Ann. Inst. Statist. Math.
Vol. 41, No. 3, 503-520 (1989)

METHODOLOGY FOR THE INVARIANT ESTIMATION OF A CONTINUOUS DISTRIBUTION FUNCTION

QIQING YU

Mathematics Department, Zhongshan University, the People's Republic of China

(Received April 21, 1987; revised November 11, 1988)

Abstract. Consider both the classical and some more general invariant decision problems of estimating a continuous distribution function, with the loss function $L(F,a)=\int(F(t)-a(t))^2h(F(t))dF(t)$ and a sample of size n from F. It is proved that any nonrandomized estimator can be approximated in Lebesgue measure by the more general invariant estimators. Some methods for investigating the finite sample problem are discussed. As an application, a proof that the best invariant estimator is minimax when the sample size is 1 is given.

Key words and phrases: Admissibility, admissibility within U_1, invariant estimator, minimaxity.

1. Introduction

Since Aggarwal (1955) found the best invariant estimator of an unknown continuous distribution function $F(t)$, under the loss

$$L(F,a)=\int(F(t)-a(t))^2h(F(t))dF(t)\,, \tag{1.1}$$

different methods have been used in investigating the decision theoretic properties of the best invariant estimator d_0. One interesting fact is that when $h(t)=t^{-1}(1-t)^{-1}$, the best invariant estimator is the same as the empirical distribution function $\hat{F}(t)$.

The asymptotical method has been used in approaching the problem. For instance, Dvoretzky *et al.* (1956) studied the asymptotical minimaxity property of the best invariant estimator for some loss function; and Read (1972) considered the asymptotical admissibility property of the best invariant estimator. However, this method does not describe the decision theoretic properties when the sample size n is finite.

Some considered the problem in another way. They considered the loss

$$L(F, a) = \int (F(t) - a(t))^2 h(F(t)) dW(t) , \tag{1.2}$$

where W is a nonzero finite measure, instead of (1.1). For example, Phadia (1973) and Cohen and Kuo (1985) considered this problem in this way. The decision theoretic properties of $\hat{F}(t)$ in this set-up are, of course, interesting. However, the loss function is no longer invariant. So it is not surprising that the conclusion in this set-up is different from that in the classical set-up.

Brown (1988) carefully formulated the discrete analogues of the continuous nonparametric estimation problem. Besides investigating the decision problems themselves, he also hoped to approximate the classical set-up by the analogues. This effort does not seem to be successful, since the property of the best invariant estimator in discrete problems does not coincide with the one in continuous problems for the special case with the loss

$$L(F, a) = \int (F(t) - a(t))^2 dF(t) . \tag{1.3}$$

The most important breakthrough in this finite sample continuous nonparametric estimation problem is made by Brown (1988). He found an estimator to improve on the best invariant estimator when the loss is (1.3). It turns out that Brown's estimator belongs to a set $U_m = U_m(s_1, \ldots, s_m)$ (see (2.8)), of nonrandomized invariant estimators of more general invariant decision problems. Yu (1986, 1987 and 1989*b*) extended Brown's result and formulated a sequence of more general invariant decision problems. This seems to be a good methodology in investigating the classical finite sample invariant estimation problems and leads to some important results, e.g., $\hat{F}(t)$ is admissible iff the sample size is 1 or 2 with the loss (1.1), where $h(t) = t^{-1}(1 - t)^{-1}$.

So it is worthwhile to study properties of the sets U_m. In Section 3, we prove that for any nonrandomized estimator d, there is a sequence of estimators $\{d_k\}$ such that $d_k \to d$ in Lebesgue measure and each d_k belongs to some $U_m(s_1, \ldots, s_m)$. In Section 4, we show that for the sample size $n = 1$, the best invariant estimator is admissible iff it is admissible within U_1. We also show that for the sample size $n = 2$, under some assumptions on $h(t)$ in the loss (1.1) similar to $h(t) = t^{-1}(1 - t)^{-1}$, the best invariant estimator is admissible iff it is admissible just within U_1 instead of U_2. We conjecture that in general for the sample size $n > 1$ the best invariant estimator is admissible iff it is admissible within U_n (within U_1 would probably be too strong to be true).

The discussion of the sets U_m is necessary both for theoretical and for practical reasons. The admissibility of the best invariant estimator has not been solved completely even for the special form of the weight function $h(t) = t^\alpha(1-t)^\beta$ in the loss (1.1). On the other hand, estimators in U_m are useful since it is relatively easy to compute the risk, especially when $m = 1$.

Yu (1989*a*) considered a subset $\tilde{Q}_d$ of $(0,1)$ (see (5.1)). It plays an important role in determining the admissibility of d_0 for $n = 1$. As pointed out by Yu (1989*a*) it leads to an important fact: there is no continuous estimator which is as good as d_0 when $n = 1$ (so even when d_0 is inadmissible, we cannot expect to find a continuous estimator which can improve on d_0, though we are estimating a continuous distribution function). In Section 5, it leads to a proof that d_0 is minimax for $n = 1$, too. Since $\tilde{Q}_d$ cannot be generalized to the case $n > 1$ directly, we formulate a subset Q_d of $(0,1)$ (see (5.1′)), and provide a sufficient condition of the minimaxity of d_0. Also we propose to consider the minimaxity of d_0 within a class of estimators in Section 5. In proving Theorem 5.4 related to Q_d, we illustrate the method for constructing a continuous distribution function F for given estimator d and given $\varepsilon > 0$ such that $R(F,d) \geq R(F,d_0) - \varepsilon$ (a similar continuous distribution function was constructed in Yu (1987) in proving the admissibility of $\hat{F}(t)$ for the sample size $n = 2$). This is a useful idea in considering the minimaxity of d_0.

It is well known that the decision problem of the invariant estimation of a continuous distribution function with the support on $(-\infty, +\infty)$ is equivalent to that with the support on $(0,1)$ (for example, see Brown (1988)). Thus for convenience we consider the latter problem in this paper.

2. Notations and formulation

Let Θ denote the parameter space. Here

$$(2.1) \qquad \Theta = \{F;\ F \text{ is a continuous distribution function on } (0,1)\}.$$

Let $X_1,\ldots,X_n$ be a sample of size n from F in Θ.

Let $A = \{a(t); a(t)$ is a nondecreasing function from $(0,1)$ into $[0,1]\}$ denote the action space.

