The Annals of Statistics
1973, Vol. 1, No. 6, 1149–1157

MINIMAX ESTIMATION OF A CUMULATIVE DISTRIBUTION FUNCTION[1]

BY E. G. PHADIA

The Ohio State University

In this paper the minimax estimators of a cumulative distribution function F is obtained for four types of loss functions. The result is quite general in that no restrictions are imposed on the unknown F. Moreover, the estimates do not depend upon the weight function used in the definition of the loss functions. It is also shown that the sample distribution function is minimax under one of these types of loss functions.

1. Introduction and summary. Suppose we are given a random sample $X_1, \cdots, X_n$ from an unknown cumulative distribution function F and the problem is to estimate the function F. One of the most frequently used estimators is the sample cumulative distribution function. This estimator has some nice properties, and, in addition, Dvoretzky, Kiefer, and Wolfowitz (1956) have shown that the sample distribution function is asymptotically minimax for a very wide class of loss functions. For some classes of loss functions Read (1972) has shown that it is asymptotically inadmissible. Aggarwal (1955) considered this problem of estimation using a decision theoretic approach but with the restriction that the unknown F be continuous. Considering a class of invariant loss functions he obtained best invariant estimators which are essentially step functions. Later, Taha (1968) obtained invariant estimators for a different class of invariant loss functions, but still under the assumption that F be continuous. The minimax estimators of F are obtained here with no assumptions on the unknown cumulative distribution function F.

To obtain the minimax estimators we use a well-known result: If a sequence of Bayes risks of Bayes estimators of F with respect to a sequence of prior distributions converges to a constant which is not smaller than the risk function of an estimator then this latter estimator is minimax. The main difficulty arises in constructing a suitable sequence of priors on the space of all distribution functions defined on the real line. Several attempts have been made in this direction. Notable among them are those of Dubins and Freedman (1963, 1967), Kraft and van Eeden (1964), Kraft (1964), and Ferguson (1973). Without going into the merits or demerits of these methods we use a method which is similar to that of Ferguson (1973).

Received December 1971; revised April 1973.

[1] This research is part of a doctoral dissertation at The Ohio State University; research is partially supported by the Air Force Office of the Scientific Research, United States Air Force, under AFOSR Grant No. AF-AFOSR-1305-67.

AMS 1970 *subject classifications.* Primary 62; Secondary C99 or G05.

Key words and phrases. Minimax, prior distribution, posterior distribution, Bayes estimators, Bayes risk, distribution function.

In this paper, we derive minimax estimators for four classes of loss functions. For one of these classes the minimax estimator is the sample distribution function. However, for the other three classes they are improper distribution functions. In the sequel we also give necessary and sufficient conditions for the risk function to be independent of F and also evaluate the minimum Bayes risk under a general class of loss functions.

2. Notation. We consider the parameter space Θ, the action space Ω, and the loss function L defined as follows.

$\Theta = \{F : F$ is a right continuous distribution function on the real line $R^1\}$.
$\Omega = \{\phi : \phi$ is a non-decreasing right continuous function on R^1 such that $0 \leqq \phi(-\infty),\ \phi(\infty) \leqq 1\}$.

$$L = L(F, \phi) = \int [F(t) - \phi(t)]^2 [F(t)]^{\gamma-1} [1 - F(t)]^{\delta-1}\, dW(t)\,, \qquad \gamma, \delta \geqq 0,\ F \in \Theta,\ \phi \in \Omega\,, \tag{2.1}$$

where W is a given non-null, finite measure on (R^1, B), where B is the Borel field on R^1. In the rest of the paper we shall refer to W as a weight function. We denote the risk function of an estimator ϕ, by $R(F, \phi)$ and the corresponding Bayes risk with respect to a prior distribution τ by $r(\tau, \phi)$.

