Statistica Sinica **5**(1995), 291-301

COMPLETE CLASSES FOR CONFIDENCE SET ESTIMATION

L. D. Brown and Arthur Cohen

University of Pennsylvania and Rutgers University

Abstract. Consider the statistical model where a statistic T has a distribution belonging to a one parameter family which has the strict monotone likelihood ratio property. We study confidence set estimation for two types of risk functions. One is a vector risk consisting of two components: one is one minus the coverage probability and the other is the expected measure of the set. The other risk is a linear combination of the two components. Under certain conditions we find that monotone procedures are a complete class and also that procedures with an interval property are a complete class. As consequences we find that nonrandomized procedures are a complete class. Furthermore, we find an analogue to the Rao-Blackwell theorem and construction for exponential family models. Other applications and observations are noted.

Key words and phrases: Admissibility, interval property, monotone procedure, randomized procedures, Rao-Blackwell construction.

1. Introduction and Summary

Consider a statistical model where a statistic T has a density belonging to a one parameter family which has the strict monotone likelihood ratio (SMLR) property. Let θ designate the parameter and $f_T(t|\theta)$ denote the density of T with respect to Lebesque measure. In Section 3 assume also that $f_T(t|\cdot)$ can be differentiated through the integral, as in (3.4). Such would arise, for example, in many situations where a sample of size n is taken from a population which has a one parameter exponential family distribution; then T would be a sufficient statistic. For such a model many theoretical and practical results exist in point estimation and hypothesis testing theory. For example, in hypothesis testing a complete class of tests for testing $H_0 : \theta \leq \theta_0$ vs. $H_1 : \theta > \theta_0$ consists of monotone tests i.e. reject if and only if $T > C$. (See Karlin (1956).) For one sided testing or point estimation with bowl-shaped loss functions, Brown, Cohen and Strawderman (BCS) (1976) showed that under mild conditions, monotone procedures are a complete class. The BCS (1976) result applies to fixed width confidence interval estimation since the latter type of inference is really point estimation. In addition to complete class results, BCS (1976) demonstrate the following, each under certain conditions:

(a) Randomized procedures can be eliminated.

(b) An analogue to the Rao-Blackwell theorem and construction regarding sufficient statistics (for non-convex loss functions).

(c) Two observations are strictly better than one observation for an entire class of bowl-shaped loss functions. (Hence n observations are strictly better than $1, 2, \ldots,$ or $n-1$ observations.)

This paper addresses the question of complete class theory, randomized procedures, Rao-Blackwell theorem and construction, etc. for confidence set estimation. Whereas there are many studies involving a loss function approach to confidence set estimation, very few, if any, results of the above type exist.

Our approach to confidence set estimation through a loss function is not new. See for example, Winkler (1972), Joshi (1970), Cohen and Strawderman (1973), and Casella and Berger (1990). A more typical approach is to require a minimum coverage probability (confidence coefficient) in confidence estimation. One advantage of the loss function approach is that as the sample size goes to infinity a reasonable procedure will have a risk that goes to zero. Furthermore, we emphasize a vector risk function in which one component is one minus the coverage probability. One can restrict the class of procedures to those for which this component is bounded above, thus insuring a minimum coverage probability.

In addition to being concerned with confidence set estimators that are monotone (in a sense defined in the next section) we shall be concerned with confidence set estimators that have an "interval" property. The interval property seems to be intuitive and perhaps, for complete class theory, more compelling than the monotone property. A confidence set has the interval property if for every θ, the set of $T = t$ for which θ is included in the confidence set, is an interval.