Let

$$(2.2) \qquad L(F,a) = \int (F(t) - a(t))^2 h(F(t))\,dF(t)$$

be the loss function, where $h(t)$ is nonnegative Lebesgue measurable on $(0,1)$ and $\int_0^1 t(1-t)h(t)\,dt < +\infty$.

Let $s_1, s_2, \dots, s_m \in [0,1]$, $s_1 < \cdots < s_m$,

$$G_0 = \{\phi_g; \phi_g(x) = (g(x_1), \dots, g(x_n)), g(t) < g(s) \text{ if } t < s \text{ and } g((0,1)) = (0,1)\}, \tag{2.3}$$

$$G_m(s_1, \dots, s_m) = \{\phi_g \in G_0; g(s_i) = s_i, i = 1, \dots, m\}. \tag{2.4}$$

Under the above assumptions and notations, the decision problem (Θ, A, L) with the distribution family Θ over $(0,1)$ is invariant under the group of transformations G_0. Then (see Aggarwal (1955)) the nonrandomized invariant estimators are of the form

$$d(t) = \sum_{i=0}^{n} u_i 1(Y_i \le t < Y_{i+1}), \tag{2.5}$$

where $1(E)$ is the indicator function of E, $Y_0 = 0$, $Y_{n+1} = 1$, $Y_1 < \cdots < Y_n$ are the order statistics of the sample $X_1, \dots, X_n$. The best invariant estimator, denoted by $d_0(t)$, has the form (2.5) with

$$u_i = \frac{\int_0^1 t^{i+1}(1-t)^{n-i} h(t)\,dt}{\int_0^1 t^i (1-t)^{n-i} h(t)\,dt}, \quad i = 0, \dots, n, \tag{2.6}$$

and has constant risk. Since G_m is a subgroup of G_0, under the above assumptions and notations, for each $m = 0, 1, \dots,$ and $s_1, \dots, s_m$, the decision problem (Θ, A, L) with the distribution family Θ over $(0,1)$, is also invariant under the group of transformations $G_m(s_1, \dots, s_m)$.

The form of nonrandomized invariant estimators d of the decision problem $(\Theta, A, L, G_m(s_1, \dots, s_m))$ can be described as follows. d assigns weight $u_{I,j}$, depending on the ranks I of $s_1, \dots, s_m$ among $(X_1, \dots, X_n, 0, 1, s_1, \dots, s_m)$, to the order statistics $Y_0^I, \dots, Y_{n+m+1}^I$ of $(X_1, \dots, X_n, 0, 1, s_1, \dots, s_m)$, i.e.,

$$d(Y, t) = \sum_{j=0}^{n+m} u_{I,j} 1(Y_j^I \le t < Y_{j+1}^I), \tag{2.7}$$

where I is a random vector, $I = (I_1, \dots, I_m)$, defined by $I_l = \max\{j \ge 0: Y_j \le s_l\}$, $l = 1, \dots, m$. So

$$\{I_j = k\} = \{Y_k \le s_j < Y_{k+1}\}, \quad k = 0, \dots, n.$$

$$Y_j^{I_1} = \begin{cases} Y_j & \text{if } 0 \le j \le I_1; \\ s_1 & \text{if } j = I_1 + 1; \\ Y_{j-1} & \text{if } I_1 + 1 < j \le n + 2; \end{cases}$$

$$Y_j^{(I_1,\dots,I_{k+1})} = \begin{cases} Y_j^{(I_1,\dots,I_k)} & \text{if} \quad 0 \le j \le I_{k+1} + k\,, \\ s_{k+1} & \text{if} \quad j = I_{k+1} + k + 1\,, \\ Y_{j-1}^{(I_1,\dots,I_k)} & \text{if} \quad I_{k+1} + k + 1 < j \le n + k + 2\,, \end{cases}$$

$$\text{for } k = 1,\dots,m-1\,.$$

Given $I = \boldsymbol{i} = (i_1,\dots,i_m)$, $u_{I,j} = u_{i,j}$ is a constant, $0 \le u_{I,j} \le u_{I,j+1} \le 1$, $m = 1,2,\dots, j = 0,1,\dots,m+n-1$. The nonrandomized invariant estimators denoted by $U_m(s_1,\dots,s_m)$ of the decision problem (Θ, A, L, G_m) are essentially complete within the class of randomized invariant estimators. Abusing notation, write

$$U_m = U_m(s_1,\dots,s_m)\,, \tag{2.8}$$

where U_0 is the set of classical invariant estimators. Given $m = 1$, define a mapping f from U_1 to $\{(u_{ij})_{(n+1)\times(n+2)};\, u_{ij} \le u_{i,j+1}, i = 0,\dots,n,\ j = 0,\dots,n\}$ by

$$f\colon \sum_{j=0}^{n+1} u_{I,j} 1\{Y_j^I \le t < Y_{j+1}^I\} \to (u_{ij})_{(n+1)\times(n+2)}\,. \tag{2.9}$$

The risk functions of d in U_0 and U_1 are interesting. If $d \in U_0$, then

$$R(F,d) = \sum_{i=0}^{n} \int_0^1 (t-u_i)^2 h(t) \binom{n}{i} (1-t)^{n-i} t^i dt\,. \tag{2.10}$$

If $d \in U_1$, then

$$\begin{aligned} R(F,d) = {} & \sum_{i=0}^{n} \sum_{j=0}^{i} \int_0^p (t-u_{ij})^2 h(t) \binom{n}{j} \\ & \cdot \binom{n-j}{i-j} t^j (p-t)^{i-j} (1-p)^{n-i} dt \\ & + \sum_{i=0}^{n} \sum_{j>i}^{n+1} \int_p^1 (t-u_{ij})^2 h(t) \binom{n}{j-1} \\ & \cdot \binom{j-1}{i} (1-t)^{n-j+1} (t-p)^{j-1-i} p^i dt\,, \end{aligned} \tag{2.11}$$

where $p = F(s_1)$. For the whole discussion above, refer to Yu (1986*b*).

DEFINITION 2.1. Let U be a subset of the space Ω of all estimators. If d belongs to U and no estimator in U can improve on d, then we say that d is admissible within U.

