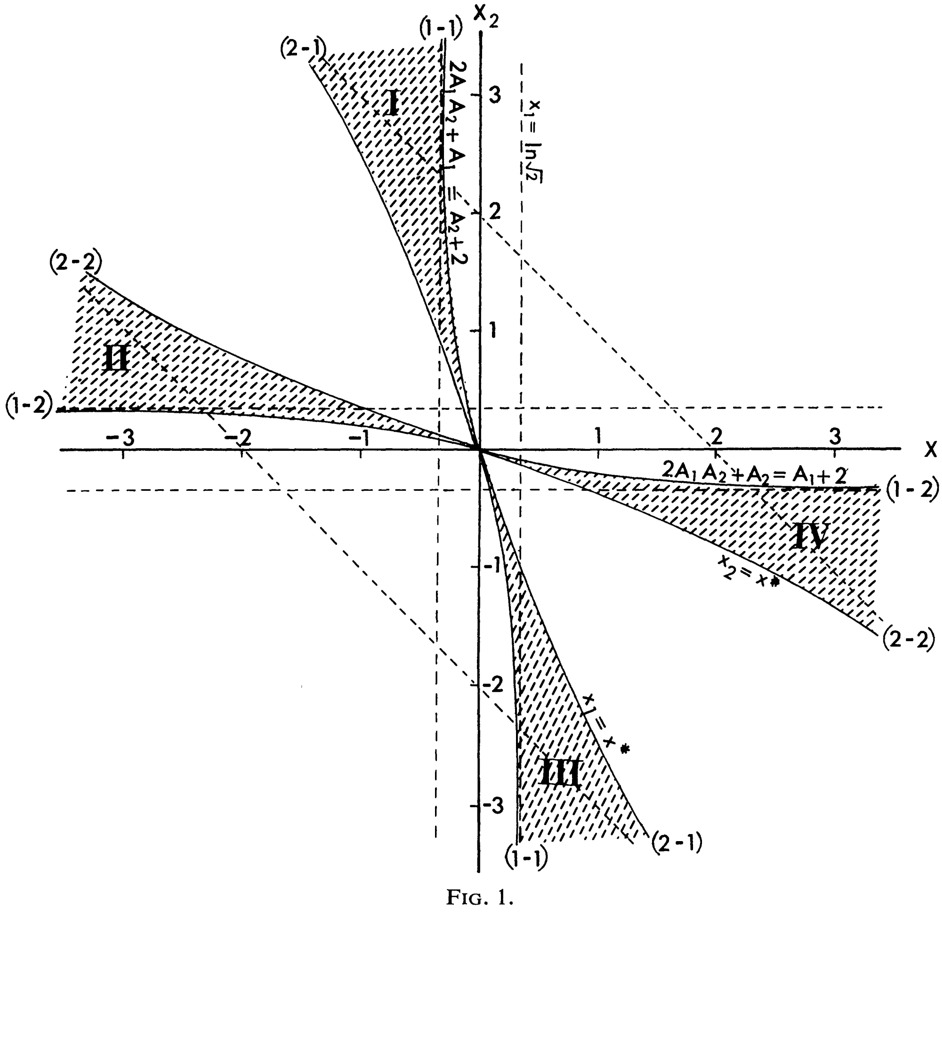

FIG. 1.

Clearly all curves pass through the origin. By elementary calculation we see that the slopes dx_2/dx_1 of (1-1) and (2-1) are both -3 at the origin. By computing the second derivatives, it is easily verified that for $x_1 > 0$, (1-1) is concave downward and (2-1) is concave upward, and vice versa for $x_1 < 0$. Since (1-1) and (1-2) are symmetric with respect to the line $x_1 + x_2 = 0$, and likewise (2-1) and (2-2), it thus follows that the four curves do not intersect elsewhere. From the defining properties of R^* and $\bar{R}$ we see that the two procedures yield different decisions when and only when $\boldsymbol{x}$ is observed in the shaded regions designated as I, II, III and IV.

At $\boldsymbol{\theta} = (1, 1)$, it is clear that Region I + IV has higher probability than Region II + III. (Note again the symmetry, and consider the conditional distribution of $U = 2^{-\frac{1}{2}}(X_1 + X_2)$ given $V = 2^{-\frac{1}{2}}(X_1 - X_2)$). Thus $L(\bar{R}, \boldsymbol{\theta}) > L(R^*, \boldsymbol{\theta})$ for $\boldsymbol{\theta} = (1, 1)$ and therefore $\bar{R}$ does not dominate R^*. Likewise at $\boldsymbol{\theta} = (-1, -1)$. At $\boldsymbol{\theta} = (-1, 1)$ and $(1, -1)$, however, the inequality is reversed due to the Bayesness of $\bar{R}$.

It could be shown, via above geometrical argument, that R^* is dominated by $\{B_\lambda \mid \lambda' \leqq \lambda \leqq \lambda''\}$, for some λ', λ'' with $\frac{1}{3} < \lambda' < \lambda'' < \frac{1}{2}$.

Acknowledgment. The author wishes to thank Dr. Dennis C. Gilliland for his helpful discussions.

REFERENCE

[1] Robbins, H. (1951). Asymptotically subminimax solutions of compound statistical decision problems. *Proc. Second Berkeley Symp. Math. Statist. Prob.*, 131–148.