A convenient notation for a confidence set estimator is the function $\psi(\theta|t)$ where $0 \leq \psi(\theta|t) \leq 1$ denotes the probability that θ is in the confidence set, given t. We first consider the loss function

$$L(\theta, \psi) = b \int \psi(\rho|t) d\rho + 1 - \psi(\theta|t), \tag{1.1}$$

where $b > 0$. Note that the quantity $\int \psi(\rho|t) d\rho$ gives the volume of a confidence set when it is nonrandomized. The risk function therefore is

$$R(\theta, \psi) = bE_\theta[\int \psi(\rho|T) d\rho] + (1 - E_\theta[\psi(\theta|T)]). \tag{1.2}$$

We also consider the vector loss function

$$L^v(\theta, \psi) = (\int \psi(\rho|t) d\rho, 1 - \psi(\theta|t)), \tag{1.3}$$

with risk

$$R^v(\theta, \psi) = (R_1^v, R_2^v) = (E_\theta \text{ (measure)}, 1 - P_\theta \text{ (coverage)}), \tag{1.4}$$

where measure means Lebesgue measure. Vector losses are discussed in Cohen and Sackrowitz (1984). The vector loss formulation is particularly valuable for set estimation. In terms of admissible procedures, if a procedure is inadmissible for loss (1.3) it is inadmissible for loss (1.1). The converse is however not true. Thus if an admissible procedure with respect to loss (1.3) must have a property (to be admissible or essentially admissible) then any procedure without the property is not very good. Furthermore the vector loss approach is appealing since oftentimes one wishes to restrict confidence set estimators whose probability of coverage is greater than or equal to a given constant.

For convenience only, we assume that the distribution of T is absolutely continuous. Among the main results are the following:

(1) For loss (1.1), Bayes (or generalized Bayes) procedures are monotone.

(2) For loss (1.1), limits of sequences of Bayes procedures are an e-essentially complete class. (Essential admissibility (e-admissibility) and an e-essentially complete class are defined in Section 2.)

(3) For loss (1.3) and hence for loss (1.1), the set of procedures having the interval property is an e-complete class.

(4) For loss function (1.3) and hence for loss (1.1), nonrandomized procedures are an e-complete class.

(5) For loss function (1.1), the set of monotone procedures having the interval property (a.e.) is an e-complete class.

As consequences of results (3), (4) and (5) we derive an analogue to the Rao-Blackwell theorem and construction. Furthermore it is shown that two observations are strictly better than one. In Section 2 we give results on monotone procedures and complete classes. In Section 3 we discuss procedures with the interval property. Section 4 contains some remarks and discussion regarding lower confidence bounds.

2. Monotone Confidence Sets

Throughout, we take the parameter space, Θ, and the range of T, τ, to be intervals in $\mathbb{R}$, and assume in all following expressions that $\theta \in \Theta, t \in \tau$. The SMLR property implies that $f_T(t|\theta) > 0$ a.e. (λ_τ) where λ_τ denotes Lebesque measure on τ. We begin with the definition of monotonicity.

Definition 2.1. The measurable function $\psi(\theta|t)$ is *monotone* if for any $\theta < \theta', t < t'$ the three inequalities

$$\psi(\theta|t) < 1 \qquad \psi(\theta'|t) > 0 \qquad \psi(\theta|t') > 0 \tag{2.1}$$

imply $\psi(\theta'|t') = 1$.

Thus, for nonrandomized confidence sets, a confidence set is nonmonotone if there exist $\theta < \theta', t < t'$ such that $\psi(\theta|t') = \psi(\theta'|t) = 1$ while $\psi(\theta|t) = \psi(\theta'|t') = 0$.

There are three other logically equivalent versions of Definition 2.1. Namely that for $\theta < \theta', t < t'$; $[\psi(\theta|t) < 1,\ \psi(\theta'|t) > 0,\ \psi(\theta'|t') < 1] \Rightarrow \psi(\theta|t') = 0$; or $[\psi(\theta|t') > 0,\ \psi(\theta'|t') < 1,\ \psi(\theta|t) < 1] \Rightarrow \psi(\theta'|t) = 0$; or $[\psi(\theta|t') > 0,\ \psi(\theta'|t') < 1,\ \psi(\theta'|t) > 0] \Rightarrow \psi(\theta|t) = 1$.

Now let G denote a prior (generalized prior) on the parameter space Θ, let $p(\theta|t)$ denote the posterior density of $\theta|T = t$, and $f(t)$ the marginal density of T. The prior G can be written as $G = G_1 + G_0$, where G_1 is singular with respect to Lebesgue measure and G_0 is absolutely continuous. G_1 is supported on S_1, a set of Lebesgue measure 0. Let $g_0(\theta) = dG_0/d\theta$ be the density of G_0.