Given the sample size n, let

(2.12) N be the collection of all nonrandomized Lebesgue (in $(0,1)^{n+1}$) measurable estimators;

and let

$$U = \bigcup_{m=0}^{\infty} \left(\bigcup_{s_1,\ldots,s_m \in (0,1)} U_m(s_1,\ldots,s_m) \right). \tag{2.13}$$

3. Relation between N and U

We are going to discuss the relation between several sets in the coming sections. Recall that $d_0 \in N$. As discussed by Brown (1988) and Yu (1986), for some special weight function $h(t)$, in order to investigate the admissibility of the best invariant estimator, it suffices to study its decision theoretic properties within $\bigcup_{m=1}^{\infty} U_m(s_1,\ldots,s_m)$. Theorem 3.1 tells an interesting fact about the relation between N and U (see (2.12) and (2.13)).

THEOREM 3.1. *Let $X=(X_1,\ldots,X_n)$ be a sample of size n. Let $Y=(Y_1,\ldots,Y_n)$ be the order statistics of X. $\forall\, d(Y,t) \in N$, $\exists\, \{d_k\} \subset U$ such that $d_k \to d$ in Lebesgue measure m^{n+1}.*

PROOF. As we know that $d \in U_k(s_1,\ldots,s_k)$ iff $d = \sum_{j=0}^{n+k} u_{I,j} 1(Y_j^I \le t < Y_{j+1}^I)$ (see (2.7)), i.e., d is constant for $t \in (Y_j^I, Y_{j+1}^I)$, where Y_j^I's are the order statistics of $0, 1, Y_1,\ldots, Y_n$ and $s_1,\ldots,s_k$. Furthermore, if the orders of $0, 1, Y_1,\ldots, Y_n$ and $s_1,\ldots,s_k$ do not change, neither do $u_{I,j}$'s. So it suffices to show $\forall\, \varepsilon$, $\delta > 0$, $\exists$ finite many points $s_1,\ldots,s_K$, integer m and a step function d_ε such that

(a) for $t \in (s_{i_0}, s_{i_0+1})$, $x_1 \in (s_{i_1}, s_{i_1+1}),\ldots, x_n \in (s_{i_n}, s_{i_n+1})$, $d_\varepsilon(\boldsymbol{x}, t)$ is constant and assumes one value of $0, 1/m,\ldots, 1$;

(b) $d_\varepsilon(\boldsymbol{x}, t)$ is nondecreasing in t;

(c) $m^{n+1}\{|d_\varepsilon - d| > \delta\} < \varepsilon$.

In the following 4 steps we try to simplify our consideration step by step.

(1) Suppose d is an arbitrary estimator of F. Let

$$d_m = \sum_{i=0}^{m} (i/m) 1((Y,t) \in B_i),$$

where $B_i = \{(Y,t)\colon d \in [i/m, (i+1)/m)\}$, $i = 0,\ldots, m$. Then $d_m \to d$ uniformly on $(0,1)^{n+1}$ as $m \to \infty$.

(2) By (1), we can assume, without loss of generality, that d assumes

$0, 1/m, \ldots, 1$ only. Let

$$E_i = \{y: d(y,t) = i/m \ \forall\, t \in (0,1)\}, \qquad i = 0,\ldots,m\,;$$
$$E_{i_1,i_2} = \{y: d(y,t) = (i_1/m)1(t < t_y) + (i_2/m)1(t_y \le t)\}\,,$$

where t_y is not 0 or 1 and depends on y, $0 \le i_1 < i_2 \le m; \ldots$

$$E_{0,\ldots,m} = \{y: d(y,t) \text{ assumes exactly } m+1 \text{ distinct values}\}\,.$$

Let Δ be the symmetric difference. $\forall\, \eta > 0$, $\exists$ finite union $O_{i_1,\ldots,i_j}$ of disjoint subsets with the form $I_1 \times \cdots \times I_n$, where I_i's are intervals, satisfying: $\bigcup_{i_1,\ldots,i_j} O_{i_1,\ldots,i_j} = (0,1)^n$;

$m^n(O_i \,\Delta\, E_i) < \eta$ and define

$$d_\eta = i/m \qquad \text{for} \qquad y \in O_i, \qquad i = 0,\ldots,m\,;$$

$m^n(O_{i_1,i_2} \,\Delta\, E_{i_1,i_2}) < \eta$ and define

$$d_\eta = \begin{cases} (i_1/m)1(t < 1/2) + (i_2/m)1(t \ge 1/2) & \text{if} \quad y \in O_{i_1,i_2} \backslash E_{i_1,i_2}\,, \\ d(y,t) & \text{if} \quad y \in O_{i_1,i_2} \cap E_{i_1,i_2}\,; \end{cases}$$

where $0 \le i_1 < i_2 \le m$;

$\vdots$

$m^n(O_{0,\ldots,m} \,\Delta\, E_{0,\ldots,m}) < \eta$ and define

$$d_\eta = \begin{cases} \sum_{i=1}^{m} (1/m)1(t \ge i/(m+1)) & \text{if} \quad y \in O_{0,\ldots,m} \backslash E_{0,\ldots,m}\,, \\ d(y,t) & \text{if} \quad y \in O_{0,\ldots,m} \cap E_{0,\ldots,m}\,. \end{cases}$$

By construction $d_\eta \to d$ as $\eta \to 0$ in Lebesgue measure on $(0,1)^{n+1}$. Note

($*$) d_η assumes exactly j values $i_1/m, \ldots, i_j/m$ for $y \in O_{i_1,\ldots,i_j}$.

(3) By (1) and (2), we can assume, without loss of generality, that d satisfies ($*$). Noting that the jump points depend on y and might be infinitely many, we try to simplify to the case where there are only a finite amount of jump points. It suffices to consider $d(y,t) = (i_1/m)1(t < t_y) + (i_2/m)1(t \ge t_y)$ for $y \in O = I_1 \times \cdots \times I_n$, where I_i's are intervals.

$\forall\, \varepsilon > 0$, $\exists\, b_1, \ldots, b_k$, where k depends on ε such that $\forall\, t_y$, $\exists\, b_i \in \{b_1,\ldots,b_k\}$ satisfying $|t_y - b_i| < \varepsilon$. Let $b_i = \min\{b_h: |t_y - b_h| < \varepsilon\}$ for the given $y \in O$ and define $d_\delta(y,t) = (i_1/m)1(t < b_i) + (i_2/m)1(t \ge b_i)$; for y not in O, define $d_\delta(y,t)=d(y,t)$. Then $d_\delta \to d$ in Lebesgue measure on $(0,1)^{n+1}$ as $\delta \to 0$.