By L_1, L_2, L_3 and L_4 we shall denote four special cases of the loss function L, obtained by substituting $\gamma = 0, 1$ and $\delta = 0, 1$ in (2.1).

$$\begin{aligned} L_1 &= L(F, \phi) = \int [F(t) - \phi(t)]^2\, dW(t) && (\text{for } \gamma = \delta = 1) \\ L_2 &= L(F, \phi) = \int \{[F(t) - \phi(t)]^2 / F(t)[1 - F(t)]\}\, dW(t) && (\text{for } \gamma = \delta = 0) \\ L_3 &= L(F, \phi) = \int \{[F(t) - \phi(t)]^2 / F(t)\}\, dW(t) && (\text{for } \gamma = 0,\ \delta = 1) \\ L_4 &= L(F, \phi) = \int \{[F(t) - \phi(t)]^2 / [1 - F(t)]\}\, dW(t) && (\text{for } \gamma = 1,\ \delta = 0)\,. \end{aligned} \tag{2.2}$$

3. Characterization of estimators ϕ which make $R(F, \phi)$ independent of F. Many of the commonly used estimators of a distribution function are step functions with jumps at the observations. They can be written in the form

$$\hat{F}(t) = a + \sum_{i=1}^n b_i \delta_{X_i}((-\infty, t]) \tag{3.1}$$

where $a, b_1, \cdots, b_n$ are nonnegative constants $\delta_{X_i}(A) = 1$ if $X_i \in A$, zero otherwise. This observation leads us to state and prove the following result for the class of estimators which are step functions of this type. This result is used in the derivation of the main result of Section 5.

THEOREM 3.1. *Let $X_1, \cdots, X_n$ be a random sample from an unknown cumulative distribution function F. Let*

$$\Phi = \{\phi : \phi(t) = a + \textstyle\sum_{i=1}^n b_i \delta_{X_i}((-\infty, t]),\ a \geqq 0,\ b_i \geqq 0 \text{ and } a + \sum b_i \leqq 1\} \tag{3.2}$$

be a class of estimators of F. Then under the loss functions L_i of (2.2), the risk function $R(F, \phi)$, $\phi \in \Phi$ is independent of F for all $F \in \Theta$ if and only if $\phi \in \Phi_i$, where

$$\Phi_i = \{\phi : \phi \in \Phi \text{ and condition } C_i \text{ on } a, b_1, \cdots, b_n \text{ is satisfied}\}, \tag{3.3}$$

$i = 1, 2, 3$ and 4, and

$$\begin{array}{llll} C_1: & 2a = 1 - \sum b_i & \text{and} & (1 - \sum b_i)^2 = \sum b_i^2 \\ C_2: & a = 0 & \text{and} & \sum b_i = 1 \\ C_3: & a = 0 & \text{and} & (1 - \sum b_i)^2 = \sum b_i^2 \\ C_4: & a = 1 - \sum b_i & \text{and} & (1 - \sum b_i)^2 = \sum b_i^2 . \end{array} \tag{3.4}$$

The summation is over $i = 1, 2, \cdots, n$.

REMARK. It can easily be seen that for the values of γ and δ other than zero and one, the risk function cannot be made independent of F. A formal proof can be provided by evaluating the risk function, expanding the factor $[1-F(t)]^{\delta-1}$, collecting the coefficients of the powers of $F(t)$, equating them to zero and solving.

PROOF. We shall prove this theorem only for the cases L_1 and L_2. The other two cases can be proven similarly.

Case of L_1: Now for this loss function and $\phi \in \Phi$,

$$\begin{aligned} R(F, \phi) &= E_F \int [F(t) - \phi(t)]^2 \, dW(t) \\ &= \int E_F[F(t) - a - \sum b_i \delta_{X_i}((-\infty, t])]^2 \, dW(t) \\ &= \int \{[(1 - \sum b_i)^2 - \sum b_i^2]F^2(t) \\ &\qquad + [-2a(1 - \sum b_i) + \sum b_i^2]F(t) + a^2\} \, dW(t) \end{aligned} \tag{3.5}$$

by expanding and using the fact that

$$E_F[\delta_{X_i}((-\infty, t])] = F(t) \qquad \text{for each} \quad i = 1, 2, \cdots, n . \tag{3.6}$$

Here the symbol E_F denotes the expectation taken with respect to the distribution function F.