Theorem 2.2. *For loss function* (1.1) *there is a version of the Bayes (generalized Bayes) procedure which is monotone. Furthermore, any Bayes procedure must agree with the monotone Bayes procedure a.e.* $(G \cdot \lambda_\tau)$, *where* $(G \cdot \lambda_\tau)$ *denotes the indicated product measure.*

Proof. Integrate (1.2) with respect to the prior $G(\theta)$, interchange the order of integration, write the absolutely continuous part of the joint density of (t, θ) as $p(\theta|t)f(t)$, and note that the expected risk is minimized for each t essentially by

$$\psi(\theta|t) = \begin{cases} 1, & \text{if } \theta \in S_1, \\ 1, & \text{if } p(\theta|t) > b, \\ 0, & \text{if } p(\theta|t) < b, \\ \text{anything}, & \text{if } p(\theta|t) = b, \end{cases} \tag{2.2}$$

where

$$p(\theta|t) = \frac{g_0(\theta) f_T(t|\theta)}{\int g_0(u) f_T(t|u)du + \int f_T(t|u) G_1(du)}.$$

Furthermore, every Bayes procedure must satisfy (2.2) a.e. $(G \cdot \lambda_\tau)$. Note that for $\psi(\theta|t)$ determined by (2.2), $\psi(\theta|t) < 1$ and $\psi(\theta'|t) > 0$ can only occur if $\theta \notin S_1$ and $p(\theta|t) \le b \le p(\theta'|t)$. Since $p(\theta|t') < p(\theta'|t')$ by the SMLR property of $f_T(t|\theta)$, $\psi(\theta|t') > 0 \Rightarrow p(\theta|t') \ge b \Rightarrow p(\theta'|t') > b \Rightarrow \psi(\theta'|t') = 1$.

In order to develop complete classes for loss function (1.1) we need to introduce the notion of essential admissibility. This is because there are no nontrivial admissible procedures for the loss functions we study. To see this, note that given $\psi(\theta|t)$, if θ_0 is any value such that $E_{\theta_0}(\psi(\theta_0|T)) < 1$, then $\psi'(\theta|t) = \psi(\theta|t)$ for $\theta \ne \theta_0$, and $\psi'(\theta|t) = 1$ if $\theta = \theta_0$ is better than ψ.

Definition 2.3. The confidence set estimator $\psi(\theta|t)$ is *essentially admissible* (*e-admissible*) if there does not exist a $\psi'(\theta|t)$ such that

$$R(\theta, \psi') \le R(\theta, \psi) \text{ a.e. } \theta \quad (\text{w.r.t. } \lambda_\Theta) \tag{2.3}$$

and

$$\lambda_\Theta\{\theta : R(\theta, \psi') < R(\theta, \psi)\} > 0. \tag{2.4}$$

An e-essentially complete class is such that for any procedure ψ outside the class, there exists one in the class, say ψ', such that (2.3) holds. Now let the set of confidence set estimators (decision rules) be $D = \{\psi : \Theta \times T \to [0,1], \psi$ jointly measurable $\}$. This can be considered as a subset of $L_\infty(\Theta \times T, d\theta \times dt)$ under the weak* topology. The subset is closed and bounded and hence compact.

Theorem 2.4. *For loss function* (1.1)*, the collection of e-admissible confidence set estimators is the minimal e-complete class i.e. every confidence set not in the class is dominated by an e-admissible procedure.*

Proof. The proof is given in the Appendix.

Given the contents of the proof we also get the following as a corollary.

Corollary 2.5. *For loss function* (1.1) *the limits (in D) of sequences of Bayes confidence sets are an e-essentially complete class.*

Proof. The proof is given in the Appendix.

3. Interval Property (Property I)

Definition 3.1. The confidence set estimator $\psi(\theta|t)$ has *the interval property* (a.e.) if (except for a θ-set of measure zero), ψ is nonrandomized and $\{t : \psi(\theta|t) = 1\}$ is an interval for every θ.

