On The Elementary Theorems Of Decision Theory

By

Narasinga Rao Chaganty

Old Dominion University

and

Florida State University

June, 1986

FSU Technical Report No. M732
USARO Technical Report No. D-91

Department of Statistics
Florida State University
Tallahassee, FL 32306-3033

Research partially supported by the U.S. Army Research Office Grant number DAAL03-86-K-0094. The United States Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon.

Keywords and Phrases: Admissible rule, Bayes rule, Minimax principle, Complete class.
Abstract. Consider a statistical decision problem in which nature has a finite number of states. The elementary theorems of decision theory, namely the Minimax theorem, the Complete class theorem, and theorems on the structure of admissible rules, are proved in most texts under the assumptions that the risk set is closed from below and bounded from below. The condition that the risk set is bounded from below is sufficient for the existence of the lower boundary points; however, that it is not necessary can be seen from simple examples. The purpose of this paper is to extend the elementary theorems of decision theory to include the case in which the risk set is not bounded from below and the set of lower boundary points is nonempty. We also show that if the risk set is bounded from above then it is necessary for the risk set to be bounded from below for the set of lower boundary points to be nonempty. We present examples to illustrate our theorems.

1. Introduction and Definitions. A statistical decision problem consists of a triplet \((\Omega, D, R)\) where \(\Omega\) is the set of all possible states of nature, also known as the parameter space. The set \(D\) is the class of all randomized decision rules available to the statistician. When the statistician chooses the randomized decision rule \(\delta \in D\) and the state of nature is \(\theta \in \Omega\), the risk of the statistician is given by the real valued function \(R(\theta, \delta)\). For a complete discussion of randomized decision rules, risk functions, the reader is referred to Ferguson (1967) and Berger (1985). Throughout this paper we shall assume that the parameter space \(\Omega\) is a finite set, say \(\{\theta_1, \theta_2, \ldots, \theta_k\}\), where \(k \geq 1\) is fixed.
Given any decision rule \( \delta \in D \), the risk vector of \( \delta \) is defined as the vector \( x = (R(\theta_1, \delta), \ldots, R(\theta_k, \delta))^t \). The collection

\[
S = \left\{ x = (R(\theta_1, \delta), \ldots, R(\theta_k, \delta))^t : \delta \in D \right\}
\]

of all risk vectors is known as the risk set. It is well known that \( S \) is a convex subset of \( \mathbb{R}^k \); that is, if \( x, y \in S \) and \( 0 \leq \lambda \leq 1 \) then \( \lambda x + (1-\lambda)y \) is also in \( S \). We adopt the following definitions from Ferguson (1967).

**Definition 1.1.** The subset \( S \) is said to be bounded from below if there exists a finite positive number \( A \), such that for every \( x = (x_1, \ldots, x_k)^t \in S \)

\[
x_i > -A \quad \text{for } i = 1, \ldots, k.
\]

We say that \( S \) is bounded from above if for every \( x = (x_1, \ldots, x_k)^t \in S \),

\[
x_i < A \quad \text{for } i = 1, \ldots, k.
\]

The set \( S \) is said to be bounded if it is bounded from above and below.

**Definition 1.2.** Let \( x \) be a point of \( \mathbb{R}^k \). The lower quantant at \( x \), denoted by \( Q_x \), is defined as the set

\[
Q_x = \{ y = (y_1, \ldots, y_k)^t : y_j \leq x_j, \ j = 1, \ldots, k \}.
\]

In the following definitions \( \overline{S} \) denotes the closure of the set \( S \), that is, \( \overline{S} \) is the smallest closed set containing \( S \).
Definition 1.3. A point $x$ is said to be a lower boundary point of a convex set $S \subseteq \mathbb{R}^k$ if $0 \in \overline{S} = \{x\}$. The set of lower boundary points of a convex set $S$ is denoted by $\text{LB}(S)$.

The following lemma can be found in Ferguson (1967).

Lemma 1.4. Let $S$ be a convex subset of $\mathbb{R}^k$. Then $\overline{S}$ is convex and the set of lower boundary points of $S$ and $\overline{S}$ is the same, that is, $\text{LB}(S) = \text{LB}(\overline{S})$.

Definition 1.5. A convex set $S \subseteq \mathbb{R}^k$ is said to be closed from below if $\text{LB}(S) \subseteq S$. We say that $S$ is closed if $S = \overline{S}$.

Let $x = (x_1, \ldots, x_k)^t$ and $y = (y_1, \ldots, y_k)^t$ be two elements of $\mathbb{R}^k$. We write $x \leq y$ if $x_i \leq y_i$ for $i = 1, \ldots, k$, and $x < y$ if $x \leq y$ and $x_j < y_j$ for at least one $j$. The relations $\leq$, $<$ induce a partial ordering of the elements of $\mathbb{R}^k$. The next few definitions are concerned with the class $D$ of randomized decision rules available to the statistician.

