A GENERAL THEORY OF ASYMPTOTIC CONSISTENCY FOR SUBSET SELECTION WITH APPLICATIONS

BY JAN F. BJØRNSTAD

The University of Texas at Austin

The problem of selecting a random nonempty subset from \( k \) populations, characterized by \( \theta_1, \ldots, \theta_k \) with possible nuisance parameters \( \sigma \), is considered using a decision-theoretic approach. The concept of asymptotic consistency is defined as the property that the risk of a procedure at \( (\theta, \sigma) \) tends to the minimum loss at \( (\theta, \sigma) \). Necessary and sufficient conditions for both pointwise and uniform (on compact sets) consistency for permutation-invariant procedures are derived with general loss functions.

Various loss functions when the goal is to select populations with \( \theta_i \) close to max \( \theta_j \) are considered. Applications are made to normal populations. It is shown that Gupta's procedure is the only procedure in Seal's class that can be consistent. Other Bayes and admissible procedures are also considered.

1. Introduction. The multiple decision problem of selecting a random nonempty subset from \( k \) populations \( \pi_1, \ldots, \pi_k \) is considered. \( \pi_1, \ldots, \pi_k \) are characterized by \( \theta_1, \ldots, \theta_k \) respectively, where \( \theta_i \in \Theta \subset \mathbb{R} \) and the parameter-space of \( \theta = (\theta_1, \ldots, \theta_k) \) is \( \Omega \subset \Theta^k \). \( \hat{X}^n \) is an estimate of \( \theta_i \), based on \( n \) observations. We shall allow for the presence of nuisance-parameters, denoted by \( \sigma \) with parameter-space \( \Sigma \). \( \sigma \) is estimated by \( \hat{S}^n \in E \). The joint distribution function of \( (\hat{X}^n, \hat{S}^n) \) is denoted by \( F_{\hat{X},\hat{S}} \). Let now \( G \) be the group of permutations \( g \) on \( \{1, \ldots, k\} \). For \( x \in \mathbb{R}^k \), \( gx \) is defined by \( (gx)_i = x_{g^{-1}i} \). For any subset \( A \) of \( \mathbb{R}^k \), \( gA = \{gx : x \in A\} \). The probability model is assumed to be invariant under \( G \), i.e. (a) if \( (\hat{X}^n, \hat{S}^n) \) has cdf \( F_{\hat{X},\hat{S}} \), then \( (g\hat{X}^n, \hat{S}^n) \) has cdf \( F_{\hat{X},\hat{S}} \), and (b) \( g\Omega = \Omega \), \( \forall g \in G \).

The decision-space is \( \mathcal{A} = \{a \subset \{1, \ldots, k\}\} \), where the decision \( a \) is interpreted as selecting the populations \( \pi_i, i \in a \). For \( a \in \mathcal{A} \), \( ga = \{gi : i \in a\} \). The loss-function \( \zeta(\theta, a) \) is assumed to be permutation-invariant, i.e. \( \zeta(\theta, ga) = \zeta(g\theta, ga) \) for all \( g \in G \). It follows that the multiple decision problem is invariant under \( G \). Furthermore, \( -\infty < \zeta(\theta, a) < \infty \), \( \forall a \in \mathcal{A}, \forall (\theta, \sigma) \in \Omega \times \Sigma \). A subset selection procedure is given by:

\[
\delta_n(a \mid x, s) = \Pr[\text{decision } a \mid X^n = x, S^n = s].
\]

We shall consider the class of invariant procedures, \( \mathcal{D}_I \), where \( \delta_n \in \mathcal{D}_I \) if and only if \( \delta_n(ga \mid gx, s) = \delta_n(a \mid x, s) \) for all \( a \in \mathcal{A}, x \in \mathbb{R}^k, g \in G \). The risk-function of \( \delta_n \) is \( r_n(\theta, \sigma \mid \delta_n) = \sum_{a \in \mathcal{A}} \zeta_n(\theta, a)E_a \delta_n(a \mid X^n, S^n) \).

Received October 1982; revised March 1984.

1 This research was supported in part by the National Science Foundation Grant No. MCS-8303820.


Key words and phrases. Subset selection, asymptotic theory, decision-theory, invariance, consistency.

1058
The purpose of this paper is to develop a theory of asymptotic consistency for different loss functions in this multiple decision problem. To define the term consistency, let \( m_\sigma(\theta) = \min_{\theta \in \Theta} \ell_\sigma(\theta, a) \). If \( \ell_\sigma(\theta, a_0) = m_\sigma(\theta) \) then \( a_0 \) is a correct decision when \( (\theta, \sigma) \) is true. Obviously, \( r_n(\theta, \sigma | \delta_n) \geq m_\sigma(\theta), \forall (n, \theta, \sigma) \). All limits in this paper are as \( n \to \infty \).

**Definition 1.1.** The sequence of procedures \( \{\delta_n\} \) is consistent at \( (\theta, \sigma) \) if 
\[ r_n(\theta, \sigma | \delta_n) \to m_\sigma(\theta). \]

We say that \( \delta_n \) is pointwise consistent on \( \Omega \times \Sigma \) if \( \delta_n \) is consistent at each \( (\theta, \sigma) \in \Omega \times \Sigma \).

We shall also consider the concept of uniform consistency. We note that the metric on \( \Omega \times \Sigma \) is the usual Euclidean distance.

**Definition 1.2.** The sequence \( \{\delta_n\} \) is uniformly consistent if
\[ \sup_{K_1, \sigma \in K_2} |r_n(\theta, \sigma | \delta_n) - m_\sigma(\theta)| \to 0 \quad \text{for all compact sets } K_1, K_2 \text{ of } \Omega, \Sigma. \]

Consistency is a desirable property universally in all decision-problems. It simply states that the decision-procedure should take the correct decision as \( n \) tends to infinity. The theory for pointwise consistency will require only the following condition:
\[ E_{\theta, \sigma} |X^n_i - \theta_i| \to 0 \quad \text{for } i = 1, \ldots, k, \forall (\theta, \sigma) \in \Omega \times \Sigma. \]

Similarly, the theory for uniformly consistent procedures will require:
\[ \sup_{K_1 \times K_2} E_{\theta, \sigma} |X^n_i - \theta_i| \to 0 \quad \text{for } i = 1, \ldots, k \]
for all compact subsets \( K_1, K_2 \) of \( \Omega, \Sigma \).

In Section 2, necessary and sufficient conditions for pointwise and uniform consistency are derived for procedures in \( \mathcal{D}_l \) with respect to general loss-functions. Let \( \pi(l) \) correspond to \( \theta(l) \) where \( \theta(l) \leq \cdots \leq \theta(k) \). Section 3 considers different loss functions reflecting the goal to select populations close to \( \pi(k) \). These loss functions have been proposed by Chernoff and Yahav (1977), Bickel and Yahav (1977), Goel and Rubin (1977), Gupta and Hsu (1978) and Bjørnstad (1981). It is noted that some of these losses imply that the classical approach, started by Seal (1955) and Gupta (1956), of employing the so-called \( P^* \)-condition is not always appropriate.