If $\phi \in \Phi_1$, the coefficient of $F^2(t)$ and $F(t)$ inside the integral sign in (3.5) vanish. The risk function is

$$R(F, \phi) = \int a^2 \, dW(t) \qquad \text{for} \quad \phi \in \Phi_1 , \tag{3.7}$$

which is independent of F.

Also, in order that $R(F, \phi)$ be independent of F for all $F \in \Theta$ and $\phi \in \Phi$, equating the coefficients of $\int F^2(t) \, dW(t)$ and $\int F(t) \, dW(t)$ in (3.5) to zero, we obtain the following conditions on $a, b_1, \cdots, b_n$

$$(1 - \sum b_i)^2 = \sum b_i^2 \qquad \text{and} \qquad 2a = 1 - \sum b_i . \tag{3.8}$$

These conditions imply that $\phi \in \Phi_1$.

Case of L_2: When the loss function is L_2 (2.2) and $\phi \in \Phi$,

$$\begin{aligned} R(F, \phi) &= \int E_F\{[F(t) - a \\ &\quad - \textstyle\sum b_i \delta_{X_i}((-\infty, t])]^2/[F(t)(1 - F(t))]\} \, dW(t) \\ &= \int \{[(1 - \textstyle\sum b_i)^2 - \sum b_i^2]F(t)/[1 - F(t)] \\ &\quad + [-2a(1 - \textstyle\sum b_i) + \sum b_i^2]/[1 - F(t)] \\ &\quad + a^2/[F(t)(1 - F(t))]\} \, dW(t) \,. \end{aligned} \tag{3.9}$$

If $\phi \in \Phi_2$, then $a = 0$, $\sum b_i = 1$, and (3.9) reduces to

$$R(F, \phi) = \int (\textstyle\sum b_i^2) \, dW(t) \tag{3.10}$$

which is independent of F.

To prove the converse, we note that in order that $R(F, \phi)$ be independent of F for all $F \in \Theta$ and $\phi \in \Phi$, a must be zero, in which case (3.9) reduces to

$$R(F, \phi) = \int [\{F(t)[(1 - \textstyle\sum b_i)^2 - \sum b_i^2] + \sum b_i^2\}/[1 - F(t)]] \, dW(t) \,. \tag{3.11}$$

In order that this be independent of F for all $F \in \Theta$, we should have

$$-[(1 - \textstyle\sum b_i)^2 - \sum b_i^2] = \sum b_i^2 \tag{3.12}$$

or

$$\textstyle\sum b_i = 1 \,. \tag{3.13}$$

This implies that $\phi \in \Phi_2$. This completes the proof of the theorem.

If we let all b_i's be equal to, say, b, then the condition C_1 in (3.4) reduces to $2a = 1 - nb$ and $(1 - nb)^2 = nb^2$. This in turn yields $b = (n + n^{\frac{1}{2}})^{-1}$ or $(n - n^{\frac{1}{2}})^{-1}$ and correspondingly $a = [2(n^{\frac{1}{2}} + 1)]^{-1}$ and $-[2(n^{\frac{1}{2}} + 1)]^{-1}$. But the second value of a is not permissible, and hence for the loss function L_1 we get only one estimator for F,

$$\phi_1(t) = [2(n^{\frac{1}{2}} + 1)]^{-1} + [n^{\frac{1}{2}}(n^{\frac{1}{2}} + 1)]^{-1} \cdot \textstyle\sum \delta_{X_i}((-\infty, t]) \tag{3.14}$$

which makes the risk function $R(F, \phi)$ independent of F. Similarly, for the loss function L_2, the condition C_2 reduces to $a = 0$, $b_i = b = n^{-1}$ for $i = 1, 2, \cdots, n$, and the resulting estimator

$$\phi_2(t) = n^{-1} \cdot \textstyle\sum_i \delta_{X_i}((-\infty, t]) \tag{3.15}$$

is the only one which makes $R(F, \phi)$ independent of F. Thus we have the following corollary.

