ON MINIMAX STATISTICAL DECISION PROCEDURES AND THEIR ADMISSIBILITY[1]

By Colin R. Blyth

University of California, Berkeley

Summary. This paper is concerned with the problem of making a decision on the basis of a sequence of observations on a random variable. Two loss functions, each depending on the distribution of the random variable, the number of observations taken, and the decision made, are assumed given. Minimax problems can be stated for weighted sums of the two loss functions, or for either one subject to an upper bound on the expectation of the other. Under suitable conditions it is shown that solutions of the first type of problem provide solutions for all problems of the latter types, and that admissibility for a problem of the first type implies admissibility for problems of the latter types. Two examples are given: Estimation of the mean of a random variable which is (1) normal with known variance, (2) rectangular with known range. The resulting minimax estimates are, with a small class of exceptions, proved admissible among the class of all procedures with continuous risk functions. The two loss functions are in each case the number of observations, and an arbitrary nondecreasing function of the absolute error of estimate. Extensions to a function of the number of observations for the first loss function are indicated, and two examples are given for the normal case where the sample size can or must be randomised among more than a consecutive pair of integers.

1. Introduction. We will consider a sequence $X_1, X_2, X_3, \cdots$ of independent random variables, each having the same distribution F, which belongs to a class Ω of possible distributions. A sequential decision procedure S is a rule by which, having observed $x_1, \cdots, x_m (m = 0, 1, 2, \cdots)$ we make one of the following decisions:

(a) Take an observation on X_{m+1}.

(b) Stop experimentation and make a terminal decision $d = d(x_1, \cdots, x_m)$.

We will consider two non-negative loss functions $W_1(n, d, F)$ and $W_2(n, d, F)$. Each can be thought of as an economic loss incurred when the X's have distribution F and the terminal decision d is made after n observations have been taken. In the simplest applications one W will be a function of n only (cost of experimentation) and the other W will be a function of d and F only (loss incurred by making the decision d when the X's have distribution F). We will denote by $E(W_i \mid F, S)$ the expected value of W_i when the X's have distribution F and the decision procedure S is used. Let ξ be any probability measure defined on some class of subsets of Ω. We will denote by $E(W_i \mid \xi, S)$ the expected value

[1] Work done under the sponsorship of the Office of Naval Research, and submitted in partial satisfaction of the requirements for the degree of Ph.D. at the University of California, Berkeley.

of W_i, given the (*a priori*) distribution ξ over Ω, when the decision procedure S is used.

Minimax problems, first considered by Wald, have been formulated in three ways for the situation just described. We may seek a decision procedure S which will (i) subject to an upper bound on $E(W_1 \mid F, S)$, minimise $\sup_\Omega E(W_2 \mid F, S)$; or (ii) subject to an upper bound on $E(W_2 \mid F, S)$, minimise $\sup_\Omega E(W_1 \mid F, S)$; or (iii) minimise $\sup_\Omega \{c\, E(W_1 \mid F, S) + E(W_2 \mid F, S)\}$, where $0 < c < \infty$, c constant. We will show that in certain cases it suffices to find solutions for all problems (iii) since these solutions provide solutions for all problems (i) and (ii).

We will first discuss the corresponding Bayes problems, not for their own interest, but because they can often be used to find solutions for the minimax problems in which we are really interested.

2. Bayes problems. The following three classes of Bayes problems have been considered: Given *a priori* the distribution ξ over Ω, we want to find a (Bayes) procedure which will

(i)′ subject to $E(W_1 \mid \xi, S) \leqq L_1$, minimise $E(W_2 \mid \xi, S)$,

(ii)′ subject to $E(W_2 \mid \xi, S) \leqq L_2$, minimise $E(W_1 \mid \xi, S)$,

(iii)′ minimise $r_c(\xi, S) = cE(W_1 \mid \xi, S) + E(W_2 \mid \xi, S)$.

Let $\mathcal{S}_c$ be the class of all solutions of problem (iii)′ for a given c, $0 < c < \infty$. Let $\mathcal{S} = \bigcup_{0<c<\infty} \mathcal{S}_c$ be the class of all solutions of problems (iii)′, $0 < c < \infty$.

LEMMA 1. *If $S' \in \mathcal{S}$ has $E(W_1 \mid \xi, S') = L_1$, then S' is a solution of the problem* (i)′ *for this L_1. If S'' is any other solution of this problem* (i)′, *then $E(W_1 \mid \xi, S'') = L_1$ and $S'' \in \mathcal{S}$. Similarly for problems* (ii)′.

PROOF. Let $S' \in \mathcal{S}_c$. Suppose there exists a procedure S^* having

$$E(W_1 \mid \xi, S^*) \leqq E(W_1 \mid \xi, S') = L_1,$$

$$E(W_2 \mid \xi, S^*) < E(W_2 \mid \xi, S').$$

Then

$$cE(W_1 \mid \xi, S^*) + E(W_2 \mid \xi, S^*) < cE(W_1 \mid \xi, S') + E(W_2 \mid \xi, S').$$

This implies $S' \notin \mathcal{S}_c$, which is false. This contradiction shows that no such S^* can exist. Hence S' is a solution of this problem (i)′.

If S'' is any other solution of this problem (i)′ we must have $E(W_2 \mid \xi, S'') = E(W_2 \mid \xi, S')$. Suppose that $E(W_1 \mid \xi, S'') < E(W_1 \mid \xi, S') = L_1$. Then

$$cE(W_1 \mid \xi, S'') + E(W_2 \mid \xi, S'') < cE(W_1 \mid \xi, S') + E(W_2 \mid \xi, S'),$$

implying the contradiction $S' \notin \mathcal{S}_c$. Hence $E(W_1 \mid \xi, S'') = E(W_1 \mid \xi, S') = L_1$. We therefore have $r_c(\xi, S'') = r_c(\xi, S')$, and so $S'' \in \mathcal{S}_c$.

LEMMA 2. *If* $S' \in \mathcal{S}$, $S'' \in \mathcal{S}$, *then*

$$E(W_1 \mid \xi, S') < E(W_1 \mid \xi, S'') \longleftrightarrow E(W_2 \mid \xi, S') > E(W_2 \mid \xi, S''),$$

$$E(W_1 \mid \xi, S') = E(W_1 \mid \xi, S'') \longleftrightarrow E(W_2 \mid \xi, S') = E(W_2 \mid \xi, S'').$$

LEMMA 3. *If* $S' \in \mathcal{S}_{c'}$, *and* $S'' \in \mathcal{S}_{c''}$ *where* $c' < c''$, *then*

$$E(W_1 \mid \xi, S') \geqq E(W_1 \mid \xi, S''),$$

$$E(W_2 \mid \xi, S') \leqq E(W_2 \mid \xi, S'').$$

PROOF. Assume one of the following:

$$L_1 = E(W_1 \mid \xi, S') < E(W_1 \mid \xi, S'') = L_1 + r,$$

$$L_2 = E(W_2 \mid \xi, S') > E(W_2 \mid \xi, S'') = L_2 - s.$$

The other then follows from Lemma 2. Write $c'' = c' + a$. Here $r > 0$, $s > 0$, $a > 0$. Then

$$r_{c'}(\xi, S') = c'L_1 + L_2,$$

$$r_{c'}(\xi, S'') = c'L_1 + L_2 + (c'r - s),$$

$$r_{c''}(\xi, S') = c'L_1 + L_2 + aL_1,$$

$$r_{c''}(\xi, S'') = c'L_1 + L_2 + aL_1 + (c'r - s + ar).$$

Now

$$S' \in \mathcal{S}_{c'} \rightarrow c'r - s \geqq 0,$$

and

$$S'' \in \mathcal{S}_{c''} \rightarrow c'r - s + ar \leqq 0.$$

Since $ar > 0$ these last two results cannot both be true. This contradiction shows that neither of the assumed inequalities can be true, and proves the lemma.

Let us write

$$\underline{L}_1 = \inf_{S \in \mathcal{S}} E(W_1 \mid \xi, S), \qquad \bar{L}_1 = \sup_{S \in \mathcal{S}} E(W_1 \mid \xi, S),$$

$$\underline{L}_2 = \inf_{S \in \mathcal{S}} E(W_2 \mid \xi, S), \qquad \bar{L}_2 = \sup_{S \in \mathcal{S}} E(W_2 \mid \xi, S),$$

where the improper value ∞ is admitted for the upper bounds.