To save space, the theory in Section 3 is applied only to the selection of means from normally distributed populations in Section 4. It is clear, however, that procedures for binomial, multinomial, multivariate normal and other selection problems can be checked for consistency in a similar way.

It is shown in Section 4, that among all the procedures in the class proposed by Seal (1955), only Gupta's procedure can be consistent. We also consider two classes of admissible procedures, derived by Bjørnstad (1981), and the Bayes procedures derived by Chernoff and Yahav (1977), Goel and Rubin (1977) and Gupta and Hsu (1978) for their respective loss functions and exchangeable normal priors.
2. Consistent invariant procedures. Our first aim is to develop necessary and sufficient conditions for pointwise consistency, with an invariant loss function, for procedures in \(D\). For \(x \in \mathbb{R}^k\), \(x^* = (x_{(1)}, \ldots, x_{(k)})\) where \(x_{(1)} \leq \cdots \leq x_{(k)}\). Since for any \(\delta_n \in D\), \(r_n(\theta, \sigma | \delta_n) = r_n(\theta^*, \sigma | \delta_n)\), we have that \(\delta_n\) is consistent at \((\theta, \sigma)\) if and only if \(\delta_n\) is consistent at \((\theta^*, \sigma)\). Let \(Y^n = (Y^n_1, \ldots, Y^n_k) = (X^n_{(1)}, \ldots, X^n_{(k)})\). We need the following result.

**Lemma 2.1.** Assume (1.1) holds, and \(\theta_1 \leq \cdots \leq \theta_k\).

(a) If \(\theta_i \neq \theta_j\) then \(P_{\theta, \sigma}(X^n_i = Y^n_i) \to 0\).

(b) If \(g\theta \neq \theta\) then \(P_{\theta, \sigma}(gX^n = Y^n) \to 0\).

**Proof.** (a): Let first \(i > j\) such that \(\theta_i > \theta_j\). Then
\[
P_{\theta, \sigma}(X^n_i = Y^n_i) \leq \sum_{h=1}^i P_{\theta, \sigma}(X^n_h \geq X^n_i)
\leq 1/(\theta_i - \theta_j) \sum_{h=1}^i |E_{\theta, \sigma}| X^n_h - \theta_h| + E_{\theta, \sigma}| X^n_i - \theta_i|
\]
from Chebyshev's inequality, and the result follows. Let next \(i < j\). Then:
\[
P_{\theta, \sigma}(X^n_i = Y^n_i) \leq 1/(\theta_j - \theta_i) \sum_{h=j}^k |E_{\theta, \sigma}| X^n_i - \theta_i| + E_{\theta, \sigma}| X^n_h - \theta_h| \to 0.
\]
Consider next part (b). Let \(g\theta = (\theta_{i1}, \ldots, \theta_{ik})\). There exists \(ij\) such that \(\theta_{ij} \neq \theta_{ij}\). Since \(gX^n = Y^n\) implies \(X^n_{ij} = Y^n_{ij}\), the result follows from (a). □

We can now state and prove the complete solution of pointwise consistency. First, let
\[
\mathcal{A}_\sigma(\theta) = \{a \in \mathcal{A}_\sigma : \mathcal{A}_\sigma(\theta, a) = m_\sigma(\theta)\}.
\]

**Theorem 2.1.** Assume (1.1) holds, and let \(\delta_n \in D\). Then (2.2) and (2.3) below are two equivalent, necessary and sufficient conditions for \(|\delta_n|\) to be consistent at \((\theta, \sigma)\).

(2.2) \[E_{\theta, \sigma}|\sum_{a \in \mathcal{A}_\sigma(\theta)} \delta_n(a | X^n, S^n)| \to 1\]

(2.3) \[E_{\theta, \sigma}|\sum_{a \in \mathcal{A}_\sigma(\theta)} \delta_n(a | Y^n, S^n)| \to 1\]

**Proof.** As remarked earlier we may assume \(\theta = \theta^*\). Now,
\[r_n(\theta, \sigma | \delta_n) = m_\sigma(\theta) + \sum_{a \in \mathcal{A}_\sigma(\theta)} |\mathcal{A}_\sigma(\theta, a) - m_\sigma(\theta)| \delta_n(a | X^n, S^n)\]
and it follows immediately that (2.2) is necessary and sufficient.

It remains to show that (2.2) ⇔ (2.3). Assume first that (2.3) holds. Let \(x \in \mathbb{R}^k\) and \(y = x^*\). The function \(I(a = b) = 1\) if \(a = b\), and 0 otherwise. Then
\[
\delta_n(a | x, s) \leq \sum_{g \in G} \delta_n(ga | y, s)I(gx = y).
\]
It is therefore enough to show that for all \(a \notin \mathcal{A}_\sigma(\theta)\), all \(g \in G\),
\[
E_{\theta, \sigma} \delta_n(ga | Y^n, S^n)I(gX^n = Y^n) \to 0.
\]
If \(ga \notin \mathcal{A}_\sigma(\theta)\), (2.5) follows directly from (2.3). If \(ga \in \mathcal{A}_\sigma(\theta)\), then \(\mathcal{A}_\sigma(\theta, ga) = \mathcal{A}_\sigma(\theta, ga)\).
$\zeta(g^{-1} \theta, a) < \zeta(\theta, a)$. Hence $g \theta \neq \theta$ and (2.5) follows from Lemma 2.1. The other way follows in the same manner. □

The individual selection functions of a subset selection procedure $\delta_n$ are given by

$$\psi^n_i(x, s) = P(\text{selecting } \pi_i \mid X^n = x, S^n = s) = \sum_{a \in \mathbb{F}} \delta_n(a \mid x, s).$$

Let $\psi^n = (\psi^n_1, \ldots, \psi^n_k)$, and let $\psi^n_0$ correspond to $\theta_0$. We note that for $\delta_n \in \mathcal{D}$,

$$(2.6) \quad \psi^n_i(x, s) = \psi^n_0(gx, s), \quad \forall g \in G, x \in \mathbb{R}^k, s \in E, i \in (1, \ldots, k).$$

When convenient, we shall denote the procedure $\delta_n$ by its selection functions $\psi^n$. Immediately from Theorem 2.1 we have the following result.

**Corollary 2.1.** Assume $\delta_n \in \mathcal{D}$ is consistent at $(\theta, \sigma)$, and that (1.1) holds. Then $E_{\theta, \sigma} \psi^n_i(X^n, S^n) \rightarrow 0$ for all $i$ such that $|a \in \mathcal{A} : a \ni i| \cap \varnothing(\theta^*) = \varnothing$.