Definition 1.6. Let $\delta_1$ and $\delta_2$ be two decision rules in $D$ with risk vectors $x_1$ and $x_2$ respectively. We say that $\delta_1$ is as good as the rule $\delta_2$ if $x_1 \leq x_2$. The rule $\delta_1$ is said to be better than $\delta_2$ if $x_1 < x_2$ and $\delta_1$ is said to be equivalent to $\delta_2$ if $x_1 = x_2$.

Definition 1.7. A rule $\delta$ is said to be admissible if there exists no rule that is better than $\delta$. A rule $\delta$ is said to be inadmissible if it is not admissible.
It is clear that inadmissible rules are undesirable. Thus in looking for a best rule, the statistician usually restricts his search to the class of admissible rules, provided the class is nonempty.

**Definition 1.8.** A class $C$ of decision rules, $C \subseteq D$, is said to be complete, if given any decision rule $\delta \in D$, not in $C$, there exists a rule $\delta_0 \in C$ that is better than $\delta$. The class $C$ of decision rules is said to be minimal complete if $C$ is complete and if no proper subclass of $C$ is complete.

We note that a complete class $C$ always exists for a statistical decision problem (take $C=D$), but a minimal complete class need not exist. However, if a minimal complete class exists then it consists exactly of the admissible rules. This result and several other results which demonstrate the relationship between minimal complete class and admissible rules can be found in Ferguson (1967) (also see Berger (1985)). We state below one important theorem that is relevant to the main theorems of this paper.

**Theorem 1.9.** If $C$ is the class of admissible rules and $C$ is a complete class then $C$ is minimal complete.

We now proceed and develop the notion of Bayes rule and Bayes risk. Any probability distribution $\pi = (\pi_1, \pi_2, \ldots, \pi_k)$ on the parameter space $\Omega$ is known as a prior distribution. We shall denote the class of all prior distributions by $P$. 
Given any prior distribution $\pi \in P$ and a decision rule $\delta \in D$, the Bayes risk of $\delta$ with respect to $\pi$ is defined as

$$(1.5) \quad BR(\pi, \delta) = E \left[ R(Z, \delta) \right]$$

where $Z$ is a random variable having distribution $\pi$. The Bayes principle requires the statistician to look for a rule which minimizes the Bayes risk. This rule if it exists is known as the Bayes rule. Formally we have

**Definition 1.10.** A decision rule $\delta_b$ is said to be Bayes with respect to the prior distribution $\pi \in P$ if

$$(1.6) \quad BR(\pi, \delta_b) = \inf_{\delta \in D} BR(\pi, \delta)$$

In searching for a best rule the statistician is also guided by another notion of best rule, known as the Minimax principle. A rule $\delta_1$ is preferred to a rule $\delta_2$ if

$$(1.7) \quad \sup_{\theta \in \Theta} R(\theta, \delta_1) < \sup_{\theta \in \Theta} R(\theta, \delta_2).$$

This relation "is preferred to" also defines a linear ordering on the class of decision rules $D$. If the statistician decides to adopt the Minimax principle then he should look for a smallest element in this ordering. A rule that is most preferred in this ordering is called a minimax decision rule. This discussion leads us to the following
**Definition 1.11.** A decision rule $\delta_0$ is said to be minimax if

\[ \sup_{\theta \in \Omega} R(\theta, \delta_0) = \inf_{\delta \in \Delta} \sup_{\theta \in \Omega} R(\theta, \delta). \]

We can also define similarly, a minimax rule for nature. This will be an element of the class $\mathcal{P}$. The minimax rule of the nature $\pi_f \in \mathcal{P}$ is also known as a least favorable distribution. The probability distribution $\pi_f \in \mathcal{P}$ satisfies the equality

\[ \inf_{\delta \in \Delta} BR(\pi, \delta) = \sup_{\pi \in \mathcal{P}} \inf_{\delta \in \Delta} BR(\pi, \delta). \]

2. **Main Theorems.** In this section we present the main results of this paper. The theorems presented in this section are analogous to the theorems in Chapter 2 of Ferguson (1967). However, Ferguson (1967) assumes that the risk set $S$ is bounded from below. We replace this condition by the weaker condition that the set of lower boundary points of $S$, $LB(S)$, is nonempty. We also present applications of our theorems which are not covered by the theorems of Ferguson (1967). The following lemma plays an important role in the proofs of our theorems.