LEMMA 4.

$$E(W_1 \mid \xi, S) < \underline{L}_1 \rightarrow E(W_2 \mid \xi, S) = \infty,$$

$$E(W_2 \mid \xi, S) < \underline{L}_2 \rightarrow E(W_1 \mid \xi, S) = \infty.$$

PROOF. Suppose that S is a procedure for which $E(W_1 \mid \xi, S) = L_1 < \underline{L}_1$ and $E(W_2 \mid \xi, S) = L_2 < \infty$.

If $\bar{L}_2 = \infty$, there exists some $S_c \in \mathcal{S}_c$ having $E(W_1 \mid \xi, S_c) \geqq \underline{L}_1$ and $E(W_2 \mid \xi, S_c) > L_2$; but we would then have $r_c(\xi, S_c) > r_c(\xi, S)$, contradicting the fact that $S_c \in \mathcal{S}_c$.

If $\bar{L}_2 < \infty$, then for $S_c \in \mathcal{S}_c$ we have

$$cE(W_1 \mid \xi, S_c) + E(W_2 \mid \xi, S_c) \geqq c\underline{L}_1 + \underline{L}_2 > cL_1 + L_2$$

for c sufficiently large, again contradicting the fact that $S_c \in \mathcal{S}_c$. This completes proof of the first part of the lemma; the second part is proved in the same way.

Lemma 4 shows that no problem (i)′ with $L_1 < \underline{L}_1$ has a solution. Lemmas 2 and 4 show that if $S \in \mathcal{S}$ has $E(W_1 \mid \xi, S) = \bar{L}_1$, then $E(W_2 \mid \xi, S) = \underline{L}_2$ and S is a solution of all problems (i)′ with $L_1 \geqq \bar{L}_1$. Similar remarks hold for problems (ii)′.

THEOREM. *If for every L_1 satisfying $\underline{L}_1 \leqq L_1 \leqq \bar{L}_1$, there exists $S \in \mathcal{S}$ having $E(W_1 \mid \xi, S) = L_1$, then the class of all solutions of problems* (i)′ *with $\underline{L}_1 \leqq L_1 \leqq \bar{L}_1$ coincides with $\mathcal{S}$. Similarly for problems* (ii)′. *If $\bar{L}_1 = \infty$ or $\bar{L}_2 = \infty$ the appropriate equality signs must be omitted.*

This theorem is an immediate consequence of Lemma 1.

NOTE. From monotonicity (Lemma 3) we know that as $c \to c^0$ from one side and $S_c \in \mathcal{S}_c$, $E(W_1 \mid \xi, S_c) \to$ some limit L_1 from one side and $E(W_2 \mid \xi, S_c) \to$ some limit L_2 from one side. If this implies the existence of a procedure S having $E(W_1 \mid \xi, S) = L_1$ and $E(W_2 \mid \xi, S) = L_2$ whenever L_1 and L_2 are finite, it is easy to show that $S \in \mathcal{S}_{c^0}$, and that the conditions for the theorem are satisfied. However, the conditions themselves are usually easy to check once we have found $\mathcal{S}$.

Suppose that for a given Ω, ξ, W_1, W_2 we have found the class $\mathcal{S}$ of all solutions of problems (iii)′, $0 < c < \infty$, and find the conditions for the above theorem satisfied. Solving any problem (i)′ or (ii)′ is now reduced to choosing the appropriate member of $\mathcal{S}$.

3. Minimax problems. The following three classes of minimax problems have been considered: We want to find a (minimax) procedure which will

(i) subject to $\sup_\Omega E(W_1 \mid F, S) \leqq L_1$, minimise $\sup_\Omega E(W_2 \mid F, S)$,

(ii) subject to $\sup_\Omega E(W_2 \mid F, S) \leqq L_2$, minimise $\sup_\Omega E(W_1 \mid F, S)$,

(iii) minimise $\sup_\Omega \{cE(W_1 \mid F, S) + E(W_2 \mid F, S)\}$.

If there is an *a priori* distribution ξ which is least favorable in problem (iii)′ for all c, $0 < c < \infty$, this distribution is also least favorable for all problems (i)′ and (ii)′. The Bayes solutions with respect to this distribution are minimax solutions of the corresponding problems stated in this section. In many problems, however, this easy approach is not available.

LEMMA 5. *Suppose some problem* (iii) *has a solution S' with*

$$\sup_\Omega E(W_1 \mid F, S') = L_1, \qquad \sup_\Omega E(W_2 \mid F, S') = L_2,$$

$$\sup_\Omega \{cE(W_1 \mid F, S') + E(W_2 \mid F, S')\} = cL_1 + L_2.$$

(These conditions will in particular hold if either $\sup_\Omega E(W_1 \mid F, S') = L_1$ *and* $E(W_2 \mid F, S') \underset{\Omega}{\equiv} L_2$*, or* $\sup_\Omega E(W_2 \mid F, S') = L_2$ *and* $E(W_1 \mid F, S') \underset{\Omega}{\equiv} L_1$*.) Then* S' *is a solution of the problem* (i) *with this* L_1*, and a solution of the problem* (ii) *with this* L_2.

PROOF. Suppose there is a procedure S having

$$\sup_\Omega E(W_1 \mid F, S) \leqq L_1, \qquad \sup_\Omega E(W_2 \mid F, S) < L_2.$$

Then we would have

$$\sup_\Omega \{cE(W_1 \mid F, S) + E(W_2 \mid F, S)\} \leqq c \sup_\Omega E(W_1 \mid F, S)$$

$$+ \sup_\Omega E(W_2 \mid F, S) < cL_1 + L_2 = \sup_\Omega \{cE(W_1 \mid F, S') + E(W_2 \mid F, S')\},$$

contradicting the fact that S' is a solution of some problem (iii). Hence no such S can exist, and S' is a solution of the problem (i) with this L_1. Similarly S' is a solution of the problem (ii) with this L_2.

Let $\mathcal{C}$ be any class of solutions of problems (iii), each member S of which satisfies the condition

$$\sup_\Omega \{E(W_1 \mid F, S) + E(W_2 \mid F, S)\} = \sup_\Omega E(W_1 \mid F, S) + \sup_\Omega E(W_2 \mid F, S).$$

Let $\mathcal{C}_c$ denote those members of $\mathcal{C}$ which are solutions of the problem (iii) for this particular c. Then the following two lemmas can be proved in exactly the same way as the corresponding lemmas of Section 2.

LEMMA 2a. *If* $S' \in \mathcal{C}$, $S'' \in \mathcal{C}$, *then*

$$\sup_\Omega E(W_1 \mid F, S') < \sup_\Omega E(W_1 \mid F, S'') \longleftrightarrow \sup_\Omega E(W_2 \mid F, S') > \sup_\Omega E(W_2 \mid F, S''),$$

and

$$\sup_\Omega E(W_1 \mid F, S') = \sup_\Omega E(W_1 \mid F, S'') \longleftrightarrow \sup_\Omega E(W_2 \mid F, S') = \sup_\Omega E(W_2 \mid F, S'').$$

LEMMA 3a. *If* $S' \in \mathcal{C}_{c'}$ *and* $S'' \in \mathcal{C}_{c''}$, *where* $c' < c''$, *then*

$$\sup_\Omega E(W_1 \mid F, S') \geqq \sup_\Omega E(W_1 \mid F, S'')$$

and

$$\sup_\Omega E(W_2 \mid F, S') \leqq \sup_\Omega E(W_2 \mid F, S'').$$

Suppose that we have found such a class $\mathcal{C}$ of solutions of problems (iii) and that there exists $S \in \mathcal{C}$ having $\sup_\Omega E(W_i \mid F, S) = L_i$ whenever $\inf_{n,d,F} W_i(n, d, F) \leqq L_i \leqq \sup_{n,d,F} W_i(n, d, F)$, $i = 1, 2$. (Omit appropriate equality signs if either upper bound is ∞). Then solving any problem (i) or (ii) is reduced to choosing the appropriate member of $\mathcal{C}$.

In order to find solutions of problems (iii) in the examples we consider, the following lemma, which is due to E. Lehmann, will be needed.