(4) By discussion in (3), we can assume, without loss of generality, that d satisfies (*) and the set of possible jump points are finite, though they still depend on y. Consider again that for $y \in O = I_1 \times \cdots \times I_n$, where I_i's are intervals, $d(y,t)=(i_1/m)1(t<b_i)+(i_2/m)1(t \geq b_i)$ and there are only k b_i's.

Let $E_i=\{y \in O\text{: The jump point is } b_i\}$. Note that E_i might not be a nice set such as a finite union of disjoint subsets like O. However, $\forall\, \varepsilon>0$, $\exists$ finite union O_i of disjoint subsets with the form $I_1 \times \cdots \times I_n$, where I_i's are intervals such that $\bigcup_{i=1}^{k} O_i = O$; $mO_i \,\Delta\, E_i < \varepsilon/(k+1)$. Define

$$d_\varepsilon(y,t)=\begin{cases} (i_1/m)1(t<b_i)+(i_2/m)1(t\geq b_i) & \text{if } \ y \in O_i, \quad i=1,\dots,k\,; \\ d(y,t) & \text{otherwise}\,. \end{cases}$$

Then $d_\varepsilon \to d$ in Lebesgue measure on $(0,1)^{n+1}$ as $\varepsilon \to 0$.

This procedure can be generalized to general cases that for $y \in$ some O defined similarly as above, $d(y,t)$ assumes exactly h values $i_1/m,\dots,i_h/m$ with jump points only at b_i's. Note that this d_ε satisfies the condition (a), (b) and (c) in the beginning of the proof ($\{b_1,\dots,b_K\}$ contains all jump points and endpoints of intervals I's mentioned above) and thus finishes our proof. □

4. U_1 and admissibility of d_0

When the sample size $n=1$, the best invariant estimator is

$$d_0(t)=\begin{cases} u_0 & \text{if} \quad t<X\,, \\ u_1 & \text{if} \quad t \geq X\,, \end{cases} \tag{4.1}$$

where $u_0=\int_0^1 t(1-t)h(t)dt \Big/ \int_0^1 (1-t)h(t)dt$ and $u_1=\int_0^1 t^2h(t)dt \Big/ \int_0^1 th(t)dt$. For convenience, define

$$V=\{d(t)\text{: } R(F,d) \leq R(F,d_0)\ \forall\, F \in \Theta\}\,. \tag{4.2}$$

When the loss function is of the form (2.2), with $h(t)=t^\alpha(1-t)^\beta$, $\alpha,\beta \geq -1$, and the sample size $n=1$, Yu (1989*a*) pointed out that d_0 is admissible iff it is admissible within U_1. In fact, it is true for general $h(t)$ under the assumptions on $h(t)$ given in Section 2. It is not hard to simulate the argument in Section 3 of Yu (1989*a*) and get the following proposition.

PROPOSITION 4.1. *Suppose $t(1-t)h(t)$ is integrable and the sample size $n = 1$. Then $d_0(t)$ is admissible iff $d_0(t)$ is admissible within U_1.*

As an application of Proposition 4.1, look at the following example.

Example 4.1. Suppose $h(t) = (\sin t)^{-1}$ and $n = 1$. Then the best invariant estimator d_0 is admissible.

PROOF. It suffices to show that d_0 is admissible within U_1.

$$d_0(t) = \begin{cases} 0 & \text{if} \quad t < X , \\ u_1 & \text{if} \quad t \geq X , \end{cases}$$

where $u_1 = \int_0^1 t^2(\sin t)^{-1}dt \Big/ \int_0^1 t(\sin t)^{-1}dt$. If $d \in V \cap U_1$, then

$$f(d) = \begin{pmatrix} u_{00} & u_{01} & u_{02} \\ u_{10} & u_{11} & u_{12} \end{pmatrix} = \begin{pmatrix} u_{00} & 0 & u_1 \\ 0 & u_1 & u_{12} \end{pmatrix}$$

by (2.9). Note $0 \leq u_{00} \leq u_{01} = 0$, so $u_{00} = 0$. Let $u_{12} = u_1 + 2c$, since $u_{12} \geq u_1$ implies $c \geq 0$. Now by (2.11)

$$\begin{aligned} R(F,d) - R(F,d_0) &= p\int_p^1 [(t - u_1 - 2c)^2 - (t - u_1)^2]h(t)dt \\ &= p\int_p^1 [-4c(t - u_1 - c)]h(t)dt \end{aligned}$$

and

$$\begin{aligned} &\lim_{p \to 0} p^{-1}[R(F,d) - R(F,d_0)] \\ &\quad = \lim_{p \to 0} \int_p^1 [4c(u_1 + c)](\sin t)^{-1}dt = +\infty > 0 , \end{aligned}$$

if $c > 0$. So $d = d_0$. This means d_0 is admissible within U_1. □

When the sample size $n = 2$ we conjecture that d_0 is admissible iff d_0 is admissible within U_2. In fact, by simulating the argument in Section 3 of Yu (1987), we can get the following conclusion.

PROPOSITION 4.2. *Suppose the sample size $n = 2$. Assume $t(1-t)h(t)$ is integrable, but $\int_0^1 th(t)dt = \infty$ and $\int_0^1 (1-t)h(t)dt = \infty$. Then the best invariant estimator is admissible iff it is admissible within U_1.*

Example 4.2. Suppose $h(t) = [(\sin t)(1 - \sin t)]^{-1}$ and the sample size $n = 2$. Then the best invariant estimator is admissible. The reader can verify it by referring to the proof that the best invariant estimator is admissible for the sample size $n = 2$ and for the case $h(t) = t^{-1}(1 - t)^{-1}$ in Yu (1987).

5. Q_d and the minimaxity of d_0

In this section we are going to investigate some methods for attacking the minimaxity of d_0, instead of as a by-product of its admissibility, since in most cases we know so far, d_0 is inadmissible.

Yu (1989*a*) introduced the notion $\tilde{Q}_d$ for given $d \in V$ (see (4.2) and (5.1)).

$$\tilde{Q}_d = \{x \in (0, 1): \lim_{t \uparrow x} d(x, t) = u_0, \text{ and } \lim_{t \downarrow x} d(x, t) = u_1\}\,, \tag{5.1}$$

where u_0 and u_1 are as (4.1). An interesting fact is that the set $\tilde{Q}_d$ plays an important role in justifying the admissibility of $d_0(t)$. We are going to investigate the relation between $\tilde{Q}_d$ and the minimaxity of d_0. It is not surprising to see that this leads to a proof of the minimaxity of the best invariant estimator for the sample size $n = 1$.