**Lemma 2.1.** Let \( \{ y_{1n}, n \geq 1 \}, \ldots, \{ y_{kn}, n \geq 1 \} \) be $k$ sequences which are bounded above. Assume that there exists $j$, $1 \leq j \leq k$, such that $y_{jn}$ $\longrightarrow - \infty$ as $n \longrightarrow \infty$. Then there exists an $r$, $1 \leq r \leq k$, and a subsequence \( \{ m \} \) such that

\[ \left[ \begin{array}{c} y_{1m} \\ \vdots \\ y_{km} \end{array} \right] \longrightarrow M_j \]

for all $i$, $1 \leq i \leq k$, where $0 \leq M_j < \infty$ and $y_{rn} \longrightarrow - \infty$ as $m \longrightarrow \infty$. 

7
Proof. We shall prove the lemma by the method of induction. The lemma trivially holds for \( k = 1 \). Assume that the lemma is true for \( k = (I-1) \), that is, there exists an \( r \), \( 1 \leq r \leq (I-1) \), and a subsequence \( \{m\} \) such that (2.1) holds for \( 1 \leq i \leq (I-1) \). Let \( \{y_{1m}, m \geq 1\} \) be another sequence.

Note that

\[
0 \leq \liminf \left( \frac{y_{1m}}{y_{rm}} \right) \leq \infty,
\]

since \( \{y_{1m}\} \) is bounded above and \( y_{rm} \to \infty \) as \( m \to \infty \). We consider the following two cases.

Case 1: \( 0 \leq \liminf \left( \frac{y_{1m}}{y_{rm}} \right) < \infty \). In this case we can find a subsequence \( \{m'\} \) of \( \{m\} \) such that

\[
(2.2) \quad \lim \left( \frac{y_{1m'}}{y_{rm'}} \right) = M_1,
\]

where \( 0 \leq M_1 < \infty \). The subsequence \( \{m'\} \) satisfies (2.1) for all \( 1 \leq i \leq I \).

Case 2: \( \liminf \left( \frac{y_{1m}}{y_{rm}} \right) = \infty \). In this case we can find a subsequence \( \{m'\} \) of \( \{m\} \) such that

\[
(2.3) \quad \lim \left( \frac{y_{rm'}}{y_{1m'}} \right) = 0.
\]

Combining (2.3) with the induction hypothesis we get that for \( 1 \leq i \leq (I-1) \),
\[ (2.4) \quad \lim \left( \frac{y_{1m}}{y_{1m}'} \right) = \lim \left( \frac{y_{2m}}{y_{2m}'} \right) \lim \left( \frac{y_{rm}}{y_{rm}'} \right) = 0 = M_i \quad \text{(say).} \]

Therefore the subsequence \( \{ m' \} \) and \( r = 1 \) satisfies (2.1) for \( 1 \leq i \leq l \).

This completes the proof of Lemma 2.1.

As an application of Lemma 2.1 we have the following result.

**Lemma 2.2.** Let \( S \) be a convex subset of \( \mathbb{R}^k \). Assume that \( S \) is bounded above. If the set of lower boundary points \( \text{LB}(S) \) is nonempty then \( S \) is bounded from below.

**Proof.** Suppose \( S \) is not bounded from below. Then there exists a sequence \( y_n = (y_{1n}, \ldots, y_{kn})^t \in S \) such that \( y_{jn} \to -\infty \) for some \( j \), \( 1 \leq j \leq k \) as \( n \to \infty \). Since the \( y_{jn} \)'s are bounded above, using Lemma 2.1 we can find an \( 1 \leq r \leq k \) and a subsequence \( \{ m \} \) such that \( y_{rm} \to -\infty \) and

\[ (2.5) \quad \left( \frac{y_{1m}}{y_{rm}} \right) \to M_i \quad \text{as} \quad m \to \infty, \]

where \( 0 \leq M_i < \infty \), for \( 1 \leq i \leq k \). Let \( z = (z_1, \ldots, z_k)^t \) be an element of \( \text{LB}(S) \). Let \( \epsilon > 0 \) be given. Define

\[ (2.6) \quad \lambda_m = \left\lfloor \left( z_r - \epsilon \right) - y_{rm} \right\rfloor \left/ \left[ \left( z_r - y_{rm} \right) \right] \right., \quad \text{for} \quad m \geq 1. \]

Note that \( 0 < \lambda_m < 1 \) for sufficiently large \( m \). Since \( \bar{S} \) is convex, \( \lambda_m z + (1-\lambda_m)y_m \in \bar{S} \). Using (2.5) and (2.6) we can easily verify that

\[ (2.7) \quad \lambda_m \to 1 \]
(2.8) \[ \lambda_a z + (1-\lambda_a) y_a \longrightarrow z - \varepsilon M \]

as \( m \longrightarrow \infty \), where \( M = (M_1, \ldots, M_k)^t \). Since \( M_\infty = 1 \), the limit point \( z - \varepsilon M \)

which belongs to \( \overline{S} \), is less than \( z \) and this contradicts the fact that \( z \in LB(S) \). The proof of Lemma 2.2 is now complete.