LEMMA 6. *Consider the minimax problem of finding a procedure which minimises* $\sup_{\Omega} r(F, S)$. *(This may be subject to conditions as in* (i) *and* (ii), *or not as in* (iii).) *Let* S_k *be a Bayes procedure with respect to the a priori distribution* ξ_k *over* Ω, $k = 1, 2, \cdots$. *Then for any procedure* S,

$$\sup_{\Omega} r(F, S) \geqq r(\xi_k, S) \geqq r(\xi_k, S_k)$$

for all k. *Therefore*

$$\sup_{\Omega} r(F, S) \geqq \limsup_{k \to \infty} r(\xi_k, S_k).$$

A sufficient condition for the procedure S_0 *to be minimax is therefore*

$$r(F, S_0) \leqq \limsup_{k \to \infty} r(\xi_k, S_k)$$

for all $F \epsilon \Omega$.

4. Admissibility. Admissible procedures (not necessarily solutions) for the problems stated in Section 3 are defined as follows:

A procedure S is admissible for a particular problem (iii) if there is no procedure S^* having

$$r_c(F, S^*) \leqq r_c(F, S) \quad \text{for all} \quad F \epsilon \Omega,$$

with strict inequality for at least one $F \epsilon \Omega$, where $r_c(F, S) = cE(W_1 \mid F, S) + E(W_2 \mid F, S)$.

A procedure S is admissible for a particular problem (i) if there is no procedure S^* having

$$\sup_{\Omega} E(W_1 \mid F, S^*) \leqq L_1,$$

and

$$E(W_2 \mid F, S^*) \leqq E(W_2 \mid F, S) \quad \text{for all} \quad F \epsilon \Omega,$$

with strict inequality for at least one $F \epsilon \Omega$. Admissibility is defined in a similar way for problem (ii).

LEMMA 7. *Suppose* S *is an admissible procedure for some problem* (iii). *Then if* $E(W_1 \mid F, S) \underset{\Omega}{\equiv} L_1$, S *is admissible for the problem* (i) *with this* L_1. *And if* $E(W_2 \mid F, S) \underset{\Omega}{\equiv} L_2$, S *is admissible for the problem* (ii) *with this* L_2.

PROOF. Suppose that $E(W_1 \mid F, S) \underset{\Omega}{\equiv} L_1$ and that S is not admissible for the problem (i) with this L_1. Then there is a procedure S^* having $\sup_{\Omega}$

$E(W_1 \mid F, S^*) \leqq L_1$; and $E(W_2 \mid F, S^*) \leqq E(W_2 \mid F, S)$ for all $F \in \Omega$, with strict inequality for at least one $F \in \Omega$. We therefore have

$$r_c(F, S^*) = cE(W_1 \mid F, S^*) + E(W_2 \mid F, S^*)$$
$$\leqq cL_1 + E(W_2 \mid F, S) = cE(W_1 \mid F, S) + E(W_2 \mid F, S) = r_c(F, S)$$

for all $F \in \Omega$, with strict inequality for at least one $F \in \Omega$. That is, S cannot be admissible for any problem (iii), a contradiction which proves the first part of the lemma. The second part is proved in the same way.

If for a problem there is a least favorable distribution for which the Bayes solution is unique, this is the unique minimax solution and is therefore admissible. If Ω is a parametric family and all possible procedures have risks continuous in the parameter θ, and λ is a least favorable distribution which assigns positive probability to every interval of values of θ, then any Bayes solution for this λ is minimax and admissible. When can we conclude that minimax solutions obtained by the method of Lemma 6 are admissible? In Sections 5 and 7 we will show for particular examples that the solutions so obtained, except for trivial exceptions, are all admissible among the class of procedures with continuous risk functions. We might hope that all constant risk minimax solutions so obtained are admissible, but will see that this is not so.

The method used here for proving admissibility of minimax solutions involves examination of the Bayes solutions used to obtain them. In the examples considered, if W_2 is continuous, this method works both for classical fixed sample size problems and for the sequential problems (i), (ii), (iii) subject to the additional restriction that the number of observations is bounded.

Admissibility is proved for a number of examples by Hodges and Lehmann in [4] by a completely different method, which involves no appeal to Bayes solutions, and which works for certain fixed sample size problems in which the method of this paper fails. Their method, however, is restricted to number of observations and squared error of estimate for loss functions, and among sequential problems will handle only (i), again subject to the additional restriction that the number of observations is bounded.

5. Example: normal. Let $X_1, X_2, \cdots$ be a sequence of independent random variables, each being $N(\theta, 1)$, i.e., normal with mean θ and variance 1. A point estimate z is wanted for the mean θ. Let

$$W_1(n, z, \theta) = n, \qquad W_2(n, z, \theta) = W(z - \theta),$$

where W is a non-decreasing function of $| z - \theta |$. The three classes of minimax problems are

(i) subject to $\sup_\theta E_\theta(n) \leqq M$, minimise $\sup_\theta E_\theta W(z - \theta)$,

(ii) subject to $\sup_\theta E_\theta W(z - \theta) \leqq L$, minimise $\sup_\theta E_\theta(n)$,

(iii) minimise $\sup_\theta \{cE_\theta(n) + E_\theta W(z - \theta)\}$.

NOTE. This problem was first considered by Stein and Wald in [1]. They solved problems (i) and (ii) for the case $W(z - \theta) = 0$ or 1 according as $|z - \theta| \leqq a/2$ or $>a/2$; their estimates are thus confidence intervals of fixed length a. For this same case Wolfowitz in [2] solved problems (iii) and showed that these solutions provide solutions for problems (i) and (ii). Wolfowitz also obtained solutions of problems (iii) for the general $W(z - \theta)$, non-decreasing in $|z - \theta|$. The question of admissibility is not considered in [1] or [2].

The remainder of this section will be concerned with proving the following results.

THEOREM. *To a given c there corresponds either an integer N or a pair of consecutive integers N, N + 1. A class of solutions of the problem* (iii) *for this c are procedures in which the only possible sample sizes are N* (*or N, N* + 1) *and in which the estimate used is* $\frac{1}{n}\sum_{i=1}^{n} X_i$ *if* $n > 0$. *If* $N \neq 0$, *all such solutions are admissible among the class of procedures with continuous risk functions. The class of solutions so obtained,* $0 < c < \infty$, *provides solutions for all problems* (i) *and* (ii).

We will find solutions for problems (iii) by first finding Bayes solutions for the corresponding problems (iii)′ when θ has the *a priori* distribution $N(0, \sigma^2)$. The Bayes problem is to find a sequential estimation procedure which will minimise the risk

$$\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}\{cE_\theta(n) + E_\theta W(z - \theta)\}e^{-(1/2\sigma^2)\theta^2}\,d\theta.$$

We will assume that $W(z - \theta)$ increases with $|z - \theta|$ slowly enough so that

$$\int_{-\infty}^{\infty} E_\theta W(z - \theta)e^{-(1/2\sigma^2)(\theta-\mu)^2}\,d\theta < \infty$$

for some σ_0, μ_0, z_0, and hence for all $\sigma < \sigma_0$, μ, z.

Let us first determine what should be our estimate z for θ if we stop after having observed $x_1, \cdots, x_m$. For this we need to know the *a posteriori* distribution

$$\begin{aligned} p(\theta \mid x_1, \cdots, x_m) &= p(\theta, x_1, \cdots, x_m)/p(x_1, \cdots, x_m) \\ &= c_1(x_1, \cdots, x_m)e^{-(1/2\sigma^2)\theta^2}e^{-\frac{1}{2}\Sigma_1^m(x_i-\theta)^2} \\ &= c_2(x_1, \cdots, x_m)e^{-((m\sigma^2+1)/2\sigma^2)(\theta-(\sigma^2/(m\sigma^2+1))\Sigma_1^m x_i)^2} \end{aligned}$$

That is, θ, given $x_1, \cdots, x_m$, is $N\left(\frac{\sigma^2}{m\sigma^2+1}\sum_1^m x_i, \frac{\sigma^2}{m\sigma^2+1}\right)$. Given that we observe $x_1, \cdots, x_m$ and then stop and estimate $z(x_1, \cdots, x_m)$ for θ, our (*a posteriori*) risk is therefore

$$cm + \frac{\sqrt{m\sigma^2+1}}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} W(z - \theta)e^{-((m\sigma^2+1)/2\sigma^2)(\theta-(\sigma^2/(m\sigma^2+1))\Sigma_1^m x_i)^2}\,d\theta.$$