DEFINITION 5.1. Suppose $r \in (0, 1]$ and I is a positive-Lebesgue-measure subset of R. A distribution function $F(t)$ is said to give mass r uniformly to I if $F(t) - F(a) = \int_a^t [1(x \in I) r / mI] dx$ for $t \in (a, b)$, where $a = \text{ess.inf } I$ and $b = \text{ess.sup } I$. If $r = 1$ in the above formula F is the uniform distribution on I.

THEOREM 5.1. *Suppose $h(t)$ is nonnegative, $t(1 - t)h(t)$ is integrable and the sample size $n = 1$. If $m\tilde{Q}_d > 0$, then*

$$\sup_{F \in \Theta} R(F, d) \geq R(F, d_0)\,.$$

PROOF. Without loss of generality, assume that $h(t)$ is integrable. Since $m\tilde{Q}_d > 0$, $\forall\, \delta > 0$, $\exists$ a subset B of $\tilde{Q}_d$, satisfying:

(1) as $t \uparrow x$, $d(x, t)$ converges to u_0 uniformly for $x \in B$;
(2) as $t \downarrow x$, $d(x, t)$ converges to u_1 uniformly for $x \in B$;
(3) $mB > 0$ and $\sup B - \inf B = \eta$;
(4) $|d(x, t) - u_0| < \delta$, $\forall\, x \in B$ and $0 < x - t < \eta$;
(5) $|d(x, t) - u_1| < \delta$, $\forall\, x \in B$ and $0 < t - x < \eta$.

Let F be the uniform distribution on B, then

$$|R(F, d) - R(F, d_0)| < 2\delta \int_0^1 h(t) dt\,.$$

Since δ is arbitrary, the result follows. □

COROLLARY 5.1. *Under the assumptions of Theorem* 5.1, *the best invariant estimator is minimax for the sample size* $n = 1$.

PROOF. By simulating the proof of Theorem 3.4 in Yu (1989*a*), we have that $d \in V$ implies $m\tilde{Q}_d = 1$. So the minimaxity follows from Theorem 5.1. □

Since the minimaxity of the best invariant estimator remains open for general sample size $n > 1$, it is natural to consider some extension of $\tilde{Q}_d$ to the sample size $n > 1$ case. For this purpose, we introduce a new notion Q_d as follows,

$$Q_d = \{x \in (0,1): \forall\, \varepsilon > 0,\ \exists\, \delta > 0 \ni |d(y,t) - d_0(y,t)| < \varepsilon \text{ for almost all } (y,t) \in N((x,\dots,x),\delta)\}, \tag{5.1'}$$

where y is an n-dimension vector, and consider the minimaxity within the family of estimators d satisfying

$$d(\boldsymbol{x},t) \text{ is nonincreasing in } x_i \text{ given } t \text{ and other } x_j\text{'s fixed}, \quad i = 1,\dots,n. \tag{5.2}$$

As we can see that this notion can be used in the higher dimension. The connection of these two sets for the case $n = 1$ can be seen from the following theorem.

THEOREM 5.2. *Suppose the sample size* $n = 1$. *Given an estimator* d, *then* $mQ_d = 1$ *implies* $m\tilde{Q}_d = 1$.

The proof of Theorem 5.2 appears in the Appendix.

The importance of the notion of Q_d for given estimator d can be seen from the following theorem which is also proved in the Appendix.

THEOREM 5.3. *Suppose* $h(t)$ *is nonnegative*, $t(1-t)h(t)$ *is integrable. For the sample size* $n > 0$, *if for any* $d \in V$ *we have that* Q_d *is not empty, then* d_0 *is minimax.*

Now we have a sufficient condition for checking the minimaxity of the best invariant estimator d_0. Of course, $mQ_d = 1$ if $d \in V$ satisfies the condition. As we know, if $h(t) = t^{-1}(1-t)^{-1}$, $d \in V$ and the sample size is 1 or 2, then $mQ_d = 1$. We can also show that the condition that $d \in V$ implies Q_d is not empty is true for general $h(t)$ when $n = 1$, if d satisfies (5.2) (thus we have a sufficient condition for checking the minimaxity of d_0 within the family of estimators satisfying (5.2)). We need some lemmas before we prove it. The proofs of these lemmas are given in the Appendix.

LEMMA 5.1. *Suppose d_1 has finite risk for all $F \in \Theta$ and $d \in V$. $\forall\, \varepsilon > 0$, $\exists\, \delta > 0$ such that*

$$|R(F,d) - R(F,d_1)| < \varepsilon \qquad \text{if} \quad |d(\mathbf{x},t) - d_1(\mathbf{x},t)| < \delta \quad a.e. \quad (dF)^{n+1}.$$

Remark 5.1. Under assumption (5.2): if $\exists\, H \subset R^2$ such that

$$|d(x,t) - u| < \delta \;\forall\, (x,t) \in H \quad \text{and} \quad m^2 H > 0,$$

then there are two points (b,p) and (a,q) in H such that $a < b$ and $p < q$; furthermore,

$$|d(x,t) - u| < \delta, \quad \forall\, (x,t) \in (a,b) \times (p,q),$$

since $u - \delta < d(b,p) \le d(x,t) \le d(a,q) < u + \delta$.

LEMMA 5.2. *Suppose d satisfies* (5.2). *Given n and $x \in (0,1)$, $\exists\, 0 \le l \le h \le 2^n$ such that $\forall\, \delta > 0$ $\exists$ closed intervals I_1 and I_2 in $(0,1)$ of positive measure satisfying*:

(a) $I_1 \cup I_2 \subset N(x,\delta)$, $\inf I_2 \ge \sup I_1$;

(b) $d(x,t) \in [l/2^n, (l+1)/2^n)$ *if* $(x,t) \in I_2 \times I_1$;

(c) $d(x,t) \in [h/2^n, (h+1)/2^n)$ *if* $(x,t) \in I_1 \times I_2$.

So by Lemma 5.2, $\exists\, h_1$ = maximum of all possible h as above;

$\exists\, h_2$ = minimum of all possible h as above;

$\exists\, l_3$ = maximum of all possible l as above;

$\exists\, l_4$ = minimum of all possible l as above.