For the remainder of this section we assume that the distribution $f_T(t|\theta)$ is strictly totally positive of order at least 3 (STP_3). (Note that exponential family densities qualify.) See Karlin (1968) for the definition of TP_3 and STP_3. Karlin (1968), Section 5.3.1 asserts that STP_3 implies SVR_3 where SVR_3 stands for the strict variation reducing property of $f_T(t|\theta)$. See Brown, Johnstone and MacGibbon (1981), Section 2 for the definition of SVR_3.

Theorem 3.1. *For the loss function* (1.3)*, and hence for the loss function* (1.1)*, the set of procedures having the interval property is an e-complete class.*

Proof. Let ψ be any procedure. Define ψ' as a procedure with property I as follows:

$$\psi'(\theta|t) = \begin{cases} 1, & \text{if } a(\theta) < t < b(\theta), \\ 0, & \text{otherwise}, \end{cases} \tag{3.1}$$

where $a(\theta), b(\theta)$ are the unique solutions to

$$\int_{a(\theta)}^{b(\theta)} f_T(t|\theta)dt = \int \psi(\theta|t) f_T(t|\theta)dt, \tag{3.2}$$

and

$$\int_{a(\theta)}^{b(\theta)} \frac{\partial}{\partial\theta} f_T(t|\theta)dt = \int \psi(\theta|t)\frac{\partial}{\partial\theta} f_T(t|\theta)dt. \tag{3.3}$$

(The STP_3 property implies that $a(\theta), b(\theta)$ are uniquely determined by (3.2)-(3.3). Measurability of $a(\cdot)$ and $b(\cdot)$ then follows as an application of the Stschegolkow selection theorem. (See, e.g. Brown and Purves (1973).)

Now let ρ be a fixed point in Θ. Let $\chi_A(\cdot)$ be the indicator function for a set A and note that $[\psi(\rho|\cdot) - \chi_{(a(\rho),b(\rho))}(\cdot)]$, has at most two sign changes, from positive to negative, to positive. Furthermore, from (3.3),

$$\frac{\partial}{\partial\theta}\int [\psi(\rho|t) - \chi_{(a(\rho),b(\rho))}(t)] f_T(t|\theta)dt|_{\theta=\rho} = 0. \tag{3.4}$$

Also, from (3.2),

$$\int [\psi(\rho|t) - \chi_{(a(\rho),b(\rho))}(t)] f_T(t|\rho)dt = 0. \tag{3.5}$$

Hence the SVR_3 property implies

$$\int [\psi(\rho|t) - \chi_{(a(\rho),b(\rho))}(t)] f_T(t|\theta)dt \geq 0, \tag{3.6}$$

with strict inequality for $\theta \neq \rho$, unless $\psi(\theta|\cdot) = \chi_{(a(\theta),b(\theta))}(\cdot)$ a.e. For the vector risk (1.4), the second components for ψ and ψ' are the same by virtue of (3.5), but (3.6) implies that

$$R_1^v(\theta,\psi) \geq R_1^v(\theta,\psi'), \tag{3.7}$$

with strict inequality unless $\psi = \psi'$ a.e. on R^2.

A rationale for Theorem 3.1 comes from the duality of confidence sets and testing, along with admissibility implications derived by Cohen and Strawderman (1973). For STP_3 densities, interval tests of a simple null hypothesis against a two sided alternative form a complete class. The criteria of size and power for testing are in one to one correspondence with one minus probability of true coverage and probability of covering false values respectively. In fact, a confidence set is admissible using these criteria as a vector risk if and only if the corresponding family of tests are. Cohen and Strawderman (1973) demonstrate that inadmissibility for a confidence set in terms of these risk criteria implies inadmissibility if the second of the criteria is replaced by expected length. Thus any noninterval procedure would be inadmissible by either set of risk criteria. The proof of Theorem 3.1 has the advantage that it is constructive.