Since $W(z - \theta)$ is a non-decreasing function of $|z - \theta|$, this risk is clearly minimised by choosing $z = \frac{\sigma^2}{m\sigma^2 + 1} \sum_1^m x_i$, where we interpret $\sum_1^m x_i = 0$ if $m = 0$. The minimum value is

$$r_{c,\sigma}(m) = cm + \frac{\sqrt{m\sigma^2 + 1}}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{\infty} W(y) e^{-((m\sigma^2+1)/2\sigma^2)y^2}\, dy.$$

This does not depend on the observations, but only on the number of observations. Since $r_{c,\sigma} \to \infty$ as $m \to \infty$ it is clear that the sequence $r_{c,\sigma}(m): m = 0, 1, 2, \cdots$ assumes a minimum value at a finite set $n_1', \cdots, n_{p'}'[p' = p'(c, \sigma)]$ of integers m. Hence if θ is $N(0, \sigma^2)$ *a priori*, any of the following procedures is Bayes: The only possible sample sizes are $n_1', \cdots, n_{p'}'$; if the sample size is m, the estimate $z = \frac{\sigma^2}{m\sigma^2 + 1} \sum_1^m x_i$ is used for θ.

To obtain minimax procedures, consider a sequence of σ's tending to ∞. As $\sigma \to \infty$,

$$r_{c,\sigma}(m) \to r_c(m) = cm + \sqrt{\frac{m}{2\pi}} \int_{-\infty}^{\infty} W(y) e^{-(m/2)y^2}\, dy$$

for $m = 1, 2, \cdots$, and $r_{c,\sigma}(0) \to r_c(0) = \sup_y W(y)$.

Clearly $r_c(m): m = 0, 1, 2, \cdots$ assumes a minimum value at a finite set $n_1, \cdots, n_p[p = p(c)]$ of integers m.

Consider the following class $\mathfrak{C}_c'$ of sequential procedures: The only possible sample sizes are $n_1, \cdots, n_p$. If the sample size is 0, estimate 0 for θ (any estimate whatever will do as well). If the sample size is $m > 0$ estimate $z = \frac{1}{m} \sum_1^m x_i$ for θ. Writing $n_1 < n_2 < \cdots < n_p$, the risk of any such procedure, if $n_1 = 0$, is

$$\begin{aligned}
r_c^*(\theta) &= P(n = 0)W(\theta) + \sum_{i=2}^{p} \left\{ P_\theta(n = n_i) \left[cn_i + E_\theta W\left(\frac{1}{n_i} \sum_1^{n_i} x_j - \theta \right) \right] \right\} \\
&= P(n = 0)W(\theta) + \sum_{i=2}^{p} \left\{ P_\theta(n = n_i) \left[cn_i + \sqrt{\frac{n_i}{2\pi}} \int_{-\infty}^{\infty} W(y) e^{-(m/2)y^2}\, dy \right] \right\} \\
&\leqq P(n = 0) \sup_y W(y) + \sum_{i=2}^{p} P_\theta(n = n_i) r_c(n_i) \\
&= \sup_y W(y) = r_c(n_i), \qquad i = 2, \cdots, p, \quad \text{for all } \theta.
\end{aligned}$$

Similarly, if $n_1 \neq 0$, it is easy to show that

$$r_c^*(\theta) = r_c(n_i), \qquad i = 1, \cdots, p, \text{ for all } \theta.$$

It follows at once from Lemma 6 that every member of $\mathfrak{C}_c'$ is a minimax procedure for the problem (iii) with this c.

We will next show that

$$r_c(m) = cm + \sqrt{\frac{m}{2\pi}} \int_{-\infty}^{\infty} W(y) e^{-(m/2)y^2}\, dy \quad \text{for} \quad m > 0,$$
$$= \sup_y W(y) \quad \text{for} \quad m = 0$$

is a convex function of m. Let m_0 be the smallest integer for which $r_c(m) < \infty$; this is the same for all c. Then $r_c(m)$ is continuous in m for all $m \geqq m_0$. Convexity of $r_c(m)$ is equivalent to convexity of

$$g(m) = \sqrt{m} \int_0^{\infty} W(y) e^{-(m/2)y^2}\, dy.$$

It is easy to show that for $m_0 \leqq m < \infty$, differentiation under the integral sign any number of times is valid for $g(m)$. Therefore

$$g'(m) = \frac{1}{2\sqrt{m}} \int_0^{\infty} W(y) e^{-(m/2)y^2}(1 - my^2)\, dy,$$

$$g''(m) = \frac{1}{4m\sqrt{m}} \int_0^{\infty} W(y) e^{-(m/2)y^2}(m^2y^4 - 2my^2 - 1)\, dy$$
$$= \frac{1}{4m^2} \int_0^{\infty} W\left(\frac{x}{\sqrt{m}}\right) e^{-\frac{1}{2}x^2}(x^4 - 2x^2 - 1)\, dx.$$

Now

$$x^4 - 2x^2 - 1 < 0 \quad \text{for} \quad 0 \leqq x < \sqrt{1 + \sqrt{2}},$$
$$x^4 - 2x^2 - 1 > 0 \quad \text{for} \quad \sqrt{1 + \sqrt{2}} < x.$$

Also, $W(y)$ is non-decreasing as $y > 0$ increases and we will exclude from consideration the trivial case $W(y) \equiv$ constant. It follows that

$$g''(m) > \frac{1}{4m^2} \int_0^{\sqrt{1+\sqrt{2}}} W\left(\sqrt{\frac{1 + \sqrt{2}}{m}}\right) e^{-\frac{1}{2}x^2}(x^4 - 2x^2 - 1)\, dx$$
$$+ \frac{1}{4m^2} \int_{\sqrt{1+\sqrt{2}}}^{\infty} W\left(\sqrt{\frac{1 + \sqrt{2}}{m}}\right) e^{-\frac{1}{2}x^2}(x^4 - 2x^2 - 1)\, dx$$
$$= \frac{1}{4m^2} W\left(\sqrt{\frac{1 + \sqrt{2}}{m}}\right) \int_0^{\infty} e^{-\frac{1}{2}x^2}(x^4 - 2x^2 - 1)\, dx = 0.$$

That is, $g(m)$ is strictly convex for all $m \geqq m_0$. Hence $r_c(m)$ is continuous and strictly convex for $m \geqq m_0$.

For any given c, it follows that $r_c(m)\colon m = 0, 1, 2, \cdots$ is smallest for at most two consecutive integers m. If at two consecutive integers, these must be on opposite sides of the m which minimises $r_c(m)$. (Thus $p = 1$ or 2. The same results are now obvious for any $r_{c,\sigma}(m)$, given c, σ.)

For all c sufficiently large, $r_c(m)\colon m = 0, 1, 2, \cdots$ is minimised by $m = m_0$

only. Now, for any given m, $r_c(m)$ and $\partial r_c(m)/\partial m$ and $r_c(m+1) - r_c(m)$ are continuous and strictly increasing functions of c, $0 < c < \infty$. Therefore as we decrease c continuously from such a sufficiently large value, $r_c(m):m = 0, 1, 2, \cdots$ remains smallest for $m = m_0$ only, until a point c^1 is reached for which $r_{c^1}(m):m = 0, 1, 2, \cdots$ is minimised by $m = m_0$ and $m = m_0 + 1$. As we continue to decrease c, for ε sufficiently small and $c^1 - \varepsilon < c < c^1$, $r_c(m):m = 0, 1, 2, \cdots$ is clearly smallest for $m = m_0 + 1$ only. This remains true until we reach a point c^2 for which $r_{c^2}(m):m = 0, 1, 2, \cdots$ is minimised by $m = m_0 + 1$ and $m = m_0 + 2$. As we continue to decrease c, $r_c(m):m = 0, 1, 2, \cdots$ is smallest for larger and larger m's, which tend to ∞ as $c \to 0$, because, for a given m, $\partial r_c(m)/\partial m < 0$ for all c sufficiently small. We note that only for a denumerable set of c's is $r_c(m):m = 0, 1, 2, \cdots$ minimised by two consecutive m's; for almost all c's this minimum occurs for only one m.