In other words, for the x given above, $\exists$ some $\boldsymbol{h} = (h_1, \dots, h_4)$, and $\boldsymbol{l} = (l_1, \dots, l_4)$, satisfying:

(*) $\forall\, \delta > 0$ $\exists$ closed intervals I_{j1} and I_{j2} in the neighborhood $N(x,\delta)$ of x with radius δ, satisfying

(a) $mI_{ji} > 0$, and $\inf I_{j2} \ge \sup I_{j1}$, $i = 1, 2$, $j = 1, \dots, 4$;

(b) $d(x,t) \in [h_j/2^n, (h_j+1)/2^n)$, if $(x,t) \in I_{j1} \times I_{j2}$;

(c) $d(x,t) \in [l_j/2^n, (l_j+1)/2^n)$, if $(x,t) \in I_{j2} \times I_{j1}$, $j = 1, \dots, 4$; and

(**) $\exists\, \delta_0 > 0$ such that

$$m^2 \left\{ (y,t){:}\; y, t \in N(x,\delta_0),\ d(y,t) \notin \left[\frac{h_2}{2^n}, \frac{h_1}{2^n} \right) \text{if } y < t, \right.$$

$$\text{or } d(y,t) \notin \left[\frac{l_4}{2^n}, \frac{l_3}{2^n}\right) \text{ if } y > t \Bigg\} = 0 .$$

Let

$$B_{nlh} = \{x \in (0,1): x \text{ satisfies } (*) \text{ and } (**)\} . \tag{5.3}$$

The sets defined in (5.3) will be used in the proof of Theorem 5.4.

THEOREM 5.4. *Suppose the sample size* $n = 1$. *Let* $d(x,t) \in V$. *Let* Q_d *be as* (5.1′). *If* d *satisfies* (5.2), *then* Q_d *is dense in* $(0,1)$.

PROOF. Fix n, and $[a_0, b_0] \subset (0,1)$. Since $[a_0, b_0]$ is a compact metric space and $[a_0, b_0] = \bigcup_{l,h}(B_{nlh} \cap [a_0, b_0])$, by the Baire Category Theorem, there is an open interval (a_n, b_n) and $B_{nl_nh_n}$, where $\boldsymbol{l}_n = (l_{1n}, \ldots, l_{4n})$ and $\boldsymbol{h}_n = (h_{1n}, \ldots, h_{4n})$, such that $B_{nl_nh_n} \cap [a_0, b_0]$ is dense in (a_n, b_n). Given $x \in [a_n, b_n]$, there is a sequence $\{x_m\}$ of $B_{nl_nh_n}$ such that $x_m \to x$. So $\forall\, \delta > 0\ \exists\, m$ such that $x_m \in N(x, \delta)$. Since $x_m \in B_{nl_nh_n}$, $\exists$ measurable closed intervals I_{j1} and I_{j2} in the neighborhood $N(x_m, \delta_m) \subset N(x, \delta)$ of x_m with radius δ_m, satisfying

(a) $mI_{ji} > 0$, and $\inf I_{j2} \geq \sup I_{j1}$, $i = 1, 2$, $j = 1, \ldots, 4$;

(b) $d(x,t) \in [h_{jn}/2^n, (h_{jn}+1)/2^n)$, if $(x,t) \in I_{j1} \times I_{j2}$;

(c) $d(x,t) \in [l_{jn}/2^n, (l_{jn}+1)/2^n)$, if $(x,t) \in I_{j2} \times I_{j1}$, $j = 1, \ldots, 4$.

This means that any x in (a_n, b_n) satisfies $(*)$ for the given $\boldsymbol{l}_n$ and $\boldsymbol{h}_n$. On the other hand, $\exists\, z$ in (a_n, b_n) satisfying $(**)$, i.e.,

$(***)$ $\exists\, 2\delta_0 > 0$ such that

$$m^2 \left\{ (y,t): y, t \in N(x, 2\delta_0),\ d(y,t) \notin \left[\frac{h_{2n}}{2^n}, \frac{h_{1n}}{2^n}\right) \text{ if } y < t, \right.$$

$$\left. \text{or } d(y,t) \notin \left[\frac{l_{4n}}{2^n}, \frac{l_{3n}}{2^n}\right) \text{ if } y > t \right\} = 0 ,$$

and $(z - 2\delta_0, z + 2\delta_0) \subset (a_n, b_n)$. So any x in $[z - \delta_0, z + \delta_0]$ satisfies $(*)$ and $(**)$ for $\boldsymbol{l}_n$ and $\boldsymbol{h}_n$ given above. That is $[z - \delta_0, z + \delta_0] \subset B_{nl_nh_n}$.

Without loss of generality, we assume $B_{nl_nh_n}$ contains $[a_n, b_n]$ (note $b_n - a_n > 0$). For the same reason, there is $[a_{n+1}, b_{n+1}] \subset [a_n, b_n]$ satisfying

$$\begin{gathered} [a_{n+1}, b_{n+1}] \subset B_{n+1\,l_{n+1}h_{n+1}} , \\ \left[\frac{h_{jn+1}}{2^{n+1}}, \frac{h_{jn+1}+1}{2^{n+1}}\right) \subset \left[\frac{h_{jn}}{2^n}, \frac{h_{jn}+1}{2^n}\right) \quad \text{and} \\ \left[\frac{l_{jn+1}}{2^{n+1}}, \frac{l_{jn+1}+1}{2^{n+1}}\right) \subset \left[\frac{l_{jn}}{2^n}, \frac{l_{jn}+1}{2^n}\right). \end{gathered} \tag{5.4}$$

So there are $\boldsymbol{w} = (w_1, \ldots, w_4)$ and $\boldsymbol{\upsilon} = (\upsilon_1, \ldots, \upsilon_4)$ such that $\boldsymbol{l}_n/2^n \to \boldsymbol{w}$ and $\boldsymbol{h}_n/2^n \to \boldsymbol{\upsilon}$ as $n \to \infty$.