Corollary 3.2. *For loss function* (1.3) *and hence for* (1.1), *nonrandomized confidence set estimators are an e-complete class.*

Corollary 3.3. *All Bayes procedures have property I a.e. and are nonrandomized.*

Lemma 3.4. *Let ψ_1 and ψ_2 be any two confidence set estimators. Then there exists a confidence set estimator ψ, such that*

$$R_1^v(\theta, \psi) < (R_1^v(\theta, \psi_1) + R_1^v(\theta, \psi_2))/2 \tag{3.8}$$

for every θ such that $\lambda\{t \in \tau | \psi_1(\theta|t) \neq \psi_2(\theta|t)\} > 0$,
and

$$R_2^v(\theta, \psi) = [R_2^v(\theta, \psi_1) + R_2^v(\theta, \psi_2)]/2. \tag{3.9}$$

Proof. Consider the procedure defined by $\psi' = (\psi_1 + \psi_2)/2$. This procedure has risk $[R(\theta, \psi_1) + R(\theta, \psi_2)]/2$. Unless ψ and ψ' are equivalent, the procedure ψ' does not have property I since it is a randomized procedure. The lemma now follows via the strict inequality in (3.7).

Corollary 3.5. *If ψ is e-admissible and $R(\theta, \psi) = R(\theta, \psi')$ a.e. then ψ and ψ' are equivalent. Consequently any e-essentially complete class is also e-complete.*

Proof. Immediate from Lemma 3.4.

Before stating the last complete class theorem we need a remark. We say that ψ is monotone (a.e.) if there is a set N of measure zero in $\mathbb{R}^2$ such that ψ satisfies conditions (2.1) everywhere on $\mathbb{R}^2 - N$ (i.e. for all (θ, x) such that $(\theta, x) \notin N$). If ψ is monotone (a.e.) and has property I (a.e.) then it is possible to construct a version which is monotone and has property I. To see this, let N be the exceptional set off which ψ is monotone with property I. Let $N' = \{\theta : \lambda\{x : (\theta, x) \in N\} > 0\}$. Note that N' is measurable and $\lambda(N') = 0$. Since ψ has property I, $\{x : \psi(\theta|x) = 1\}$ must be an interval a.e. Let $M(\theta)$ be the closed interval such that $\lambda\{M(\theta)\Delta[x : \psi(\theta|x) = 1]\} = 0$. Here Δ means symmetric difference. Now define

$$\psi'(\theta|x) = \begin{cases} 1, & \text{if } \theta \notin N' \text{ and } \ x \in M(\theta), \\ 1, & \text{if } \theta \in N', \\ 0, & \text{if } \theta \notin N' \text{ and } x \notin M(\theta). \end{cases}$$

One can check that ψ' has the desired property.

Now we prove

Theorem 3.6. *For loss function* (1.1)*, the set of monotone procedures having property I a.e. is an e-complete class.*

Proof. As a consequence of Corollaries 3.2 and 3.5 we need only show that if ψ has property I and is the limit in D of a sequence of Bayes procedures, then ψ is also monotone. Hence assume ψ has property I and there exist Bayes procedures $\psi_i, i = 1, 2, \ldots$ such that $\psi_i \to \psi$ in D (i.e. $\psi_i \to \psi$ in the weak* sense meaning $\int g(t,\theta)[\psi(\theta|t) - \psi_i(\theta|t)]d\theta dt \to 0$, for every finitely integrable g on $\mathbb{R}^2$. Since ψ is nonrandomized and since $0 \le \psi_i \le 1$, this convergence occurs if and only if $\psi_i \to \psi$ in measure (i.e. $\lambda(\{(t,\theta) : |\psi(\theta|t) - \psi_i(\theta|t)| > \epsilon\}) \to 0$, for every $\epsilon > 0$). However, then there exists a subsequence $\{i'\}$ such that $\psi_{i'} \to \psi$ a.e. Since each ψ_i is monotone and has property I it is then easy to check that ψ is monotone a.e. with property I a.e. The remark preceding Theorem 3.6 claims that ψ can be modified to be exactly monotone with property I.