We now go on to develop necessary and sufficient conditions for uniform consistency in $\mathcal{D}$. We shall assume

$$(2.7) \quad \zeta(\theta, a) \text{ is continuous in } (\theta, \sigma) \text{ for each } a \in \mathcal{A}.$$  

Let now $\Omega^* = \{\theta \in \Omega : \theta_1 \leq \cdots \leq \theta_k\}$. Then $\delta_n \in \mathcal{D}$ is uniformly consistent if and only if

$$\sup_{\theta \in K_i, \sigma \in K_j} \{r_n(\theta, \sigma \mid \delta_n) - m_\sigma(\theta)\} \rightarrow 0$$  

$$(2.8) \quad \text{for all compact sets } K_1, K_2 \text{ of } \Omega^*, \Sigma.$$  

Let $d$ be the Euclidean distance in $\mathbb{R}^k$. Define for any compact set $K_1, g \in G$ and $\delta > 0$, $M_{g, \delta} = \{\theta \in K_1 : d(g \theta, \theta) \geq \delta\}$.

We need the following modification of Lemma 2.1.

**Lemma 2.2.** Assume (1.2) holds. Let $K_1, K_2$ be compact subsets of $\Omega^*, \Sigma$ respectively. Let $K^{i,j}_1 = \{\theta \in K_1 : |\theta_i - \theta_j| \geq \epsilon\}$. Then

(a)  $$\sup_{K^{i,j}_1 \times K_2} P_{\theta, \sigma}(X^n = Y^n) \rightarrow 0 \quad \text{for all } \epsilon > 0$$

(b)  $$\sup_{M_{g, \delta} \times K_2} P_{\theta, \sigma}(gX^n = Y^n) \rightarrow 0, \quad \forall g \in G, \quad \forall \delta > 0.$$  

**Proof.** (a) We follow the same idea as in the proof of Lemma 2.1. Let first $i > j$. Then

$$\sup_{K^{i,j}_1 \times K_2} P_{\theta, \sigma}(X^n = Y^n)$$

$$\leq \frac{1}{\epsilon} \sum_{h=1}^j \sup_{K^{i,j}_1 \times K_2} E_{\theta, \sigma} |X^n_h - \theta_h| + \frac{j}{\epsilon} \sup_{K^{i,j}_1 \times K_2} E_{\theta, \sigma} |X^n_i - \theta_i| \rightarrow 0,$$
from (1.2). Similarly for i < j,
\[ \sup_{K_1 \times K_2} P_{\theta, \sigma}(X_i^n = Y_j^n) \leq (1/\varepsilon) \sum_{k=1}^{\infty} \sup_{K_1 \times K_2} E_{\theta, \sigma} \| X_h^n - \theta_h \| \\
+ ((k - j + 1)/\varepsilon) \sup_{K_1 \times K_2} E_{\theta, \sigma} \| X_i^n - \theta_i \|. \]

(b) For any fixed \((\theta^0, \sigma^0) \in M_{\delta, \delta} \times K_2\), let \(Q = \{i, j: |\theta_i^0 - \theta_j^0| \geq \delta/\sqrt{k}\}\). Then
\[ P_{\theta, \sigma}(gX^n = Y^n) \leq P_{\theta, \sigma}(\cup Q(X_i^n = Y_j^n)) \leq \sum_{i \in Q} \sup_{K_1 \times K_2} P_{\theta, \sigma}(X_i^n = Y_j^n), \]

where \(\delta_0 = \delta/\sqrt{k}\). The result now follows from part (a). [square]

The necessary and sufficient conditions for uniform consistency can now be stated. Define for compact sets \(K_1, K_2\) of \(\Omega^*, \Sigma^*\);
\[(2.9) \quad K_1^* = \{(\theta, \sigma) \in K_1 \times K_2: \mathcal{L}_\sigma(\theta, a) - m_\sigma(\theta) \geq \varepsilon\}. \]

**THEOREM 2.2.** Assume (1.2) holds, and that \(\mathcal{L}\) satisfies (2.7). Let \(\delta_n \in \mathcal{D}_1\). Then (2.10) and (2.11) below are two equivalent, necessary and sufficient conditions, for \(\delta_n\) to be uniformly consistent.
\[(2.10) \quad \sup_{K_2} E_{\theta, \sigma, \delta_n}(a | X^n, S^n) \rightarrow 0 \quad \text{as} \quad n \rightarrow \infty \]
\[(2.11) \quad \sup_{K_2} E_{\theta, \sigma, \delta_n}(a | Y^n, S^n) \rightarrow 0 \]

for all \(a \in \mathcal{A}, \varepsilon > 0\), and all compact sets \(K_1, K_2\) of \(\Sigma^*\) such that \(K_1^* \neq \emptyset\).

**PROOF.** Now,
\[ r_n(\theta, \sigma | \delta_n) - m_\sigma(\theta) = \sum_{\sigma \in \mathcal{A}'} [\mathcal{L}_\sigma(\theta, a) - m_\sigma(\theta)] E_{\theta, \sigma, \delta_n}(a | X^n, S^n). \]

Using the fact that \(\mathcal{L}\) and \(m\) are bounded on \(K_1 \times K_2\) we readily get from (2.10)
\[ \lim \sup_{K_1 \times K_2} [\mathcal{L}_\sigma(\theta, a) - m_\sigma(\theta)] E_{\theta, \sigma, \delta_n}(a | X^n, S^n) \leq \varepsilon; \quad \forall \varepsilon > 0. \]

Hence (2.8) holds. The other way is obvious.

To show that (2.11) \(\Rightarrow\) (2.10), it is enough from (2.4) to show
\[(2.12) \quad \sup_{K_2} E_{\theta, \sigma, \delta_n}(ga | Y^n, S^n) I(gX^n = Y^n) \rightarrow 0, \quad \forall g \in G. \]

Let \(g \in G\) be arbitrary and define \(\Delta_\sigma(\theta, a) = \mathcal{L}_\sigma(\theta, a) - m_\sigma(\theta)\). Since the loss-function is continuous, \(\exists\) a \(\delta > 0\) such that: \(d(\theta, \theta') < \delta\) and \(\theta \in K_1 \Rightarrow |\Delta_\sigma(\theta), ga) - \Delta_\sigma(\theta, ga)| \leq \varepsilon/2; \forall \sigma \in K_2\). Then, for \((\theta, \sigma) \in K_1 \times K_2\), \(\Delta_\sigma(\theta, a) \geq \varepsilon \iff \Delta_\sigma(\theta, ga) \geq \varepsilon \iff d(\theta, \theta') \geq \delta\) or \(\Delta_\sigma(\theta, ga) \geq \varepsilon/2\). Hence, \(K_1^* \subset K_{\varepsilon/2} \cup M_{\delta, \delta} \times K_2\) and (2.12) follows from (2.11) and Lemma 2.2. (2.10) \(\Rightarrow\) (2.11) in a similar way by showing that
\[ \sup_{K_1^*} E_{\theta, \sigma, \delta_n}(ga | X^n, S^n) I(X^n = gY^n) \rightarrow 0, \quad \forall g \in G. \quad \square \]