COROLLARY 3.1. *Let $X_1, \cdots, X_n$ be a random sample from an unknown cumulative distribution function F. Let*

$$\Phi^* = \{\phi : \phi(t) = a + b \textstyle\sum \delta_{X_i}((-\infty, t]), a \geqq 0, b \geqq 0 \text{ and } a + bn \leqq 1\} \tag{3.16}$$

be a class of estimators of F. Then under the loss function L_i, the risk function $R(F, \phi)$, $\phi \in \Phi^*$, is independent of F for all $F \in \Theta$ if and only if $\phi = \phi_i$ for $i = 1$, 2, 3 *and* 4 *where* ϕ_1 *and* ϕ_2 *are defined in* (3.14) *and* (3.15) *respectively, and*

$$\phi_3(t) = (n^{\frac{1}{2}} + n)^{-1} \sum \delta_{X_i}((-\infty, t]) \tag{3.17}$$
$$\phi_4(t) = (n^{\frac{1}{2}} + 1)^{-1} + (n^{\frac{1}{2}} + n)^{-1} \sum \delta_{X_i}((-\infty, t]) .$$

It is easy to see that the risks of ϕ_1, ϕ_2, ϕ_3, ϕ_4 under their respective loss functions L_1, L_2, L_3 and L_4, are R_1, R_2, R_3 and R_4 respectively, where

$$\begin{aligned} R_1 &= R_1(\phi_1) = \{1/4(n^{\frac{1}{2}} + 1)^2\} \int dW(t) \\ R_2 &= R_2(\phi_2) = \{1/n\} \int dW(t) \\ R_3 &= R_3(\phi_3) = \{1/(n^{\frac{1}{2}} + 1)^2\} \int dW(t) \\ R_4 &= R_4(\phi_4) = \{1/(n^{\frac{1}{2}} + 1)^2\} \int dW(t) . \end{aligned} \tag{3.18}$$

4. Minimum Bayes risk. In this section we define a sequence of prior distributions for F on Θ and evaluate the minimum Bayes risk for this sequence.

The sequence of prior distributions $\{\tau_k\}$ may be described in the following way.[2] The kth member of the sequence, τ_k, chooses a distribution function at random as follows. First choose p at random from the beta distribution with parameters $\alpha, \beta > 0$ (to be denoted by Be (α, β)) and then let

$$\begin{aligned} F(t) &= 0 \quad \text{for} \quad t < -k \\ &= p \quad \text{for} \quad -k \leqq t < k \\ &= 1 \quad \text{for} \quad t \geqq k . \end{aligned} \tag{4.1}$$

Thus we see that the prior distribution τ_k is essentially concentrated on a subspace of all distributions which have jumps only at the two points $-k$ and $+k$. In other words, under this prior all the observations would be $\pm k$ with probability one.

Next we obtain the minimum Bayes risk corresponding to the prior distribution τ_k.

THEOREM 4.1. *Let* $X_1, \cdots, X_n$ *be a random sample of size* $n > 2 - \alpha - \beta - \gamma - \delta$ *from an unknown cumulative distribution function* $F \in \Theta$ *and let* τ_k *be the prior distribution on* Θ. *Then for the loss function* L *in* (2.1) *the minimum Bayes risk is given by*

$$\begin{aligned} \int \frac{1_{[-k,k)}(t)}{[(\alpha + \beta + \gamma + \delta + n - 2)B(\alpha, \beta)]} &\{[B(\alpha + \gamma, \beta + \delta) \\ - (1 - \alpha - \gamma)B(\alpha + \gamma, \beta + \delta - 1 + n) \cdot 1_{(0,1]}(\alpha + \gamma) \\ - (1 - \beta - \delta)B(\alpha + \gamma - 1 + n, \beta + \delta) \cdot 1_{(0,1]}(\beta + \delta)]\} &\, dW(t) \end{aligned} \tag{4.2}$$

where $1_A(\cdot)$ *is the indicator function of set* A *and*

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1 - x)^{\beta-1}\, dx . \tag{4.3}$$