The above lemma shows that if a convex subset \( S \) of \( \mathbb{R}^k \) is bounded above, then either \( S \) is bounded below or \( LB(S) \) is empty. The following example shows that we cannot relax the hypothesis that \( S \) is bounded above in Lemma 2.2.

**Example 2.3.** Let \( k = 2 \) and \( S_2 = \{(x_1, x_2)^t : x_1 \geq -x_2\} \). In this example \( LB(S_2) = \{(x_1, x_2)^t : x_1 = -x_2\} \), but \( S_2 \) is neither bounded from above nor bounded from below.

The converse of Lemma 2.2 is well known and is stated in the Lemma 2.4. The proof of Lemma 2.4 can be found on page 69, Ferguson (1967).

**Lemma 2.4.** Let \( S \) be a convex subset of \( \mathbb{R}^k \). If \( S \) is bounded from below, then \( LB(S) \) is nonempty.

The next lemma is crucial to the proofs of our elementary theorems of decision theory. The proof follows easily from Lemma 2.2.
Lemma 2.5. Let $S$ be a convex subset of $\mathbb{R}^k$ with nonempty lower boundary points $\text{LB}(S)$. Let $x = (x_1, \ldots, x_k)^t \in \overline{S}$ and let $S_1 = Q_x \cap \overline{S}$. Then $S_1$ is bounded from below.

Proof. Suppose $S_1$ is not bounded from below. Then we can find a sequence $y_n = (y_{1n}, \ldots, y_{kn})^t \in S_1$ such that for some $j$, $y_{jn} \to -\infty$ as $n \to \infty$. Let $z = (z_1, \ldots, z_k)^t \in \text{LB}(S)$. Since $y_n$ is bounded from above, imitating the proof of Lemma 2.2 we can find $z' \in \overline{S}$ such that $z' < z$. This contradicts the fact that $z \in \text{LB}(S)$. The proof of Lemma 2.5 is complete.

Lemma 2.5 is best understood when $k=2$ and $S$ is a convex subset of $\mathbb{R}^2$. Let $(x_1, x_2)^t \in \overline{S}$ and let $S_1 = Q_{x} \cap \overline{S}$. Let $(z_1, z_2)^t \in \text{LB}(S)$. It is easy to verify that either $x_1 > z_1$ or $x_2 > z_2$. Let us assume that $x_2 > z_2$ and $x_1 \leq z_1$, the other cases can be handled similarly. If $S_1$ is not bounded from below then there exists a sequence $y_n = (y_{1n}, y_{2n})^t \in S_1$ such that $y_{1n} \to -\infty$ and $z_2 < y_{2n} \leq x_2$. Since $\overline{S}$ is convex, the line $L_n$ joining $z$ and $y_n$ is contained in $\overline{S}$. Letting $n$ converge to $\infty$, we can check that $L_n$ coincides with the line $y = z_2$ (see Figure 1) and therefore we can find points of $\overline{S}$ which are less than $z$. This is a contradiction because $z \in \text{LB}(S)$.
The next theorem is the first of our elementary theorems of decision theory. It is well known that if the risk set $S$ is bounded from below and closed from below then the risk vector of every admissible rule is contained in $LB(S)$. Theorem 2.6 shows that we can extend this result to the case where the risk set $S$, is not bounded below but is assumed to have nonempty lower boundary points.

**Theorem 2.6.** Assume that the set of lower boundary points, $LB(S)$, of the risk set $S$ is not empty and $S$ is closed from below. Then a rule $\delta \in D$ is admissible if and only if the risk vector $x = (x_1, \ldots, x_k)^t$ of $\delta$ is contained in $LB(S)$.

**Proof.** If $x = (x_1, \ldots, x_k)^t \in LB(S)$ then it follows trivially that the decision rule $\delta$ corresponding to $x$ is admissible. We shall prove the converse by the method of contradiction. Let $\delta$ be an admissible rule and assume that the risk vector $x$ of $\delta$ is not in $LB(S)$. Let $S_1 = Q_x N^S$. By Lemma 2.5, the set $S_1$ is bounded from below. It is easy to verify that $S_1$ is convex and closed. Therefore by Lemma 2.4, $LB(S_1)$ is not empty.

If $y \in LB(S_1)$ then

\begin{equation}
(2.9) \quad \langle y \rangle = Q_y N S_1 = Q_y N Q_x N^S = Q_y N^S.
\end{equation}

Thus $y \in LB(S)$ and since $S$ is closed from below, $y \in S$. Since $y < x$, $\delta$ is not admissible, which is a contradiction. This completes the proof of the theorem.