3. Some specific loss functions for selecting \(\pi_1\) close to \(\pi_{(k)}\). In this section we shall apply the theory in the previous section to loss functions that more or less reflect the desire to have \(\pi_{(k)}\) in the selected subset \(a\), while keeping
the size $|a|$ of the subset small. The invariant loss-functions to be considered are:

\begin{align}
(3.1) \quad & \zeta_l(\theta, a) = \theta_{(k)} - (1/|a|)\sum_{j \in a} \theta_j + r(\theta_{(k)} - \max_{i \in a} \theta_i); \quad r > 0 \\
(3.2) \quad & \zeta_2(\theta, a) = \theta_{(k)} - (1/|a|) \sum_{j \in a} \theta_j + LI(\max_{i \in a} \theta_i < \theta_{(k)}); \quad L > 0.
\end{align}

Here $I(a < b) = 1$ if $a < b$ and 0 otherwise.

\begin{align}
(3.3) \quad & \zeta_3(\theta, a) = c|a| + \theta_{(k)} - \max_{i \in a} \theta_i; \quad c > 0 \\
(3.4) \quad & \zeta_4(\theta, a) = c_1 I(\max_{i \in a} \theta_i < \theta_{(k)}) + c_2|a|; \quad c_1, c_2 > 0 \\
(3.5) \quad & \zeta_5(\theta, a) = |a| + c \sum_{i \in a} I(\theta_i = \theta_{(k)}); \quad c > 0 \\
(3.6) \quad & \zeta_6(\theta, a) = \sum_{i \in a} (\theta_{(k)} - \theta_i) + \alpha \sum_{i \in a} I(\theta_i = \theta_{(k)}); \quad \alpha > 0.
\end{align}

$\zeta_1$ was considered by Chernoff and Yahav (1977). They derived a Bayes procedure for normal populations. We show in Section 4 that this Bayes procedure is uniformly consistent for $\zeta_1$.

$\zeta_2$ was proposed by Bickel and Yahav (1977). This loss is not continuous in $\theta$ for given $a$, so for $\zeta_2$ only pointwise consistency will be discussed. The loss $\zeta_2$ has been used by Goel-Rubin (1977), who derived a Bayes procedure. In the case of normal populations we show in Section 4 that the Bayes procedure is uniformly consistent for $\zeta_2$. Gupta and Hsu (1978) employed $\zeta_4(\theta, a)$. $\zeta_3$ and $\zeta_6$ are members of the class of additive loss-functions considered by Björnstad (1981). We note that $\zeta_4$, $\zeta_5$, $\zeta_6$ are not continuous in $\theta$ for fixed $a$. Since all the loss-functions are independent of $\sigma$, we will use the notation $m(\theta)$ and $\mathcal{M}(\theta)$ (see (2.1)).

\[ E^*_\epsilon = \{|\theta \in K_1: \zeta(\theta, a) - m(\theta) \geq \epsilon| \} \] for any compact set $K_1$ of $\Omega^*$ such that, from (2.9), $K_2^* = E^*_\epsilon \times K_2$. Define for any compact set $K_1$ of $\Omega^*$ and $\epsilon > 0$,

\begin{equation}
(3.7) \quad K_1^* = \{\theta \in K_1: \theta_{(k)} - \theta_i \geq \epsilon\}.
\end{equation}

**Theorem 3.1.** Let the loss be $\zeta_1$, given by (3.1), and let $\psi^n \in D_l$. (a) Assume (1.1) holds. Then $\psi^n$ is consistent at $(\theta, \sigma)$ with $\theta_{(p-1)} < \theta_{(p)} = \theta_{(k)}$ if and only if

\begin{equation}
(3.8) \quad E^*_{\theta, \sigma} \psi^n(X^n, S^n) \rightarrow 0 \quad \text{for} \quad i \leq p - 1.
\end{equation}

(b) Assume (1.2) holds. Then $\psi^n$ is uniformly consistent if and only if

\begin{equation}
(3.9) \quad \sup_{K_1^* \times K_2} E^*_{\theta, \sigma} \psi^n(X^n, S^n) \rightarrow 0 \quad \text{for} \quad i = 1, \ldots, k - 1,
\end{equation}

for all compact sets $K_1$, $K_2$ of $\Omega^*$, $\Sigma$ and all $\epsilon > 0$ such that $K_1^* \times K_2 \neq \emptyset$.

**Proof.** (a) Using Theorem 2.1, the result follows from Corollary 2.1 and the fact that $\delta_\epsilon(a | x, s) \leq \psi^n(x, s)$ if $a \ni i$.

(b) Since $K_1^* \subset E^*_\epsilon(h_k)$, for $a \ni i$, (3.9) follows from (2.10). Assume now (3.9). Let $E^*_\epsilon \times K_2 \neq \emptyset$, and $a = \{i_1, \ldots, i_q\}$ where $i_1 < \cdots < i_q$. Then $\delta_\epsilon(a | x, s) \leq \psi^n(x, s)$.
ψ^n_i(x, s). Furthermore, E^n_\theta \subset K^n_{\theta_i} where θ = ε/(1 + r), and
\[
\sup_{E^n_\theta \times K^n_\theta} E^{\theta, \delta_n}(a \mid X^n, S^n) \leq \sup_{K^n_{\theta_i} \times K^n_\theta} E^{\theta, \psi^n_i}(X^n, S^n) \to 0
\]
from (3.9). □

**Remark.** m(θ) and \( \mathcal{S}(\theta) \) are the same for \( \mathcal{A}_2 \), given by (3.2), as for \( \mathcal{A}_1 \). Hence Theorem 3.1 (a) is valid also for \( \mathcal{A}_2 \).

**Theorem 3.2.** Let the loss be \( \mathcal{L}_3 \), given by (3.3), and assume \( \psi^n \in \mathcal{D}_1 \).

(a) Assume (1.1) holds. Then \( \psi^n \) is consistent at \( (θ, σ) \) with \( θ_{(p-1)} < θ_{(p)} = θ_{(k)} \) if and only if
\[
E^{\theta, \sigma}_{\theta, \delta_n}(\sum_{i=1}^k \psi^n_i(X^n, S^n)) \to 1 \quad \text{and} \quad E^{\theta, \sigma}_{\theta, \delta_n}(\psi^n_i(X^n, S^n)) \to 0 \quad \text{for i ≤ p - 1}.
\]
(b) Assume (1.2) holds. Then \( \psi^n \) is uniformly consistent if and only if
\[
\sup_{K^n_1 \times K^n_2} E^{\theta, \sigma}_{\theta, \delta_n}(\sum_{i=1}^k \psi^n_i(X^n, S^n)) \to 1
\]
and
\[
\sup_{K^n_1 \times K^n_2} E^{\theta, \sigma}_{\theta, \delta_n}(\psi^n_i(X^n, S^n)) \to 0 \quad \text{for i ≤ k - 1}
\]
and for all compact sets \( K_1, K_2 \) of \( Ω^* \), \( Σ \) and all \( ε < 0 \) such that \( K_1 \times K_2 \neq \emptyset \).