[2] Originally the results were derived using the method of the Dirichlet Process (Ferguson, 1973). The author is grateful to Professor Ferguson for pointing out this alternative form for the priors, which simplifies the argument considerably.

PROOF. The Bayes risk of an estimator ϕ with respect to the prior τ_k is

$$r(\tau_k, \phi) = E_{\tau_k} E_F[L(F, \phi)] = \int E_{\tau_k} E_F[L_t(F(t), \phi(t))]\, dW(t) \tag{4.4}$$

where $L_t(F(t), \phi(t))$ is the integrand of the loss function L. It is minimized by choosing $\phi(t)$ for each t so that

$$E_{\tau_k} E_F[L_t(F(t), \phi(t))] \tag{4.5}$$

is minimized.

For $t < -k$, according to the prior τ_k, $F(t) = 0$ with probability one. So $E_{\tau_k} E_F[L_t(F(t), \phi(t))]$ will be minimized by taking $\phi(t) \equiv 0$. Similarly for $t \geqq k$, $F(t) = 1$ with probability one, and hence (4.5) is minimized by taking $\phi(t) \equiv 1$.

Now for $-k \leqq t < k$, according to the prior τ_k $F(t)$ is distributed as Be(α, β) and the posterior distribution given the sample is Be$(\alpha + \sum \delta_{x_i}((-\infty, t]), \beta + \sum \delta_{x_i}((t, \infty)))$. $E_{\tau_k} E_F[L_t(F(t), \phi(t))]$ will be minimized for this interval if the conditional expectation of $L_t(F(t), \phi(t))$ given the sample is minimized. That is, if

$$\begin{aligned}[B(\alpha + \textstyle\sum \delta_{x_i}((-\infty, t]), \beta + \sum \delta_{x_i}((t, \infty)))]^{-1} \int_0^1 [p - \phi(t)]^2 \\ \times [p]^{\alpha+\gamma+\sum \delta_{x_i}((-\infty,t])-2}[1-p]^{\beta+\delta+\sum \delta_{x_i}((t,\infty))-2}\, dp\end{aligned} \tag{4.6}$$

is minimized. However, the integral involved in (4.6) will be finite only if $\alpha + \gamma - 1 + \sum \delta_{x_i}((-\infty, t]) > 0$ and $\beta + \delta - 1 + \sum \delta_{x_i}((t, \infty)) > 0$. Since $\sum \delta_{x_i}((-\infty, t])$ and $\sum \delta_{x_i}((t, \infty))$ take only nonnegative integral values, this would amount to requiring t to be in the interval $[x_{([2-\alpha-\gamma]^+)}, x_{(n-[2-\beta-\delta]^++1)})$ where $x_{(r)}$ is the value of the rth order statistic, $x_{(0)} = -\infty$, $x_{(n+1)} = \infty$, $[A]^+ = \max([A], 0)$ and $[A]$ denotes the greatest integer less than or equal to A.