Let $\mathcal{C}_c$ consist of those members of $\mathcal{C}_c'$ in which the sample size does not depend on θ. Included are the procedures in which the sample size is randomised, independently of the observations, between the possible sample sizes. Let $\mathcal{C} = \bigcup_{0<c<\infty} \mathcal{C}_c$. Now $E_\theta(n)$ is constant for any member of $\mathcal{C}$, implying $\sup_\theta \{E_\theta(n) + E_\theta(W)\} = \sup_\theta E_\theta(n) + \sup_\theta E_\theta(W)$. Lemmas 5, 2a and 3a are therefore valid for $\mathcal{C}$.

For every M, $m_0 \leqq M < \infty$ there is clearly a member of $\mathcal{C}$ having $E_\theta(n) \equiv M$. Using continuity considerations it is easy to show that for every L, $0 < L < \infty$, there is a member of $\mathcal{C}$ having $\sup_\theta E_\theta(W) = L$. It follows from Lemma 5 that $\mathcal{C}$ contains a solution for every problem (i) with $M \geqq m_0$ (problems (i) with $M < m_0$ have no solutions) and a solution for every problem (ii). Selection of the appropriate member of $\mathcal{C}$ is obvious for any particular problem (i), requires successive approximation for any particular problem (ii).

Are the members of $\mathcal{C}' = \bigcup_{0<c<\infty} \mathcal{C}_c'$ admissible for the problems (iii) for which they are solutions? We will answer this question first for those members of $\mathcal{C}'$ for which 0 is not a possible sample size.

For a given c, suppose that $r_c(m):m = 0, 1, 2, \cdots$ is minimised by $m = N \neq 0$ only, or by $m = N \neq 0$ and $m = N + 1$ only. Since, for every m, $r_{c,\sigma}(m) \to r_c(m)$ as $\sigma \to \infty$, it is clear that if θ has the distribution $\lambda_\sigma = N(0, \sigma^2)$ *a priori* with σ sufficiently large, say $\sigma > K_1$, then N and $N + 1$ are the only possible sample sizes for a Bayes solution. We observe further that

$$r_{c,\sigma}(N) = cN + \frac{1}{\sqrt{2\pi}} \sqrt{N + \frac{1}{\sigma^2}} \int_{-\infty}^{\infty} W(y) e^{-((N+1/\sigma^2)/2)y^2}\, dy,$$

$$r_{c,\sigma}(N+1) = c(N+1) + \frac{1}{\sqrt{2\pi}} \sqrt{N + \frac{1}{\sigma^2} + 1} \int_{-\infty}^{\infty} W(y) e^{-((N+1/\sigma^2+1)/2)y^2}\, dy.$$

If $r_c(N) \leqq r_c(N+1)$, as we are assuming, it follows from the convexity of $g(m) = \sqrt{m} \int_0^\infty W(y)\, e^{-(m/2)y^2}\, dy$ that $r_{c,\sigma}(N) < r_{c,\sigma}(N+1)$. Hence N is the only possible sample size for a Bayes procedure, $\sigma > K_1$. Therefore, for this given c the (minimax) risk function for every member of $\mathcal{C}_c'$ is

$$r(\theta) \equiv r = cN + \frac{1}{\sqrt{2\pi}} \sqrt{N} \int_{-\infty}^{\infty} W(y) e^{-(N/2)y^2}\, dy,$$

and the Bayes risk for *a priori* λ_σ, $\sigma > K_1$, is

$$r_\sigma = cN + \frac{1}{\sqrt{2\pi}} \sqrt{N + \frac{1}{\sigma^2}} \int_{-\infty}^{\infty} W(y) e^{-((N+1/\sigma^2)/2)y^2}\, dy.$$

If the procedures in $\mathcal{C}_c'$ are non-admissible for this problem (iii) there must exist a procedure S^* having risk function $r^*(\theta) \leqq r$ for all θ, with strict inequality for at least one θ. Assuming $r^*(\theta)$ continuous this implies strict inequality for some interval of values of θ. We can therefore find two constants a and k, $0 < a < r$ and $0 < k < \infty$ such that

$$\frac{1}{2k} \int_{-k}^{k} r^*(\theta)\, d\theta = a.$$

Also, given any fixed ε, $0 < \varepsilon < 1 - a/r$, we can find $K > K_1$ so large that for $-k \leqq \theta \leqq k$,

$$1 - \varepsilon < e^{-(1/2\sigma^2)\theta^2} < 1 \quad \text{whenever} \quad \sigma > K.$$

Then for all $\sigma > K$ we have

$$\begin{aligned}
\int_{-\infty}^{\infty} r^*(\theta)\lambda_\sigma(\theta)\, d\theta &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} r^*(\theta) e^{-(1/2\sigma^2)\theta^2}\, d\theta \\
&\leqq \frac{1}{\sqrt{2\pi}\sigma} \int_{-k}^{k} r^*(\theta) e^{-(1/2\sigma^2)\theta^2}\, d\theta + \frac{2}{\sqrt{2\pi}\sigma} \int_{k}^{\infty} r e^{-(1/2\sigma^2)\theta^2}\, d\theta \\
&= \frac{1}{\sqrt{2\pi}\sigma} \int_{-k}^{k} r^*(\theta) e^{-(1/2\sigma^2)\theta^2}\, d\theta + r - \frac{2r}{\sqrt{2\pi}\sigma} \int_{0}^{k} e^{-(1/2\sigma^2)\theta^2}\, d\theta \\
&\leqq \frac{1}{\sqrt{2\pi}\sigma} \int_{-k}^{k} r^*(\theta) \cdot 1\, d\theta + r - \frac{2r}{\sqrt{2\pi}\sigma} \int_{0}^{k} (1 - \varepsilon)\, d\theta \\
&= \frac{1}{\sqrt{2\pi}\,\sigma} 2ka + r - \frac{2r}{\sqrt{2\pi}\,\sigma} k(1 - \varepsilon) \\
&= r - \frac{b}{\sigma},
\end{aligned}$$

where $b = \dfrac{2k(r - a - r\varepsilon)}{\sqrt{2\pi}} > 0$ is a constant.

Now the Bayes risk for λ_σ, $\sigma > K$, is

$$r_\sigma = r - \frac{2}{\sqrt{2\pi}} \left\{ \sqrt{N} \int_{0}^{\infty} W(y) e^{-(N/2)y^2}\, dy - \sqrt{N + \frac{1}{\sigma^2}} \int_{0}^{\infty} W(y) e^{-((N+1/\sigma^2)/2)y^2}\, dy \right\}.$$

We have seen that for $m \geqq N$, the function $g(m) = \sqrt{m} \int_0^\infty W(y)\, e^{-(m/2)y^2}\, dy$ has continuous derivatives $g'(m) < 0$ and $g''(m) > 0$. It follows that

$$r_\sigma \geqq r + \frac{2}{\sqrt{2\pi}}\, g'(N)\, \frac{1}{\sigma^2},$$

$g'(N)$ being a negative constant. It is clear that for σ sufficiently large,

$$r_\sigma \geqq r + \frac{2}{\sqrt{2\pi}}\, g'(N)\, \frac{1}{\sigma^2} > r - \frac{b}{\sigma}$$
$$\geqq \int_{-\infty}^{\infty} r^*(\theta)\lambda_\sigma(\theta)\, d\theta.$$

But this contradicts the fact that r_σ is the Bayes risk for λ_σ, and so no such S^* can exist. Therefore, if 0 is not a possible sample size for members of $\mathcal{C}_c'$, every member of $\mathcal{C}_c'$ is admissible among the class of procedures with continuous risk functions, for the problem (iii) with this c.

Furthermore, $E_\theta(n)$ and $E_\theta(W)$ are both constants for members of $\mathcal{C}$ which belong to such a $\mathcal{C}_c'$. It follows from Lemma 7 that such members of $\mathcal{C}$ are admissible among the class of procedures with continuous risk functions, for the problems (i) and (ii) for which they are minimax.

If W is continuous and the number of observations is bounded, it can be shown that $r^*(\theta)$ is continuous. Thus if W is continuous, we have proved admissibility among the class of procedures with n bounded.