If $w_i = u_0$ and $\upsilon_i = u_1$, $i = 1, \ldots, 4$, we are done by (**), (5.3) and (5.4). Otherwise one of the following four cases would be true:

$$(1)\ \upsilon_1 > u_1; \quad (2)\ \upsilon_2 < u_1; \quad (3)\ w_3 > u_0; \quad (4)\ w_4 < u_0.$$

Suppose case (1) is true. Consider an estimator

$$d_1(x,t) = \begin{cases} \upsilon_1 & \text{if} \quad t \geq x, \\ w_1 & \text{otherwise}. \end{cases}$$

Note $d_1 \in U_0$, so

$$R(F, d_1) - R(F, d_0) = 3\varepsilon > 0\,. \tag{5.5}$$

It can be shown that d_1 has finite risk. Note $d \in V$ and $R(F, d) < \infty$, so by Lemma 5.1, given ε as above, $\exists\, \delta > 0$ such that

$$|R(F, d_1) - R(F, d)| < \varepsilon \quad \text{if} \quad |d(y,t) - d_1(y,t)| < \delta \quad \text{a.e.} \quad (dF)^2\,. \tag{5.6}$$

There is n large enough such that $1/n < \delta$. For such n, $\exists$ closed intervals I_{11} and $I_{12} \subset [a_n, b_n]$ satisfying

(a) $mI_{1i} > 0$, $i = 1, 2$, and $\inf I_{12} \geq \sup I_{11}$;

(b) $d(y,t) \in [h_{1n}/2^n, (h_{1n}+1)/2^n)$ if $(y,t) \in I_{11} \times I_{12}$;

(c) $d(y,t) \in [l_{1n}/2^n, (l_{1n}+1)/2^n)$ if $(y,t) \in I_{12} \times I_{11}$.

Thus $|d(y,t) - d_1(y,t)| < 1/n$ if $(y,t) \in \left(\bigcup_{j=1}^{2} I_{1j}\right)^2 \setminus \bigcup_{j=1}^{2} (I_{1j})^2$.

It can be shown that there are closed intervals $I_{2,2i}$ and $I_{2,2i-1} \subset [a_n, b_n] \cap I_{1j}$ satisfying

(a) $mI_{2j} > 0$, $j = 1, 2, 3, 4$, and $\inf I_{2,2i} \geq \sup I_{2,2i-1}$;

(b) $d(y,t) \in [h_{1n}/2^n, (h_{1n}+1)/2^n)$ if $(y,t) \in I_{2,2i-1} \times I_{2,2i}$;

(c) $d(y,t) \in [l_{1n}/2^n, (l_{1n}+1)/2^n)$ if $(y,t) \in I_{2,2i} \times I_{2,2i-1}$.

Thus, $|d(y,t) - d_1(y,t)| < 1/n$ if $(y,t) \in \left(\bigcup_{j=1}^{2^2} I_{2j}\right)^2 \setminus \bigcup_{j=1}^{2^2} (I_{2j})^2$ by (b) and (c) in step 1 and 2.

$\vdots$

Similarly, we can construct a sequence of closed intervals $\{I_{kj}\}$. Let F_k give mass 2^{-k} uniformly to I_{kj}, $j = 1, \ldots, 2^k$ (see Definition 5.1). By the construction we have

$$
\text{(5.7)}\quad
\begin{aligned}
&(1)\quad F_k(t)=\int_0^t\left[\sum_{j=1}^{2^k}1(x\in I_{kj})/2^k mI_{kj}\right]dx\,;\\
&(2)\quad dF_k(t)dF_k(x)\quad\text{has support only on}\quad\left(\bigcup_{j=1}^{2^k}I_{kj}\right)^2;\\
&(3)\quad |d(y,t)-d_1(y,t)|<1/n\quad\text{if}\quad (y,t)\in\left(\bigcup_{j=1}^{2^k}I_{kj}\right)^2\Big\backslash\left(\bigcup_{j=1}^{2^k}I_{kj}^2\right).
\end{aligned}
$$

By the continuity of the integration w.r.t. $(dF)^2$, $\exists\ \delta_2>0$ such that

$$
\text{(5.8)}\quad \iint_B[(F(t)-d(x,t))^2-(F(t)-d_1(x,t))^2]h(F(t))dF(t)dF(x)<\varepsilon\,,
$$

if $\iint_B dF(t)dF(x)<\delta_2$.

Let $B=\bigcup_{i=1}^{2^k}I_{ki}^2$, then there is k large enough such that $\iint_B dF_k(t)dF_k(x)<\delta_2$. Since $|d(y,t)-d_1(y,t)|<1/n$ a.e. $(dF_k)^2$ in B^c by (1), (2) and (3) of (5.7), where B^c is the complement of set B, by (5.6) and (5.8)

$$
\begin{aligned}
\text{(5.9)}\qquad &|R(F_k,d_1)-R(F_k,d)|\\
&\le\left(\iint_B+\iint_{B^c}\right)[(F_k(t)-d(x,t))^2-(F_k(t)-d_1(x,t))^2]\\
&\quad\cdot h(F_k(t))dF_k(t)dF_k(x)\le\varepsilon+\varepsilon\,.
\end{aligned}
$$

By (5.5) and (5.9), we have

$$
\begin{aligned}
R(F_k,d)-R(F_k,d_0)&=R(F_k,d_1)-R(F_k,d_0)+R(F_k,d)-R(F_k,d_1)\\
&\ge 3\varepsilon-2\varepsilon=\varepsilon>0\,.
\end{aligned}
$$

Contradiction: This means case (1) is not true. Similarly we can show that cases (2) through (4) are not true. □

Acknowledgements

The author is greatly indebted to his advisor Professor T. S. Ferguson for all his help and suggestions in writing this paper. Thanks also go to the referee for his (or her) invaluable suggestions and comments.

Appendix

PROOF OF THEOREM 5.2. Suppose that given an estimator d, $mQ_d=1$ but $m\tilde{Q}_d<1$. Then $m\{(0,1)\backslash\tilde{Q}_d\}>0$. So

$$mQ_d \cap \{(0,1)\backslash \tilde{Q}_d\} > 0 .$$

There is a closed subset $G \subset Q_d \cap \{(0,1)\backslash \tilde{Q}_d\}$ of positive measure such that $d(x,t)$ converges uniformly on G as $t \uparrow x$ and as $t \downarrow x$. Given x in G, $\forall\, n\ \exists\ \delta_x$ such that $|d(y,t) - d_0(y,t)| < 1/n$ for almost all $(y,t) \in N((x,x),\delta_x)$ since $x \in Q_d$.