So for each t in $[x_{([2-\alpha-\gamma]^+)}, x_{(n-[2-\beta-\delta]^++1)})$, (4.6) will be finite and it can easily be seen that the minimum is achieved if we let $\phi(t)$ be the mean of the Be$(\alpha + \gamma - 1 + \sum \delta_{x_i}((-\infty, t]), \beta + \delta - 1 + \sum \delta_{x_i}((t, \infty)))$. The minimum value of (4.6) is the variance of this distribution times the factor

$$\frac{B(\alpha + \gamma - 1 + \sum \delta_{x_i}((-\infty, t]), \beta + \delta - 1 + \sum \delta_{x_i}((t, \infty)))}{B(\alpha + \sum \delta_{x_i}((-\infty, t]), \beta + \sum \delta_{x_i}((t, \infty)))}$$

or

$$\frac{B(\alpha + \gamma + \sum \delta_{x_i}((-\infty, t]), \beta + \delta + \sum \delta_{x_i}((t, \infty)))}{(\alpha + \beta + \gamma + \delta + n - 2)B(\alpha + \sum \delta_{x_i}((-\infty, t]), \beta + \sum \delta_{x_i}((t, \infty)))}\,. \tag{4.7}$$

For values of t in $[-k, x_{([2-\alpha-\gamma]^+)})$, we take $\phi(t) \equiv 0$ and for t in $[x_{(n-[2-\beta-\delta]^++1)}, k)$, $\phi(t) \equiv 1$ in order that ϕ be non-decreasing for all values of α, β, γ, and δ.

Thus for $-k \leqq t < k$, the minimum of (4.5) will be equal to

$$\begin{aligned} E_{\tau_k} E_F \Big\{ & 1_{[-k, X_{([2-\alpha-\gamma]^+)})}(t)[F(t)]^{\gamma+1}[1-F(t)]^{\delta-1} + 1_{[X_{([2-\alpha-\gamma]^+)}, X_{(n-[2-\beta-\delta]^++1)})}(t) \\ & \times \frac{B(\alpha + \gamma + \sum \delta_{X_i}((-\infty, t]), \beta + \delta + \sum \delta_{X_i}((t, \infty)))}{[(\alpha + \beta + \gamma + \delta + n - 2)B(\alpha + \sum \delta_{X_i}((-\infty, t]), \beta + \sum \delta_{X_i}((t, \infty)))]} \\ & + 1_{[X_{(n-[2-\beta-\delta]^++1)}, k)}(t)[F(t)]^{\gamma-1}[1 - F(t)]^{\delta+1} \Big\}\,. \end{aligned} \tag{4.8}$$

Noting that $\sum \delta_{X_i}((-\infty, t])$ is a binomial random variable with parameters n and $F(t)$, evaluating the expectations and simplifying we can show that (4.8) is equal to

$$[B(\alpha+\gamma, \beta+\delta) - (1-\alpha-\gamma)B(\alpha+\gamma, \beta+\delta-1+n) \times 1_{(0,1]}(\alpha+\gamma) - (1-\beta-\delta)B(\alpha+\gamma-1+n, \beta+\delta) \times 1_{(0,1]}(\beta+\delta)]/(\alpha+\beta+\gamma+\delta+n-2)B(\alpha, \beta) . \tag{4.9}$$

This together with the above arguments for the cases $t < -k$ and $t \geqq k$, gives the minimum Bayes risk as in (4.2). This completes the proof.

The above proof is simplified considerably if we assume $\alpha, \beta \geqq 1$. Also, since the integral in (4.6) is finite for values of t in the interval $[x_{([2-\gamma-\alpha]^+)}, x_{(n-[2-\beta-\delta]^++1)})$, we may allow negative values for γ and δ with appropriate restrictions on α and β. A typical Bayes estimator with respect to the prior τ_k would be of the form

$$\phi_k(t) = \frac{\alpha+\gamma-1+\sum \delta_{X_i}((-\infty, t])}{\alpha+\beta+\gamma+\delta+n-2} .$$

5. Minimax estimators. In this section we prove our main result; namely, the estimators ϕ_i defined in (3.14), (3.15) and (3.17) are minimax under loss function L_i for $i = 1, 2, 3$ and 4 respectively. To do this we use a well-known lemma which we state here without proof.