There remains the question of admissibility for those $\mathcal{C}_c'$ in which the possible sample sizes are 0 and 1, or 0 only. If 0 ane 1 are both possible sample sizes, two members of $\mathcal{C}_c'$ are A: take 0 observations and estimate 0 for θ; and B: take 1 observation and estimate x_1 for θ. Procedure A has risk function $r(\theta \mid A) = W(\theta)$. Procedure B has risk function $r(\theta \mid B) = c + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} W(y) e^{-\frac{1}{2}y^2}\, dy = \sup_y W(y)$. It easily follows that, except for A, all members of $\mathcal{C}_c'$ are non-admissible. The procedure A is admissible. For let S be any procedure which assigns probability $\alpha > 0$ to sample sizes > 0. Then we have

$$r(0 \mid S) \geqq \alpha c + W(0) > W(0) = r(0 \mid A),$$

so that no such S could make A non-admissible. Let T be any procedure which assigns probability 1 to the sample size 0. For any such procedure the risk $r(\theta \mid T)$ is an average, for some distribution of z, of $W(z - \theta)$. Let $(-\theta_0, \theta_0)$ be the interval or point on which $W(\theta) = W(0)$. Clearly we cannot have $r(\theta \mid T) = W(0)$ for all $\theta \in (-\theta_0, \theta_0)$ unless T coincides with A with probability 1. Hence no such T could make A non-admissible, and it now follows that A is admissible. This proof also shows that A is admissible when 0 is the only possible sample size for members of $\mathcal{C}_c'$.

Similar arguments show that every member of $\mathfrak{C}$ which belongs to a $\mathfrak{C}_c'$ of the above types, is admissible for the problems (i) and (ii) for which it is minimax.

6. Extensions of normal example. An outline of the solution of the example of section 5 shows that the same method can be used for other examples. Let X_1, X_2, $\cdots$ be independent random variables, each having the same density $p_\theta(x)$ with respect to a fixed measure μ. A point estimate z is wanted for the real parameter θ. Let

$$W_1(n, z, \theta) = W_1(n), W_2(n, z, \theta) = W_2(z, \theta)$$

and define the three classes of minimax problems as usual.

Suppose that we can find a sequence ξ_1, ξ_2, $\cdots$ of *a priori* distributions and a double sequence $z_{k,0}$, $z_{k,1}(x_1)$, $z_{k,2}(x_1, x_2)$, $\cdots$; $k = 1, 2, \cdots$ of estimates of θ, such that if θ has *a priori* distribution ξ_k and we observe x_1, $\cdots$, x_m and then stop, the *a posteriori* risk is minimised by estimating $z_{k,m}$ for θ, and the minimum value is

$$r_{c,k}(m) = cW_1(m) + \int_{-\infty}^{\infty} W_2(z_{k,m}; \theta)p(\theta \mid x_1, \cdots, x_m; \xi_k)\, d\theta,$$

depending not on the observations but only on the number m of observations (and c, k). Clearly the same sequences will do for all c, $0 < c < \infty$, and for all functions $W_1(n)$.

Then the following procedures are Bayes for the problem (iii)′ with this c, and with θ having *a priori* distribution ξ_k : The only possible sample sizes are those which minimise $r_{c,k}(m): m = 0, 1, 2, \cdots$; if the sample size is m estimate $z_{k,m}$ for θ.

Suppose for a particular ξ_k and for some particular c, that these possible sample sizes are $n_1 < n_2 < \cdots$. Since $r_{c,k}(m)$ is continuous in c for any k, m it is clear that for ε sufficiently small and $c < c' < c + \varepsilon$, no value of m other than n_1, n_2, $\cdots$ could minimise $r_{c',k}(m)$: $m = 0, 1, 2, \cdots$. And a minimum for any $m > n_1$ would provide a contradiction of Lemma 3. Hence for $c < c' < c + \varepsilon$, $r_{c',k}(m): m = 0, 1, 2, \cdots$ is minimised by $m = n_1$ only. It follows that randomisations in sample size for Bayes solutions are possible only for a denumerable set of c's; for almost all c, only one fixed sample size is possible.

Suppose that as $k \to \infty$ every sequence $z_{1,m}$, $z_{2,m}$, $\cdots$ tends to a limit z_m, and that $r_{c,k}(m) \to r_c(m) = cW_1(m) + L(m)$, for $m = 0, 1, 2, \cdots$. If the procedure: take a sample of fixed size m and estimate z_m for θ has risk function $r_c^*(\theta) = cW_1(m) + L_\theta(m) \leqq r_c(m)$ for all θ, the following procedures are minimax for the problem (iii) with this c: The only possible sample sizes are those which minimise $r_c(m): m = 0, 1, 2, \cdots$. If the sample size is m estimate z_m for θ.

Extension to problems (i) and (ii) can now be carried out as in section 5. We note that a problem of this type when solved for one $W_1(m)$ can be solved for any other $W_1(m)$ by merely determining the proper sample sizes. If $r_c(m)$ is a convex function of m, the possible sample sizes are always one integer or two consecutive integers. But if $r_c(m)$ is not convex, practically any set of integers can be possible sizes, as indicated in the following examples.

EXAMPLE. Let $X_1, X_2, \cdots$ be independent random variables, each being $N(\theta, 1)$. A point estimate z is wanted for the mean θ. Let

$$\begin{aligned} W_1(n) &= \frac{1}{3}n \qquad \text{for} \quad n = 0, 1, 2, 3, \\ &= 1 + \frac{n-3}{105} \quad \text{for} \quad n = 4, 5, \cdots, \\ W_2(z, \theta) &= (z - \theta)^2. \end{aligned}$$

Thus the first three observations each cost $\frac{1}{3}$, subsequent observations each cost $\frac{1}{105}$. Making the necessary substitutions in section 5, we get

$$\begin{aligned} r_c(m) &= c\frac{m}{3} + \frac{1}{m} \qquad \text{for} \quad m = 1, 2, 3, \\ &= c + \frac{c(m-3)}{105} + \frac{1}{m} \quad \text{for} \quad m = 4, 5, \cdots. \end{aligned}$$

For $c = 1$ it is easy to show that $r_1(m): m = 1, 2, \cdots$ is minimised by $m = 2$ and $m = 10$. For $c \neq 1$, $r_c(m): m = 1, 2, \cdots$ is minimised by one integer or by a pair of consecutive integers. Solutions are obtained for all problems (i), (ii), (iii) as in section 5. The solution obtained for any problem (i) with $\frac{2}{3} \leqq M \leqq \frac{16}{15}$ is the following:

$$\text{with probability } \frac{16 - 15M}{6} \text{ take 2 observations,}$$

$$\text{with probability } \frac{15M - 10}{6} \text{ take 10 observations,}$$

$$\text{estimate } z = \frac{1}{n}\sum_{i=1}^{n} x_i \quad \text{for} \quad \theta.$$

EXAMPLE. Let $X_1, X_2, \cdots$ be independent random variables each being $N(\theta, 1)$. A point estimate z is wanted for the mean θ. Let

$$\begin{aligned} W_1(n) &= 1 - \frac{1}{n} \quad \text{for} \quad n = 1, 2, \cdots, \\ &= 0 \qquad \text{for} \quad n = 0, \\ W_2(z, \theta) &= (z - \theta)^2. \end{aligned}$$

Making the necessary substitutions in Section 5,

$$r_c(m) = c + (1 - c)\frac{1}{m} \quad \text{for} \quad m = 1, 2, \cdots.$$

Clearly $r_1(m): m = 1, 2, \cdots$ is constant. Thus any procedure in which the sample size is at least 1 and the estimate $z = \frac{1}{n} \sum_{i=1}^{n} x_i$ is used for θ, is minimax for the problem (iii) with $c = 1$. If $c < 1$, problem (iii) has no solution. (The larger the sample size, the smaller is the risk.) If $c > 1$, $r_c(m): m = 1, 2, \cdots$ is minimised by $m = 1$ only. (In both these examples the possibility $n = 0$ is excluded because $\sup_\theta$ (risk) is then ∞.)

7. Example: rectangular. Let $X_1, X_2, \cdots$ be a sequence of independent random variables, each being $R(\theta - \frac{1}{2}, \theta + \frac{1}{2})$, i.e., rectangular with range $\theta - \frac{1}{2}$ to $\theta + \frac{1}{2}$. A point estimate z is wanted for the parameter θ. Let

$$W_1(n, z, \theta) = n, \qquad W_2(n, z, \theta) = W(z - \theta),$$

where W is a non-decreasing function of $| z - \theta |$. The three classes of minimax problems are

(i) subject to $\sup_\theta E_\theta(n) \leqq M$, minimise $\sup_\theta E_\theta W(z - \theta)$,

(ii) subject to $\sup_\theta E_\theta W(z - \theta) \leqq L$, minimise $\sup_\theta E_\theta(n)$,

(iii) minimise $\sup_\theta \{cE_\theta(n) + E_\theta W(z - \theta)\}$.