**Proof.** (a) Let \( θ_{(p-1)} < θ_{(p)} = θ_{(k)} \) and \( θ = θ^* \). Then \( \mathcal{S}(\theta) = \{(p), \ldots, (k)\} \).

Now using the property that \( δ_n(a \mid x, s) \leq ψ^n_i(x, s) \) if \( a \exists i \) and the equation
\[
\sum_{i=1}^k \psi^n_i(x, s) = 1 + \sum_{q=2}^k (q - 1) \sum_{|a| = q = 0} \delta_n(a \mid x, s),
\]
the result follows immediately from Theorem 2.1.

(b) Assume (2.10) holds. Let \( E^n_\theta = \{θ ∈ K^n_1: L_θ(θ, a) ≥ c + ε\} \). Then \( K^n_\theta = E^n_\theta \times K_2 \). For \( |a| ≥ 2 \) and \( ε < c, E^n_\theta = K^n_1 \), and (3.11) follows. Also, (3.12) follows from the fact that \( E^n_{\theta_i} = K^n_1 \).

Now, let us assume that (3.11) and (3.12) hold. Clearly for \( |a| ≥ 2 \), (3.11) \( ⇒ \) (2.10). For \( a = \{i\}, i ≤ k - 1; δ_n(|i| \mid X^n, S^n) ≤ ψ^n_i(X^n, S^n) \) and (2.10) follows from the fact that \( E^n_{\theta_i} = K^n_1 \). □

**Remark.** \( \mathcal{S}(\theta) \) is the same for \( \mathcal{A}_4 \) as for \( \mathcal{A}_3 \). Hence Theorem 3.2 (a) is valid also for \( \mathcal{A}_4 \).

Most of the research on subset selection has assumed that the procedures satisfy a certain control condition. The most common is the so-called \( P^* \)-condition, due primarily to Gupta (1956, 1965) and Seal (1955). Let a subset that includes \( π_{(k)} \) be called a correct selection, CS. The \( P^* \)-condition is:
\[
\inf_{O \times \Sigma} P^n_\theta(\{CS \mid δ_n\}) = \inf_{O \times \Sigma} E^{\theta, \delta_n}(\psi^n_{(k)}) ≥ P^*; \quad 1/k < P^* < 1.
\]
Suppose \( Ω ⊃ Ω_0 = \{θ: θ_1 = \cdots = θ_k\} \). If \( ψ^n \in \mathcal{D}_1 \) is pointwise consistent for \( \mathcal{A}_3 \) and \( \mathcal{A}_4 \) on \( Ω \times Σ \), it follows from Theorem 3.2 (a) that for \( θ ∈ Ω_0, E^{θ, ψ^n}_{θ, δ_n} = \cdots = \)
$E_{\theta,\psi_\theta^n} \to 1/k$, and therefore any pointwise consistent invariant procedure $\delta_n$ must have $\lim \sup_n [\inf_{0 \leq x \leq P_{\theta,\sigma}(CS | \delta_n)}] \leq 1/k$, and cannot satisfy (3.13) as $n \to \infty$. Here we used that fact, derived from (2.6), that for any $\psi^n \in D_I$, $E_{\theta,\psi_\theta^n} = E_{\theta,\psi_\theta^n}$ if $\theta_i = \theta_j$. In a similar way we see that if $P^* > \frac{1}{2}$, no procedure satisfying (3.13) for all $n$ can be consistent at any $\theta$ where $\theta_{(k-1)} = \theta_{(k)}$.

Let $\Omega_1 = \{ \theta \in \Omega: \theta_{(k-1)} < \theta_{(k)} \}$. Procedures that are consistent on $\Omega_1$ can of course satisfy the $P^*$-condition. Now, for any compact set $K_1$ in $\Omega_1^*$ there exists $\epsilon > 0$ such that $K_1 \subset \{ \theta \in \Omega_1^*: \theta_k - \theta_{k-1} \geq \epsilon \}$. From Theorem 3.2(b), we readily see that if (1.2) holds and $\psi^n \in D_I$, then $\psi^n$ is uniformly consistent on $\Omega_1 \times \Sigma$ for $\delta_5$ iff (3.12) holds.

Let us now consider $\delta_5$ with $c > 1$ and $\delta_6$.

**Theorem 3.3.** Let the loss be $\delta_5$ with $c > 1$ or $\delta_6$, given by (3.5) and (3.6). Assume (1.1) holds, and $\psi^n \in D_I$. Then $\psi^n$ is consistent at $(\theta, \sigma)$ with $\theta_{(p-1)} < \theta_{(p)}$ if and only if

$$E_{\theta,\sigma} \psi_\theta^n(X^n, S^n) \to \begin{cases} 0 & \text{for } i \leq p - 1 \\ 1 & \text{for } i \geq p. \end{cases}$$

(**Remark.**) Comparing (3.14) with (3.8) and (3.10), we see that $\delta_5$, $\delta_6$ requires one to select all populations $\pi_i$ with $\theta_i = \theta_{(k)}$ and excluding all others, while $\delta_1$, $\delta_2$, $\delta_3$, $\delta_4$ essentially requires one to exclude all $\pi_i$ with $\theta_i < \theta_{(k)}$ and including only at least one $\pi_i$ with $\theta_i = \theta_{(k)}$.

**Proof.** Let $\theta_{(p-1)} < \theta_{(p)} = \theta_{(k)}$ and $\theta = \theta^*$. Let $a_0 = \{p, \ldots, k\}$. By expressing $\delta_5$ as $\delta_5(\theta, a) = \#\{i \in a: \theta_i < \theta_k\} + (1 - c)\#\{i \in a: \theta_i = \theta_k\} + c\#\{i: \theta_i = \theta_k\}$ we see that $\delta_5(\theta) = a_0$ for both $\delta_5$ and $\delta_6$, since $1 - c < 0$. From Theorem 2.1 it remains to show that (3.14) is equivalent to

$$E_{\theta,\sigma} \delta_n(a_0 | X^n, S^n) \to 1.$$ 

Obviously, (3.15) $\Rightarrow$ (3.14). Assume now that (3.15) does not hold. Then there exists $a_1 \neq a_0$ such that $\lim \sup E_{\theta,\sigma} \delta_n(a_1 | X^n, S^n) = \beta > 0$. If there is an $i \in a_1$, $i \leq p - 1$, then $\lim \sup E_{\theta,\sigma} \psi_i^n \geq \beta$, violating (3.14). If $\{i \in a_1: i \geq p\}$ there must exist $j \geq p$, $j \not\in a_1$ and therefore $\lim \inf E_{\theta,\sigma} \psi_j^n \leq 1 - \beta$, implying again that (3.14) does not hold. $\square$

$\psi^n$ is said to be a just procedure if $x_i \geq x_i'$ and $x_j \leq x_j'$ for $j \neq i$ implies that $\psi_i^n(x) \geq \psi_i^n(x')$.