LEMMA 5.1. *Let $\hat{F}_k$ be the sequence of Bayes estimators of F with respect to the priors τ_k and let $r(\tau_k, \hat{F}_k)$ be the corresponding sequence of Bayes risks for any loss function. If $r(\tau_k, \hat{F}_k) \to c$, a constant, as $k \to \infty$ and if $\hat{F}$ is any other estimator such that $R(F, \hat{F})$ is $\leqq c$ for all $F \in \Theta$, then $\hat{F}$ is minimax.*

THEOREM 5.1. *Let $X_1, \cdots, X_n$ be a random sample from an unknown cumulative distribution function F on R^1. Then for the loss function L_i and action space Ω, ϕ_i as defined in* (3.15), (3.16) *and* (3.17) *is minimax estimator of F for $i = 1, 2, 3$ and* 4 *respectively.*

PROOF. For the loss function L_1 we take $\alpha = \beta = n^{\frac{1}{2}}/2$ for the prior distributions τ_k. The corresponding sequence of minimum Bayes risk is (by proper substitution in (4.2))

$$\begin{aligned} \int 1_{[-k,k)}(t)\{B(n^{\frac{1}{2}}/2+1, n^{\frac{1}{2}}/2+1)/[(n^{\frac{1}{2}}+n)B(n^{\frac{1}{2}}/2, n^{\frac{1}{2}}/2)]\}\, dW(t) \\ = \int 1_{[-k,k)}(t)\{1/4(n^{\frac{1}{2}}+1)^2\}\, dW(t) \end{aligned} \tag{5.1}$$

which converges as $k \to \infty$, to

$$\{1/4(n^{\frac{1}{2}}+1)^2\} \cdot \int dW(t) . \tag{5.2}$$

From (3.18) we see that the risk of ϕ_1 for this loss function is (5.2) and hence by Lemma 5.1 we conclude that ϕ_1 is minimax estimator of F.

For the loss function L_2, we take $\alpha = \beta = 1$ for the parameters in the prior distribution τ_k. Then, the corresponding sequence of minimum Bayes risk is obtained from (4.2) as

$$(5.3)\qquad \int 1_{[-k,k)}(t)\{B(1,1)/nB(1,1)\}\,dW(t) = n^{-1}\int 1_{[-k,k)}(t)\,dW(t) \to n^{-1}\int dW(t) \quad \text{as} \quad k \to \infty\,.$$

Again from (3.18) we have the risk of ϕ_2 for the loss function L_2 as $n^{-1}\int dW(t)$ and hence by Lemma 5.1 we conclude that ϕ_2 is minimax.

For the loss functions L_3 and L_4, by taking $\alpha = 1$, $\beta = n^{\frac{1}{2}}$ and $\alpha = n^{\frac{1}{2}}$, $\beta = 1$ respectively, one can prove similarly that ϕ_3 and ϕ_4 are minimax estimators of F.

6. Discussion and remarks. The above minimax results obtained in Section 5 are for an arbitrary unknown F. Both Aggarwal (1955) and Taha (1968) have restricted their consideration to unknown F belonging to a family of continuous cumulative distribution functions and to invariant loss functions. Our results are stronger in the sense that the estimators do not depend upon the weight function W unlike the best invariant rules of Aggarwal and Taha who have used the unknown F as the weight function.

The most frequently used estimator of a cumulative distribution function is the sample distribution function. Aggarwal (1955) has shown that if F is assumed to be continuous then the sample distribution function is the best invariant estimator under the loss function

$$(6.1)\qquad L(F, \hat{F}) = \int_{-\infty}^{\infty}\{[F(t) - \hat{F}(t)]^2/F(t)[1 - F(t)]\}\,dF(t)\,.$$

Dvoretzky, Kiefer and Wolfowitz (1956) have shown that the sample distribution function is asymptotically minimax for a very wide class of loss functions. We have shown here that the sample distribution function is minimax under the loss function L_2 described in Section 2.

