The Annals of Statistics
1992, Vol. 20, No. 4, 2192–2195

MINIMAXITY OF THE BEST INVARIANT ESTIMATOR OF A DISTRIBUTION FUNCTION UNDER THE KOLMOGOROV–SMIRNOV LOSS[1]

BY QIQING YU[2] AND ESWAR PHADIA

SUNY Stony Brook and William Paterson College

For the invariant decision problem of estimating a continuous distribution function with the Kolmogorov–Smirnov loss, it is proved that the best invariant estimator is minimax.

1. Introduction. The best invariant estimators for a continuous cumulative distribution function under monotone transformations and the weighted Cramér–von Mises loss function or more general invariant loss functions were introduced by Aggarwal (1955). Since then there has been a longstanding conjecture that the best invariant estimator, d_0, is minimax for $n \geq 1$ under the loss

$$L(f, a) = \int |F(t) - a(t)|^k h(F(t))\, dF(t), \tag{1.1}$$

where $h(t)$ is a nonnegative weight function and $a(t)$ is a nondecreasing function from $(-\infty, \infty)$ into $[0, 1]$ [see, for example, Ferguson (1967), page 197]. This conjecture was proved recently under the loss (1.1) [see Yu (1989, 1992) and Yu and Chow (1991)].

A parallel problem was to consider the Kolmogorov–Smirnov loss function

$$L(F, a) = \sup_t \{|F(t) - a(t)|\}, \tag{1.2}$$

which is also invariant under the monotone transformations. This loss function is difficult to handle analytically and therefore not much was accomplished for a long time. Brown (1988) obtained the best invariant estimator under this loss for the sample size $n = 1$ by hand and investigated its admissibility under the assumption that the unknown distribution function is discrete. This was followed up by Friedman, Gelman and Phadia (1988) who obtained the best invariant estimator d_0 for sample sizes $n > 1$ and proved its uniqueness. Again, the obvious question is whether d_0 is minimax.

In this note, the question is answered affirmatively. Thus the minimaxity conjecture is solved completely in the finite sample classical invariant estima-

---

Received May 1990; revised March 1991.
[1]Partially supported by the Governor's Challenge for Excellence Grant and NSF Grant DMS-90-01194.
[2]Part of the work was done while the author was visiting William Paterson College.
AMS 1980 *subject classifications.* Primary 62C15; secondary 62D05.
Key words and phrases. Best invariant estimator, Kolmogorov–Smirnov loss, minimaxity.

tion problems. (Similar, but partial, results have been obtained regarding the admissibility of d_0 and they will be published elsewhere as a separate article.)

2. Main result. Let $X_1, \dots, X_n$ be a sample of size n from a continuous distribution function F. Let $Y_1 < \cdots < Y_n$ be the order statistics of the X_i's. For convenience, we write $Y = (Y_1, \dots, Y_n)$. Let Θ denote the family of all continuous distribution functions, dF the measure induced by the distribution function F, that is, $\{dF\{(a,b)\} = F(b) - F(a)$, $(dF)^k$ the product measure $dF \times \cdots \times dF$ with k factors and S^k the product set $S \times \cdots \times S$ with k factors.

To prove the minimaxity of d_0, we need the following results.

THEOREM 1 [Yu and Chow (1991)]. *Suppose that the sample size $n \geq 1$ and $d = d(Y, t)$ is a nonrandomized estimator with finite risk and is a (measurable) function of the order statistic Y. For any $\varepsilon, \delta > 0$, there exist a uniform distribution function F on a positive-Lebesgue-measure subset J and an invariant estimator d_1 such that*

$$(dF)^{n+1}(\{(Y_1, \dots, Y_n, t) : |d(Y,t) - d_1(Y,t)| \geq \varepsilon\}) \leq \delta.$$

LEMMA 1. *Suppose that the sample size $n \geq 1$ and $\varepsilon \in (0,1)$. For any two arbitrary estimators d and d_n, if*

$$\exists\, F \in \Theta \text{ such that } (dF)^{n+1}\Big(\big\{|d(\vec{x},t) - d_n(\vec{x},t)| > \varepsilon\big\}\Big) < \varepsilon, \tag{2.1}$$

then $|R(F,d) - R(F,d_n)| \leq 3\sqrt{\varepsilon}$.

PROOF. Let J be the support of F. Then, (2.1) yields

$$(dF)^n\Big(\Big\{\vec{x} \in J^n : dF\Big(\big\{t : |d(\vec{x},t) - d_n(\vec{x},t)| > \varepsilon\big\}\Big) \geq \sqrt{\varepsilon}\Big\}\Big) < \sqrt{\varepsilon}, \tag{2.2}$$

that is, except for a set of small measure dF, most of $\vec{x} \in J^n$ satisfy $dF(\{t : |d(\vec{x},t) - d_n(\vec{x},t)| > \varepsilon\}) < \sqrt{\varepsilon}$. Let

$$S = \Big\{\vec{x} \in J^n : dF\Big(\big\{t : |d(\vec{x},t) - d_n(\vec{x},t)| > \varepsilon\big\}\Big) < \sqrt{\varepsilon}\Big\}. \tag{2.3}$$