NOTE. The problems (iii) are solved by Wald in [3] for the case $W(z - \theta) = (z - \theta)^2$. We will show that Wald's solution holds for any $W(z - \theta)$ which is non-decreasing in $| z - \theta |$, and will obtain solutions of (i) and (ii). In addition, admissibility results will be proved as in Section 5.

The remainder of this section will be concerned with proving the following results.

THEOREM. *The following procedures are admissible solutions of problem* (iii) *among the class of all procedures with continuous risk functions. If* $\phi^* = \sup_\alpha W(\alpha) - 2 \int_0^{\frac{1}{2}} W(\alpha)\, d\alpha - c < 0$ *take* 0 *observations and estimate* 0 *for* θ. *If* $\phi^* > 0$ *take at least one observation and after the* m^{th} *observation* ($m = 1, 2, \cdots$) *compute the range* r_m *of* $x_1, \cdots, x_m$. *If* $r_m > 1 - \bar{l}$ *stop and estimate the mid-range for* θ; *if* $r_m < 1 - \bar{l}$ *take another observation*; *if* $r_m = 1 - \bar{l}$ *do either. If* $\phi^* = 0$ *follow either procedure.* (*Here* $\bar{l}$, *to be defined later, is a constant depending on c and W.*) *The class of procedures so obtained,* $0 < c < \infty$, *provides admissible solutions among the class of procedures with continuous risk functions, for all problems* (i) *and* (ii).

Solutions are found for problems (iii) by first finding Bayes solutions for the corresponding problems (iii)′ when θ has *a priori* distribution $R(a, b)$. The Bayes problem is to find a sequential estimation procedure which minimises the risk

$$E\{r_c(\theta \mid S) \mid \theta \sim R(a, b)\} = \frac{1}{b - a} \int_a^b \{cE_\theta(n \mid S) + E_\theta(W \mid S)\}\, d\theta.$$

Let us first determine what should be our estimate z if we stop after having observed $x_1, \cdots, x_m$. For this we will need to know the *a posteriori* distribution $p(\theta \mid x_1, \cdots, x_m)$. Writing $u_m = \min(x_1, \cdots, x_m)$ and $v_m = \max(x_1, \cdots, x_m)$, this distribution is easily found to be $R(u'_m, v'_m)$, where $(u'_m, v'_m) = (a, b) \cap (v_m - \frac{1}{2}, u_m + \frac{1}{2})$ for $m = 1, 2, \cdots$, and $(u'_0, v'_0) = (a, b)$. Clearly a best estimate, i.e., one minimising the *a posteriori* risk $cm + \int W(z - \theta)p(\theta \mid x_1, \cdots, x_m)\, d\theta$ is $z = \dfrac{u'_m + v'_m}{2}$, the mid-point of (u'_m, v'_m). The minimum value is

$$r_m = cm + \frac{1}{v'_m - u'_m} \int_{u'_m}^{v'_m} W\left(\theta - \frac{u'_m + v'_m}{2}\right) d\theta$$

$$= cm + \frac{2}{t_m} \int_0^{\frac{1}{2}t_m} W(\alpha)\, d\alpha,$$

where $t_m = v'_m - u'_m$ for $m = 0, 1, 2, \cdots$.

To determine an optimum stopping rule we will need to know, for all $t > 0$, the conditional expected value of r_{m+1} given $t_m = t$. Now

$$p(x_{m+1} \mid t_m = t) = \frac{1}{t}\left(\text{length of}\left(\frac{u'_m + v'_m}{2} - \frac{t}{2}, \frac{u'_m + v'_m}{2} + \frac{t}{2}\right) \cap (x_{m+1} - \tfrac{1}{2}, x_{m+1} + \tfrac{1}{2})\right).$$

From this it is easy to show that

$$E(r_{m+1} \mid t_m = t) = c(m + 1) + \frac{2(t - 1)}{t} \int_0^{1/2} W(\alpha)\, d\alpha + \frac{4}{t} \int_0^1 \left[\int_0^{x/2} W(\alpha)\, d\alpha\right] dx$$

for $t \geqq 1$; and that for $t \leqq 1$,

$$E(r_{m+1} \mid t_m = t) = c(m + 1) + \frac{2(1 - t)}{t} \int_0^{t/2} W(\alpha)\, d\alpha + \frac{4}{t} \int_0^t \left[\int_0^{x/2} W(\alpha)\, d\alpha\right] dx.$$

Let

$$\phi(t) = cm + \frac{2}{t} \int_0^{t/2} W(\alpha)\, d\alpha - E(r_{m+1} \mid t_m = t),$$

the expected decrease in *a posteriori* risk due to taking $m + 1$ instead of m observations when $t_m = t$. We have

$$\theta(t) = \frac{2}{t} \int_0^{t/2} W(\alpha)\, d\alpha + \left(\frac{2}{t} - 2\right) \int_0^{1/2} W(\alpha)\, d\alpha - \frac{4}{t} \int_0^1 \left[\int_0^{x/2} W(\alpha)\, d\alpha\right] dx - c$$

for $t \geqq 1$; and for $t \leqq 1$,

$$\phi(t) = 2\int_0^{t/2} W(\alpha)\,d\alpha - \frac{4}{t}\int_0^t \left[\int_0^{x/2} W(\alpha)\,d\alpha\right] dx - c.$$

Now $W(\alpha)$, being non-decreasing for $\alpha \geqq 0$, has at most a denumerable set of discontinuities. If $W(\alpha)$ is continuous at $\alpha = t/2$ we have, for $t > 1$:

$$\phi'(t) = \frac{1}{t} W\left(\frac{t}{2}\right) - \frac{2}{t^2}\int_0^{t/2} W(\alpha)\,d\alpha - \frac{2}{t^2}\int_0^{1/2} W(\alpha)\,d\alpha$$
$$+ \frac{4}{t^2}\int_0^1 \left[\int_0^{x/2} W(\alpha)\,d\alpha\right] dx$$
$$= \frac{1}{t} W\left(\frac{t}{2}\right) - \frac{2}{t^2}\int_0^1 \left[\int_0^{1/2} W(\alpha)\,d\alpha\right] dx - \frac{2}{t^2}\int_{1/2}^{t/2} W(\alpha)\,d\alpha$$
$$- \frac{2}{t^2}\int_0^1 \left[\int_0^{1/2} W(\alpha)\,d\alpha\right] dx + \frac{4}{t^2}\int_0^1 \left[\int_0^{x/2} W(\alpha)\,d\alpha\right] dx$$
$$= \frac{1}{t} W\left(\frac{t}{2}\right) - \frac{2}{t^2}\int_{1/2}^{t/2} W(\alpha)\,d\alpha - \frac{4}{t^2}\int_0^1 \left[\int_{x/2}^{1/2} W(\alpha)\,d\alpha\right] dx$$
$$\geqq \frac{1}{t} W\left(\frac{t}{2}\right) - \frac{2}{t^2}\,\frac{t-1}{2}\, W\left(\frac{t}{2}\right) - \frac{4}{t^2} W\left(\tfrac{1}{2}\right)\cdot\tfrac{1}{2}$$
$$= \frac{1}{t^2} W\left(\frac{t}{2}\right) - \frac{1}{t^2} W\left(\tfrac{1}{2}\right) \geqq 0;$$

and if $t < 1$ we have

$$\phi'(t) = W\left(\frac{t}{2}\right) - \frac{4}{t}\int_0^{t/2} W(\alpha)\,d\alpha + \frac{4}{t^2}\int_0^t \left[\int_0^{x/2} W(\alpha)\,d\alpha\right] dx$$
$$= W\left(\frac{t}{2}\right) - \frac{4}{t^2}\int_0^t \left[\int_0^{t/2} W(\alpha)\,d\alpha\right] dx + \frac{4}{t^2}\int_0^t \left[\int_0^{x/2} W(\alpha)\,d\alpha\right] dx$$
$$= W\left(\frac{t}{2}\right) - \frac{4}{t^2}\int_0^t \left[\int_{x/2}^{t/2} W(\alpha)\,d\alpha\right] dx$$
$$\geqq W\left(\frac{t}{2}\right) - \frac{4}{t^2} W\left(\frac{t}{2}\right)\frac{t^2}{4} = 0.$$

If $t/2$ is a discontinuity point of $W(\alpha)$, the same inequalities hold for the one-sided derivatives of $\phi(t)$, both of which exist. We observe that these inequalities are strict unless $W(\alpha)$ is constant on the open interval $(0, t/2)$. Noting that $\phi(t) \to -c$ as $t \to 0$, we have proved that $\phi(t)$ is continuous and non-decreasing for $t > 0$, being strictly increasing whenever $\phi(t) > -c$.