So $\{N(x,\delta_x): x \in G\}$ is a cover of G. Thus, there is a finite cover $N(x_i,\delta_{x_i})$, $i = 1,\ldots,k$. In other words, $\exists\ \delta > 0$ such that $|d(y,t) - d_0(y,t)| < 1/n$ for almost all (y,t) in $\{(y,t): y \in G$ and $|t-y| < \delta\}$. This implies that $\exists$ closed subset $G_n \subset G$ of positive measure such that $|d(x,t) - d_0(x,t)| < 1/n$ for all $x \in G_n$ and $|t-x| < \delta$. Furthermore, we can assume $G_{n+1} \subset G_n$. So there is $x \in \bigcap_n G_n$ satisfying

$$\lim_{t \uparrow x} d(x,t) = u_0 \quad \text{and} \quad \lim_{t \downarrow x} d(x,t) = u_1 .$$

This contradicts that $G \cap \tilde{Q}_d$ is empty. □

PROOF OF THEOREM 5.3. Since d_0 has constant risk, it is not hard to see

$$\inf_d \sup_F R(F,d) = \inf_{d \in V} \sup_F R(F,d) .$$

We consider three cases: (1) $\int_0^1 th(t)dt = \infty$; (2) $\int_0^1 (1-t)h(t)dt = \infty$; (3) $\int_0^1 h(t)dt < \infty$.

Suppose (2) is true but $\int_0^1 th(t)dt < \infty$. Let $d \in V$, and note (2) is true, so $d(Y,t) = 0$ a.s. in $\{(Y,t): Y_1 > t\}$. By assumption Q_d is not empty. So $\exists\, x$ in Q_d. Thus $\forall\, \varepsilon > 0\ \exists\ \delta > 0$ such that for almost all $(y,t) \in \{(y,t): \|(y,t) - (x,\ldots,x)\| < \delta\}$, where $\|(x_1,\ldots,x_{n+1})\| = \sup_i |x_i|$, we have

$$|d(y,t) - d_0(y,t)| < \varepsilon \Big/ \sum_{i=1}^n 2\binom{n}{i}\int_0^1 t^i(1-t)^{n-i}h(t)dt .$$

Let G be the uniform distribution on $(x-\delta, x+\delta)$, then

$$\begin{aligned}|R(G,d) - R(G,d_0)| &\le E\int_{Y_1}^1 |(G(t)-d(t))^2 - (G(t)-d_0(t))^2| h(G(t))dG(t) \\ &\le 2E\int_{Y_1}^1 h(G(t))dG(t)\varepsilon \Big/ \sum_{i=1}^n 2\binom{n}{i}\int_0^1 t^i(1-t)^{n-i}h(t)dt \\ &= \varepsilon .\end{aligned}$$

Now $\sup_F R(F,d) \geq R(G,d) \geq R(G,d_0) - \varepsilon$ for any ε. Note that d, in fact, is arbitrary. This implies $R(G,d_0) = \inf_{d \in V} \sup_F R(F,d)$. This means d_0 is minimax. Similarly, we can get the same conclusion for the case that (1) and (2) are true and the case that (1) or (3) is true. □

PROOF OF LEMMA 5.1. $\int th(t)dt = \infty$ $\left(\text{or} \int (1-t)h(t)dt = \infty\right)$, and d and d_1 have finite risk for all $F \in \Theta$ imply that $d = d_1$ for $t < Y_1$ (or $t > Y_n$). So we assume, without loss of generality, that $\int h(t)dt < \infty$.

$\forall\, \varepsilon > 0$ let $\delta = \varepsilon / 2 \int h(t)dt$, if $|d - d_1| < \delta$ a.e. $(dF)^{n+1}$, then

$$
\begin{aligned}
&|R(F,d) - R(F,d_1)| \\
&\quad \leq E\int |(F(t) - d(Y,t))^2 - (F(t) - d_1(Y,t))^2|\, h(F(t))dF(t) \\
&\quad \leq E\int 2\delta h(F(t))dF(t) \\
&\quad \leq \varepsilon\,.
\end{aligned}
$$
□

PROOF OF LEMMA 5.2. By (5.2),

$$d(y - 1/2k, y + 1/2k) \leq d(x,t) \leq d(y - 1/k, y + 1/k)\,,$$

if $(x,t) \in I_1 \times I_2 = [y - 1/k, y - 1/2k] \times [y + 1/2k, y + 1/k]$, and

$$d(y + 1/k, y - 1/k) \leq d(x,t) \leq d(y + 1/2k, y - 1/2k)\,,$$

if $(x,t) \in I_2 \times I_1$. Also there are a and b such that

$$d(y + 1/k, y - 1/k) \uparrow a \quad \text{and} \quad d(y - 1/k, y + 1/k) \downarrow b\,,$$

as k tends to $+\infty$. Without loss of generality, we can assume that $a \in (l/2^n, (l+1)/2^n]$ and $b \in [h/2^n, (h+1)/2^n)$. So there is k_0 such that

$$
\begin{aligned}
&d(y + 1/k_0, y - 1/k_0) > l/2^n \quad \text{and} \\
&d(y - 1/k_0, y + 1/k_0) < (h+1)/2^n\,.
\end{aligned}
$$

We can check that $\delta = 1/k_0$, I_1 and I_2 $(k > k_0)$ satisfy condition (a), (b) and (c). □

REFERENCES

Aggarwal, O. P. (1955). Some minimax invariant procedures 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.

Cohen, M. P. and Kuo, L. (1985). The admissibility of the empirical distribution function, *Ann. Statist.*, **13**, 262-271.

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.

Phadia, E. G. (1973). Minimax estimator of a cumulative distribution function, *Ann. Statist.*, **1**, 1149-1157.

Read, R. R. (1972). The asymptotic inadmissibility of the sample distribution function, *Ann. Math. Statist.*, **43**, 89-95.

Yu, Q. (1986). Admissibility of the best invariant estimator of a distribution function, Ph. D. Thesis, University of California, Los Angeles.

Yu, Q. (1987). Admissibility of the empirical distribution function in the invariant problem (to appear in *Statist. Decisions*).

Yu, Q. (1989*a*). Admissibility of the best invariant estimator of a distribution function, *Statist. Decisions*, 7, 1-14.

Yu, Q. (1989*b*). Inadmissibility of the empirical distribution function in continuous invariant problems (to appear in *Ann. Statist.*).