For $\vec{x} \in S$, $\exists$ m points $t_1 < \cdots < t_m$ such that

$$\big|d\big(\vec{x},t_j\big) - d_n\big(\vec{x},t_j\big)\big| \leq \varepsilon,\ j = 1, \dots, m, \quad \text{and} \quad F(t_{j+1}) - F(t_j) < \sqrt{\varepsilon},\ \forall\, j,$$

where $t_0 = -\infty$ and $t_{m+1} = \infty$. Then, for $j = 0$, we have

$$\begin{aligned}
&\sup\big\{|F(t) - d(\vec{x},t)| : t \leq t_1\big\} \\
&\quad \leq \sup\Big\{|F(t_1) - 0|, \big|F(t_1) - \sqrt{\varepsilon} - \big(d_n(\vec{x},t_1) + \varepsilon\big)\big|\Big\} \\
&\quad \leq \sup\big\{|F(t) - d_n(\vec{x},t)| : t \leq t_1\big\} + 2\sqrt{\varepsilon};
\end{aligned}$$

for $j = 1, \dots, m-1$, we have

$$\begin{aligned}
&\sup\left\{\left|F(t) - d(\vec{x}, t)\right| : t_j \le t \le t_{j+1}\right\} \\
&\le \sup\left\{\left|F(t_j) + \sqrt{\varepsilon} - \left(d_n\left(\vec{x}, t_j\right) - \varepsilon\right)\right|, \left|F(t_{j+1}) - \sqrt{\varepsilon} - \left(d_n\left(\vec{x}, t_{j+1}\right) + \varepsilon\right)\right|\right\} \\
&\le \sup\left\{\left|F(t) - d_n(\vec{x}, t)\right| : t_j \le t \le t_{j+1}\right\} + 2\sqrt{\varepsilon}\,.
\end{aligned}$$

Additionally, $\sup\{|F(t) - d(\vec{x}, t)|:\ t \ge t_m\} \le \sup\{|F(t) - d_n(\vec{x}, t)|:\ t \ge t_m\} + 2\sqrt{\varepsilon}$. As a consequence,

$$\sup_t \left\{\left|F(t) - d(\vec{x}, t)\right|\right\} \le \sup_t \left\{\left|F(t) - d_n(\vec{x}, t)\right|\right\} + 2\sqrt{\varepsilon}\,, \quad \text{for } \vec{x} \in S.$$

Similarly we can show that

$$\sup_t \left\{\left|F(t) - d(\vec{x}, t)\right|\right\} \ge \sup_t \left\{\left|F(t) - d_n(\vec{x}, t)\right|\right\} - 2\sqrt{\varepsilon} \quad \text{for } \vec{x} \in S.$$

Thus

$$\left|\sup_t \left\{\left|F(t) - d(\vec{x}, t)\right|\right\} - \sup_t \left\{\left|F(t) - d_n(\vec{x}, t)\right|\right\}\right| \le 2\sqrt{\varepsilon} \quad \text{for } \vec{x} \in S. \tag{2.4}$$

Now the lemma follows from (2.2), (2.3) and (2.4). □

THEOREM 2. *Under loss* (1.2), *the best invariant estimator d_0 is minimax for sample size $n \ge 1$.*

PROOF. Given an estimator d which is a function of order statistics Y and has finite risk for any $F \in \Theta$, by Theorem 1, there exists an $F \in \Theta$ and there exists an invariant estimator d_1 (and thus of constant risk) such that

$$(dF)^{n+1}\left(\left\{(\vec{x}, t): \left|d(\vec{x}, t) - d_1(\vec{x}, t)\right| > \varepsilon\right\}\right) < \varepsilon.$$

By Lemma 1, $|R(F, d) - R(F, d_1)| \le 3\sqrt{\varepsilon}$. Thus it follows that

$$3\sqrt{\varepsilon} + R(F, d) \ge R(F, d_1) \ge R(F, d_0).$$

Note that ε and d are arbitrary, so $\inf_d \sup_{F \in \Theta} R(F, d) = R(F, d_0)$. □

Acknowledgment. We gratefully acknowledge the assistance of a referee in simplifying the conditions in Lemma 1.

REFERENCES

AGGARWAL, O. P. (1955). Some minimax invariant procedure of estimating a cumulative distribution function. *Ann. Math. Statist.* **26** 450–462.

BROWN, L. D. (1988). Admissibility in discrete and continuous invariant nonparametric problems and in their multivariate analogs. *Ann. Statist.* **16** 1567–1593.

FERGUSON, T. S. (1967). *Mathematical Statistics, A Decision Theoretic Approach*. Academic, New York.

FRIEDMAN, D., GELMAN, A. and PHADIA, E. (1988). Best invariant estimator of a distribution function under the Kolmogorov–Smirnov loss function. *Ann. Statist.* **16** 1254–1261.

YU, Q. (1989). Methodology for the invariant estimation of a continuous distribution function. *Ann. Inst. Statist. Math.* **41** (3) 503–520.

YU, Q. and CHOW, M. S. (1991). Minimaxity of the empirical distribution function in invariant problem. *Ann. Statist.* **19** 935–951.

YU, Q. (1992). Minimax invariant estimator of a continuous distribution function. *Ann. Inst. Statist. Math.* To appear.

DEPARTMENT OF APPLIED MATHEMATICS
AND STATISTICS
SUNY
STONY BROOK, NEW YORK 11794

DEPARTMENT OF MATHEMATICS
WILLIAM PATERSON COLLEGE
WAYNE, NEW JERSEY 07470