**Corollary 3.1.** Assume $\psi^n$ is just, invariant and pointwise consistent for $\delta_5$ with $c > 1$ and $\delta_6$ on $\Omega \times \Sigma$, $\Omega \supset \Omega_0 = \{\theta: \theta_1 = \cdots = \theta_k\}$. Then $\inf_{0 \leq x \leq P_{\theta,\sigma}} CS | \psi^n \to 1$.

**Proof.** Nagel (1970) showed that for any just procedure, $\inf_{0 \leq x \leq P_{\theta,\sigma}} CS | \psi^n$ occurs at some $\theta \in \Omega_0$. From (3.14) we have that for $\theta \in \Omega_0$,

$$E_{\theta,\sigma} \psi^n_\theta = \cdots = E_{\theta,\sigma} \psi^n_k \to 1.$$ $\square$
Corollary 3.1 means that no just, invariant procedure satisfying (3.13) with equality can be pointwise consistent for \( \mathcal{I}_0 \) or \( \mathcal{I}_a \), if \( P^* \) is chosen independent of \( n \). Hence, if these loss functions reflect the true losses involved in the selection problem the \( P^* \)-condition is not appropriate. It seems clear that it is the term \( \sum_{i \in \theta} I(\theta_i = \theta(k)) \) that makes (3.13) inappropriate.

Finally, consider \( \mathcal{I}_0 \) with \( c \leq 1 \). The following result is needed.

**Lemma 3.1.** \( Y^n = (X^n)^* \).

(a) \( (1.1) \Rightarrow E_{\theta, \sigma} | Y^n - \theta^* | \to 0 \) for \( i = 1, \ldots, k \).

(b) \( (1.2) \Rightarrow \sup_{K_1 \times K_2} E_{\theta, \sigma} | Y^n_i - \theta_i | \to 0 \) for \( i = 1, \ldots, k \) and all compact sets \( K_1, K_2 \) in \( \Omega^* \), \( \Sigma \).

**Proof.** Let \( \theta = \theta^* \). Then \( | Y^n_i - \theta_i | \leq \sum_{j=1}^{k} | X^n_j - \theta_j | + \sum_{j \neq i} | \theta_j - \theta_i | \cdot I(X^n_j = Y^n_j) \). Hence

\[
E_{\theta, \sigma} | Y^n_i - \theta_i | \leq \sum_{j=1}^{k} E_{\theta, \sigma} | X^n_j - \theta_j | + \sum_{j \neq i} | \theta_j - \theta_i | P_{\theta, \sigma}(X^n_j = Y^n_j).
\]

Then (a) follows directly from (1.1) and Lemma 2.1, and (b) follows from (1.2) and Lemma 2.2. \( \Box \)

For this particular loss we shall assume that \( X^n_1, \ldots, X^n_k \) are independent, each \( X^n_i \) has density \( f^n_n(\cdot, \theta_i) \) with respect to a \( \sigma \)-finite measure, and \( \Omega = \theta^k \). It is assumed that for fixed \( (n, \sigma) \), \( f^n_n \) has the monotone likelihood-ratio property.

Bjørnstad (1981) showed that there is a uniformly minimum risk procedure in \( D_f \) for \( \mathcal{I}_0 \) when \( c \leq 1 \). It is given by:

\[
\delta_0([i] | y) = 1/q \quad \text{for} \quad i \geq k - q + 1,
\]

when

\[
y_{k-q} < y_{k-q+1} = \cdots = y_k; \quad \forall y \in D = \{ x \in \mathbb{R}^k : x_1 \leq \cdots \leq x_k \}.
\]

Obviously, \( \delta_0 \) is the only interesting procedure in \( D_f \) for this loss. Even though \( \mathcal{I}_0 \) is not continuous in \( \theta \) we can say something about uniform consistency of \( \delta_0 \) as the next result shows.

**Theorem 3.4.** The loss is \( \mathcal{I}_0(\theta, a) = | a | + c \sum_{i \in \theta} I(\theta_i = \theta(k)) \) with \( 0 < c \leq 1 \).

(a) Assume (1.1) holds. Then \( \delta_0 \) is pointwise consistent on \( \Omega \times \Sigma \).

(b) Assume (1.2) holds and that \( f^n_n(x, \theta) \) is a continuous function of \( (\theta, \sigma) \). Then \( \delta_0 \) is uniformly consistent on \( \Omega_1 \times \Sigma \), where \( \Omega_1 = \{ \theta \in \Omega : \theta(k) > \theta(k-1) \} \).

**Proof.** (a) Let \( \theta = \theta^* \) and assume \( \theta_{k-1} < \theta_p = \theta_k \). We see that when \( c \leq 1 \),

\[
\mathcal{I}_0(\theta) = |\theta| + c \sum_{i=1}^{k} \delta_0([i] | Y^n).
\]

From (2.3) of Theorem 2.1 we need to show that \( E_{\theta, \sigma} \sum_{i=p}^{k} \delta_0([i] | Y^n) \to 1 \). \( \sum_{i=1}^{k} \delta_0([i] | Y^n) = 1 \) so we must show that \( E_{\theta, \sigma} \delta_0([i] | Y^n) \to 0 \) for \( i \leq p - 1 \). Now, \( E_{\theta, \sigma} \delta_0([i] | Y^n) \leq P_{\theta, \sigma}(Y^n_i = Y^n_k) \to 0 \) since \( Y^n_k - Y^n_i \to P \theta_k - \theta_i > 0 \), from Lemma 3.1.

(b) On \( \Omega_1 \), \( m(\theta) = 1 \) so we must show that \( \sup_{K_1 \times K_2} f_n(\theta, \sigma | \delta_0) \to 1 \) for all
compact sets $K_1, K_2$ in $\Omega_1, \Sigma$. We readily derive that
\[ r_n(\theta, \sigma | \delta_0) \leq P_{\theta, \sigma}(X^*_n \geq \max_{1 \leq j \leq k-1} X^*_j) + (1 + c) P_{\theta, \sigma}(X^*_n \leq \max_{1 \leq j \leq k-1} X^*_j). \]
It follows that it is sufficient to show $\inf_{K_1 \times K_2} P_{\theta, \sigma}(X^*_n > X^*_j) \to 1$ for $j \leq k - 1$.