We have obtained the minimax results only for four special cases of the general loss function L. For other cases, however, the risk function cannot be made independent of F and hence the above method fails to give the minimax estimators. It should be observed that the estimators are of the form $a + b\sum \delta_{X_i}$ and in some cases are not proper distribution functions. The loss functions of the form

$$(6.2)\qquad L(F, \phi) = \int |F(t) - \phi(t)|\,dW(t)$$

are not easy to handle. The main difficulty arises in evaluating the Bayes risk. Moreover, the Bayes estimator with respect to this loss function is the median of the posterior distribution which is not necessarily of the form $a + b\sum \delta_{X_i}$ and hence the above method of choosing a and b which makes the risk function independent of F fails.

If we take the weight function W to assign all the mass at a point, say t_0, then the problem reduces to that of finding the minimax estimator of a binomial parameter $F(t_0)$ which is seen as a particular case of the above result,

$$(6.3)\qquad \hat{F}(t_0) = [1/2(n^{\frac{1}{2}} + 1)] + [1/(n^{\frac{1}{2}} + n)]\sum \delta_{X_i}((-\infty, t_0])\,.$$

Here $\delta_{X_i}((-\infty, t_0])$ are independent and identically distributed Bernoulli random variables, each with the success probability of $F(t_0)$. $\hat{F}(t_0)$ can also be written in a more familiar form

$$\hat{F}(t_0) = 1/2(n^{\frac{1}{2}} + 1)] + [n^{\frac{1}{2}}/(n^{\frac{1}{2}} + 1)]\bar{\delta}(t_0) \tag{6.4}$$

where $\bar{\delta}(t_0) = [1/n] \sum \delta_{X_i}((-\infty, t_0])$. In fact in the case where W is degenerate at a single point, the minimax estimator of $F(t_0)$ is unique, and so admissible, minimax estimator (Lehmann, 1949–50). For more complicated W, however, the problem is not simple.

Since the sequence of prior distributions used above depends on the sample size n, it is not possible to extend this result to the sequential case, except for the loss function L_2. The extension of these results to higher dimensions should be straightforward.

Acknowledgments. The author is grateful to Professor J. S. Rustagi for his guidance and encouragement, and to Professors T. S. Ferguson, J. Kiefer and the referee for their valuable comments and suggestions on the earlier version of the paper. The author wishes to thank Mr. Dudley C. Ward of Technology Incorporated for providing the typing facility.

REFERENCES

[1] AGGARWAL, O. P. (1955). Some minimax invariant procedures for estimating a cumulative distribution function. *Ann. Math. Statist.* **26** 450–462.

[2] DUBINS, L. E. and FREEDMAN, D. A. (1963). Random distribution function. *Bull. Amer. Math. Soc.* **69** 548–551.

[3] DUBINS, L. E. and FREEDMAN, D. A. (1967). Random distribution function. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **2** 183–214. Univ. of California Press.

[4] DVORETZKY, A., KIEFER, J. and WOLFOWITZ, J. (1956). Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *Ann. Math. Statist.* **27** 642–669.

[5] FERGUSON, T. S. (1973). A Bayesian analysis of some non-parametric problems. *Ann. Statist.* **1** 209–230.

[6] KRAFT, C. H. and VAN EEDEN, C. (1964). Bayesian bio-assay. *Ann. Math. Statist.* **35** 886–890.

[7] KRAFT, C. H. (1964). A class of distribution function processes which have densities. *J. Appl. Probability* **1** 385–388.

[8] LEHMANN, E. L. (1949–50, reprinted 1966). Notes on the theory of estimation, recorded by Colin Blyth, Univ. of California, Berkeley, 4–30.

[9] READ, R. R. (1972). The asymptotic inadmissibility of the sample distribution function. *Ann. Math. Statist.* **43** 89–95.

[10] TAHA, M. (1968). On estimating the accumulative distribution (abstract). *Ann. Math. Statist.* **39** 1368.

DEPARTMENT OF THE AIR FORCE
AEROSPACE RESEARCH LABORATORIES (AFSC)
WRIGHT PATTERSON AFB, OHIO 45433