Hence $\phi(t) < 0$ for all t, or else $\phi(t) = 0$ has a unique root $\bar{t}$. Using also the fact that $t_{m+1} \leqq t_m$, we now obtain, by the methods of [3], the following results.

If $\phi(t) < 1$ for all t, a Bayes solution is: Take 0 observations, estimate $\dfrac{a+b}{2}$

for θ. If $\phi(t) > 0$ for some t, a Bayes solution is: After the mth observation ($m = 0, 1, 2 \cdots$) compute $t_m = v_m' - u_m'$. If $t_m < \bar{t}$ stop and estimate $\frac{u_m' + v_m'}{2}$ for θ; if $t_m > \bar{t}$ take another observation; if $t_m = \bar{t}$ do either.

Consider now the following procedures S_0: If $\phi^* = \sup_\alpha W(\alpha) - 2\int_0^{1/2} W(\alpha)\, d\alpha - c < 0$ take 0 observations and estimate 0 for θ. If $\phi^* > 0$ take at least one observation, and after each observation ($m = 1, 2, \cdots$) compute $t_m^* = u_m + \frac{1}{2} - (v_m - \frac{1}{2}) = u_m - v_m + 1$; if $t_m^* < \bar{t}$ stop and estimate $\frac{u_m + v_m}{2}$ for θ, if $t_m^* > \bar{t}$ take another observation, and if $t_m^* = \bar{t}$ do either. Finally, if $\phi^* = 0$ use either of these two procedures.

If $\phi^* > 0$ it is easy to show that $E_\theta(n \mid S_0)$, $E_\theta(W \mid S_0)$ and

$$r(\theta \mid S_0) = cE_\theta(n \mid S_0) + E_\theta(W \mid S_0) \equiv r$$

are all constants. Also, for any particular c, there is always an S_0 for which $E_\theta(n \mid S_0)$ is constant.

Let S_k be a Bayes procedure when θ has the distribution $\xi_k = R(-k, k)$ *a priori*. If $\phi^* \leqq 0$, then for all k the procedure S_k is: take 0 observations and estimate 0 for θ; it thus coincides with an S_0. (Other possible S_0 have the same $\sup_\theta r(\theta \mid S_0)$.) If $\phi^* > 0$, then for all k sufficiently large the procedure S_k coincides with S_0 for $-(k - 1) \leqq \theta \leqq k - 1$. Taking a sequence of S_k's with $k \to \infty$, it easily follows from Lemma 6 that all procedures S_0 are minimax for the problem (iii) in question.

By the same methods as are used in section 5 it is easy to show that the procedures S_0 obtained above provide solutions for all solvable problems (i) and (ii).

In the case $\phi^* > 0$, for the procedure S_0 to be non-admissible for the problem (iii) for which it is minimax, there must exist a procedure S_0^* having risk function

$$r(\theta \mid S_0^*) \leqq r \quad \text{for all} \quad \theta$$

with strict inequality for at least one θ and so, if $r(\theta \mid S_0^*)$ is continuous, for an interval of values of θ. We can therefore find $h > \frac{1}{2}$ such that

$$\frac{1}{2h - 1} \int_{-h+1/2}^{h-1/2} r(\theta \mid S_0^*)\, d\theta = a < r.$$

Now for $\alpha = \pm 2, \pm 4, \cdots$ define the procedure S_α^* as follows. If $x_1, x_2, \cdots$ are observed, use the decision procedure S_0^* for the sequence $x_1 - \alpha h$, $x_2 - \alpha h, \cdots$ and add αh to the resulting estimate. We clearly have

$$r(\theta \mid S_\alpha^*) = r(\theta - \alpha h \mid S_0^*).$$

Now define the procedure S^* as follows. Take at least one observation. If $x_1 \epsilon (\overline{\alpha - 1}h, \overline{\alpha + 1}h]$, $\alpha = 0, \pm 2, \pm 4, \cdots$, use the procedure S_α^*. If $\theta \epsilon (\overline{\alpha - 1}h + \frac{1}{2}, \overline{\alpha + 1}h - \frac{1}{2})$, then $x_1 \epsilon (\overline{\alpha - 1}h, \overline{\alpha + 1}h]$ and so the procedure S^* reduces to S_α^*. Hence $r(\theta \mid S^*)$ coincides with $r(\theta \mid S_\alpha^*)$ for

$$\theta \epsilon (\overline{\alpha - 1}h + \tfrac{1}{2}, \overline{\alpha + 1}h - \tfrac{1}{2}), \qquad \alpha = 0, \pm 2, \pm 4, \cdots .$$

And $r(\theta \mid S^*) \leqq r$ for all θ. Therefore

$$\frac{1}{2(2k+1)h} \int_{-(2k+1)h}^{(2k+1)h} r(\theta \mid S^*) \, d\theta \leqq \frac{(2h-1)a + r}{2h} = r - \frac{(r-a)(2h-1)}{2h}.$$

But if θ has the distribution $R(-\overline{2k+1}h, \overline{2k+1}h)$ *a priori*, the Bayes solution coincides with S_0 for $\theta \epsilon (-\overline{2k+1}h + 1, \overline{2k+1}h - 1)$. We therefore have for this *a priori* distribution

$$\text{Bayes risk} \geqq \frac{2(2k+1)h - 2}{2(2k+1)h} r = r - \frac{r}{(2k+1)h}.$$

For k sufficiently large this clearly exceeds $r - (2h-1)(r-a)/2h$, contradicting the above inequality on the Bayes risk. It follows that no such S_0^* as assumed can exist and therefore that the procedure S_0 is admissible, among the class of procedures with continuous risk functions, for the problems (iii) for which it is minimax and also, by Lemma 7, for the problems (i) and (ii) for which it is minimax.

If W is continuous and the number of observations is bounded it can be shown that $r^*(\theta)$ is continuous. Thus if W is continuous, S_0 is admissible, among the class of procedures with n bounded, for the three problems.

It remains to consider admissibility for procedures S_0 where $\phi^* \leqq 0$. Proofs for these cases can be given in the same way as for the corresponding cases in Section 5.

The solution for this example still works if we replace $W_1(n) = n$ by some other $W_1(n)$, but only so long as the resulting function $\phi(t)$ is non-decreasing.

NOTE. In the above examples, a procedure is called cogredient if for every c the same number of observations is taken for $x_1 + c, x_2 + c, \cdots$ as for $x_1, x_2, \cdots$ and $z(x_1 + c, \cdots, x_n + c) = z(x_1, \cdots, x_n) + c$. Such procedures have constant risk functions; so it follows that all the constant risk procedures obtained in Sections 5, 6, 7 have uniformly minimum risk among all cogredient procedures for the problems for which they are minimax.

The author wishes to express his appreciation to Professor E. L. Lehmann for valuable advice and assistance in the preparation of this paper, and also to Jack C. Kiefer, who read the manuscript and pointed out an important correction.

REFERENCES

[1] C. STEIN AND A. WALD, "Sequential confidence intervals for the mean of a normal distribution with known variance," *Annals of Math. Stat.*, Vol. 18 (1947), pp. 427–433.

[2] J. WOLFOWITZ, "Minimax estimates of the mean of a normal distribution with known variance," *Annals of Math. Stat.*, Vol. 21 (1950), pp. 218–230.

[3] A. WALD, *Statistical Decision Functions*, John Wiley and Sons, 1950.

[4] J. HODGES AND E. LEHMANN, "Extensions and applications of the Cramér-Rao inequality," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, 1951.