As mentioned earlier, there exists $\epsilon > 0$ such that $\theta \in K_1 \Rightarrow \theta_k - \theta_{k-1} \geq \epsilon$. Since $f^*_\alpha(\cdot, \theta)$ is continuous in $(\theta, \sigma)$, $P_{\theta, \sigma}(X^*_n > X^*_j)$ is a continuous function of $(\theta, \sigma)$. Hence infimum occurs at some $(\theta^n, \sigma^n) \in K_1 \times K_2$. From (1.2), $X^*_n - \theta^n_i \to P_0$ under $(\theta^n, \sigma^n)$, and
\[ P_{\theta^n, \sigma^n}(X^*_n - X^*_j > 0) \geq P_{\theta^n, \sigma^n}((X^*_n - \theta^n_i) - (X^*_n - \theta^n_j)) > -\epsilon \to 1. \]

4. Selection of means from normal populations. The $k$ populations are now assumed to be normally distributed, and $X^*_n$ is the sample mean of size $n$ from $\pi_i$. Hence $X^*_n, \ldots, X^*_k$ are independent and $X^*_n$ is $N(\theta, \sigma^2/n)$, where $\sigma$ is unknown, $\sigma \in (0, \infty)$. Moreover, let $S^2 = S^2_n$, the usual U.M.V.U. estimate of $\sigma^2$. Then $S^2_n \to P_0 \sigma^2$. In this section $\Omega = \mathbb{R}^k$ and (1.2) clearly holds. We shall apply the theory in the previous section for the loss functions $\zeta_1 - \zeta_6$ on some subset selection procedures that have been studied in the literature. It is now assumed that $c > 1$ in $\zeta_6$.

Consider the class $\mathcal{S}$, proposed by Seal (1955). $\mathcal{S} = \{\psi^{c, n}: \sum_{i=1}^{k-1} c_i = 1, c_i \geq 0; \forall i\}$, where
\[ \psi^{c, n}_i = 1 \iff X^*_i \geq \sum_{j=1}^{k-1} c_j X^*_j - S_n D_n(c); \quad D_n(c) \geq 0. \]
Here $X^*_1 \leq \cdots \leq X^*_k$ are the ordered $X^*_j$, $j \neq i$. Seal assumed $D_n(c)$ is determined such that the $P$-condition (3.13) holds with equality. We shall, however, consider $\psi^{c, n}$ for any sequence $\{D_n(c)\}$. If we want (3.13) to be satisfied, it is readily seen, since infimum of $P(CS | \psi^{c, n})$ occurs when $\theta_1 = \cdots = \theta_k$, that $\sqrt{n}D_n(c) \to t(c)$ where
\[ P[\sum_{i=1}^{k-1} c_i Z^*_j - Z_k \leq t(c)] = P^*. \]
Here $Z^*_1 < \cdots < Z^*_k$ are the ordered $Z_1, \cdots, Z_k$, and $Z_1, \cdots, Z_k$ are i.i.d. $N(0, 1)$.

One procedure in $\mathcal{S}$ has received special attention in the literature by many authors. Gupta (1956, 1965) suggested the use of $c_{k-1} = 1$. Let us call this procedure $\psi^{G, n}$, and denote $D_n(c)$ by $d_n$, such that
\[ \psi^{G, n}_i = 1 \iff X^*_i \geq X^*_k - S_n d_n. \]
Applying Lemma 2.1, the following two results can be readily shown, using Theorems 3.1–3.3.

**Theorem 4.1.** Let the loss be one of $\zeta_1 - \zeta_6$, given by (3.1)–(3.6), where $c > 1$ for $\zeta_6$. Assume $c_{k-1} < 1$. Then $\psi^{c, n}$, given by (4.1), is not pointwise consistent on $\mathbb{R}^k \times (0, \infty)$ for any sequence $\{D_n(c)\}$.

This result shows that no procedure in $\mathcal{S}$, except $\psi^{G, n}$, has a chance of being consistent for the losses $\zeta_1 - \zeta_6$. 

The cases of \( \{d_n\} \) when \( \psi^{G,n} \) is consistent for the different losses are specified in the next result.

**Theorem 4.2.** Let \( \psi^{G,n} \) be given by (4.3).
(a) \( \psi^{G,n} \) is uniformly (pointwise) consistent on \( \mathbb{R}^k \times (0, \infty) \) for \( \zeta_1(\zeta_2) \), given by (3.1), ((3.2)), if and only if \( d_n \to 0 \).
(b) \( \psi^{G,n} \) is uniformly (pointwise) consistent on \( \mathbb{R}^k \times (0, \infty) \) for \( \zeta_3(\zeta_4) \), given by (3.3), ((3.4)), if and only if \( \sqrt{n}d_n \to 0 \).
(c) \( \psi^{G,n} \) is pointwise consistent on \( \mathbb{R}^k \times (0, \infty) \) for \( \zeta_c \) with \( c > 1 \) and \( \zeta_6 \) if and only if \( d_n \to 0 \) and \( \sqrt{n}d_n \to \infty \).

**Remark.** If \( d_n \) is determined such that \( \psi^{G,n} \) satisfies (3.13) with equality, then, from (4.2), \( \psi^{G,n} \) is uniformly (pointwise) consistent for \( \zeta_1(\zeta_2) \), but not consistent for any of the other losses.

For the rest of this section we assume \( \sigma \) is known. Two classes of invariant, admissible procedures for \( \zeta_8 \) and \( \zeta_6 \) with \( c > 1 \) are given below.

(4.4) \[ \psi^{1,n}_i = 1 \iff c \exp(b_nX^n_i) \geq \sum_{j=1}^{k} \exp(b_nX^n_j) \text{ or } X^n_i = X^n_{(k)}. \]
(4.5) \[ \psi^{2,n}_i = 1 \iff (1 + (\alpha/b_n))\exp(b_nX^n_i) \geq \sum_{j=1}^{k} \exp(b_nX^n_j) \text{ or } X^n_i = X^n_{(k)}. \]

Bjørnstad (1981) showed that \( \psi^{1,n} \) is admissible for \( \zeta_8 \) and \( \psi^{2,n} \) is admissible for \( \zeta_6 \), for all \( b_n > 0 \). From Theorem 3.3, the following result is easily shown.

**Theorem 4.3.** (a) Let \( \psi^{1,n} \) be given by (4.4), and assume \( c > k \) in \( \zeta_5 \). Then \( \psi^{1,n} \) is pointwise consistent on \( \mathbb{R}^k \) for \( \zeta_8 \), \( \zeta_6 \) if and only if \( b_n \to \infty \) and \( b_n/\sqrt{n} \to 0 \).
(b) Let \( \psi^{2,n} \) be given by (4.5). Then \( \psi^{2,n} \) is pointwise consistent on \( \Omega_1 = \{ \theta \in \mathbb{R}^k, \theta_{(k-1)} < \theta_{(k)} \} \) for \( \zeta_8 \), \( \zeta_6 \) if and only if \( \lim \inf b_n \geq \alpha \).

**Remark 1.** It is readily seen that \( \lim \inf b_n \geq \alpha \) implies
\[ E_{\theta' \psi^{2,n}_k} \to 1/(k - p + 1), \]
for some subsequence, for all \( p \leq k - 1 \). By (b) of Theorem 4.3 this implies, from Theorem 3.3, that \( \psi^{2,n} \) is not pointwise consistent on \( \mathbb{R}^k \) for any \( \{b_n\} \).

**Remark 2.** It can be shown that if \( c \leq k \), then \( \psi^{1,n} \) is not pointwise consistent on \( \mathbb{R}^k \) for any \( \{b_n\} \).

At last in this section we consider the Bayes-procedures derived for normal exchangeable priors for \( \zeta_1, \zeta_3, \zeta_4 \) by Chernoff and Yahav (1977), Goel and Rubin (1977) and Gupta and Hsu (1978) respectively. The prior is: \( \theta' \sim N_{\alpha}(me, rI + tU) \), where \( e = (1, \ldots, 1)' \) and \( U = ee' \), \( r > 0 \), \( t \geq 0 \). As shown by the authors mentioned above, the risks of the Bayes-procedures do not depend on \( m, t \) so we may let \( m = 0, t = 0 \). If so, \( (\theta' | X^n = x) \) is
\[ N \left( \frac{r}{q_n + r}, \frac{r q_n}{q_n + r} I \right), \]
where \( q_n = \frac{\sigma^2}{n} \) and \( \hat{\theta} = \frac{r}{q_n + r} X^n \)
is the usual squared error loss Bayes estimate.
Let us consider first $\mathcal{A}$ and its Bayes procedure $\delta^B_n$. Consider $T_n(a \mid X^n) = E[\zeta(a, a) \mid X^n] = (1 + r)E[\theta(a) \mid X^n] - (1/a) \sum_{j:a_j} \theta_j - rE[\max_{i:a_i} \theta_i \mid X^n]$. Then, $\delta^B_n(a \mid X^n) = 1$ iff $a$ minimizes $T_n(a' \mid X^n)$ for all $a' \in \mathcal{A}$. Clearly, $\delta^B_n$ is permutation-invariant. Lemma 3.1(b) implies that $T_n(a \mid Y^n) \rightarrow_{d} \zeta(\theta, a) \rightarrow_{p} 0$ for any sequence $\{\theta_n\} \in K$, and compact set $K$ of $\Omega^*$. We shall show that $\delta^B_n$ is uniformly consistent for $\mathcal{A}$ on $\mathbb{R}^k$. In order to do so it is enough, from Theorem 2.2, to show that $\sup_{\mathcal{E}} E_{\theta} \delta^B_n(a \mid Y^n) \rightarrow 0$, where $E_{\theta} = \{\theta \in K : \zeta(\theta, a) \geq \varepsilon\}$. Now,

$$E_{\theta} \delta^B_n(a \mid Y^n) \leq P_{\theta} T_n(a \mid Y^n) \leq T_n(\{k\} \mid Y^n),$$

which is a continuous function in $\theta$. Hence, for some $\theta^* \in E_{\theta}$,

$$\sup_{\mathcal{E}} E_{\theta} \delta^B_n(a \mid Y^n) \leq P_{\theta^*} T_n(a \mid Y^n) \leq T_n(\{k\} \mid Y^n) \leq -\varepsilon \rightarrow 0.$$

Next, we consider the Bayes-procedure $\delta^B_n$ for $\zeta$. Let $\gamma^2_n = q_n/(q_n + r) = 1/(r^{-1} + n\sigma^{-2})$. According to Corollary 5 of Goel and Rubin (1977), $\delta^B_n(\{k\} \mid y) = 1$, for $y = x^*$, $\forall x \in \mathbb{R}^k$, provided $c/\gamma_n \geq 1/\sqrt{\pi}$. Since $\gamma_n \rightarrow 0$, it now follows immediately from (2.11) of Theorem 2.2 that $\delta^B_n$ is uniformly consistent for $\mathcal{A}$ on $\mathbb{R}^k$, since $K^2 \neq \emptyset$ implies $a \neq \{k\}$.

Finally, let $\delta^B_n$ be the Bayes-procedure for $\mathcal{A}_4$, given by (3.4). Gupta and Hsu (1978) showed that $\delta^B_n$ is given by

$$\psi^B_{i,p} = 1 \iff X^n \geq \max_{j<i} X^n \quad \text{or} \quad P(\theta_i = (\theta_a) \mid X^n) \geq c_2/c_1.$$ 

Clearly $\psi^B_{i,p}$ is permutation-invariant. Since $\mathcal{A}_4$ is not continuous in $\theta$ we shall discuss only the pointwise consistency properties of $\psi^B_{i,p}$. Let now $\theta_{(p-1)} < \theta_{(p)} = \theta_{(k)}$, and $\theta = \theta^*$. Then for $i \leq p - 1$,

$$P(\theta_i = (\theta_{(k)} \mid X^n) \leq \Phi((\hat{\theta}_i - \hat{\theta}_k)/\gamma_n\sqrt{2}) \rightarrow 0,$$

since $\hat{\theta}_i \rightarrow \theta_{(k)}$ and $\gamma_n \rightarrow 0$. Therefore $E_\theta \psi^B_{i,p} \rightarrow 0$ for $i \leq p - 1$. This implies that (3.10) holds if $p = k$, and from the remark after Theorem 3.2, we have shown that $\psi^B_{i,p}$ is pointwise consistent on $\Omega_1 = \{\theta \in \mathbb{R}^k : \theta_{(k)} > \theta_{(k-1)}\}$. If $c_2/c_1 \geq \gamma_2$ then $P(\theta_k = (\theta_{(k)} \mid X^n) \geq c_2/c_1 \rightarrow X^n \geq \max_{j<k-1} X^n$ and $E_\theta \psi^B_{k,k} \rightarrow 1/(k - p + 1)$. Therefore, from (3.10), $\psi^B_{k,k}$ is pointwise consistent on $\mathbb{R}^k$, if $c_2/c_1 \geq \gamma_2$. However, if $c_2/c_1 < \gamma_2$ it is straightforward to show that $\psi^B_{k,k}$ is not consistent on all points in $\mathbb{R}^k - \Omega_1$.

Acknowledgment. The author would like to thank two referees for helpful comments.

REFERENCES