Corollary 3.2 yields a constructive analogue of the Rao-Blackwell theorem. Take any confidence set procedure not based only on a sufficient statistic. By considering the projection of the procedure onto the space of the sufficient statistic a randomized procedure, with the same risk function as the original procedure, is determined. The construction in Theorem 3.1 applied to the above randomized procedure yields a better procedure based on the sufficient statistic.

Similar reasoning yields the fact that a procedure based on two observations is strictly better than a procedure based on just one. This follows since one can randomize between two separate procedures each based on one observation, thereby arriving at a randomized procedure whose risk equals the risk of the procedure based on one observation. Again the construction in Theorem 3.1 yields a strictly better procedure based on two observations.

We note that the Rao-Blackwell analogue and the fact that two observations are better than one work for the vector risk function (1.4). This is satisfying in the sense that in confidence estimation one frequently seeks confidence procedures whose probability of coverage exceeds a given value. The construction of Theorem 3.2 entails matching the second component of the vector risk, and so the better procedure satisfies the probability of coverage constraint.

4. Lower Confidence Bounds

In this section we make some observations in connection with lower confidence bounds. For the statistical model given in Section 1 we consider lower confidence bounds for the loss function

$$B(\theta - a)^+ + [1 - I_{(0,\infty)}(\theta - a)], \tag{4.1}$$

where a is the action (the lower confidence bound) and B is a predetermined constant. (Variations on (4.1), to accommodate scale parameters or other special problems can be made.)

Our first observation is that the loss function in (4.1) is bowl shaped and also satisfies the conditions in BCS (1976). This means that the monotone procedures are a complete class and all results of BCS (1976) apply in this situation.

Our next observation is in connection with the model where $T \sim N(\theta, \sigma^2/n)$, σ^2 known or unknown. Here it is of interest to derive the best invariant lower confidence bound. A discussion of invariant confidence bounds is in Casella and Berger (1990). First, let σ^2 be known and equal to one. The best invariant lower confidence bound is $(T + C_n/\sqrt{n})$ where C_n is chosen to minimize

$$BE_\theta(\theta - [T + C_n/\sqrt{n}])^+ + 1 - P_\theta((\theta - [T + C_n/\sqrt{n}]) > 0). \tag{4.2}$$

By changing variables in (4.2) and differentiating with respect to C_n the appropriate C_n is the solution to the following equation:

$$\varphi(C_n) = (B/\sqrt{n})\Phi(-C_n), \tag{4.3}$$

where φ and Φ are the p.d.f. and c.d.f. respectively of a standard normal variable. We note that for $B/\sqrt{n}$ sufficiently small, C_n is the unique negative solution to (4.3).

It is interesting to note that $C_n \to -\infty$ as $n \to \infty$ at the rate $(\log n)^{1/2}$ which means that $C_n/\sqrt{n} \to 0$ as $n \to \infty$ at the rate $(\log n/n)^{1/2}$. The facts that $C_n \to -\infty$ and $C_n/\sqrt{n} \to 0$ imply that the risk of the procedure tends to zero as $n \to \infty$. This is consistent with the idea that in a loss function approach to a problem, the risk tends to zero as the sample size tends to infinity. This is unlike typical approaches to confidence estimation where one minus probability of coverage does not usually go to zero. The particular rate of $(\log n/n)^{1/2}$ is the same as found in confidence sequences. See Robbins (1970).

When σ^2 is unknown, we consider the loss function

$$B[(\theta - a)^+/\sigma] + [1 - I_{(0,\infty)}((\theta - a)/\sigma)] \ . \tag{4.4}$$

The best invariant lower confidence bound is of the form $T + C_n s/\sqrt{n}$. (See Casella and Berger (1990) for the discussion of the group leaving the problem invariant.) Here s^2 is such that $\nu s^2/\sigma^2$ is χ^2_ν and is independent of T. A calculation similar to that performed in the case σ^2 known yields the optimal value of C_n to be the solution of

$$\begin{aligned}&(B/\sqrt{n}) \int_0^\infty \{\Phi(-C_n u^{1/2}) e^{-u/2}\, u^{\frac{\nu+1}{2}-1}/\sqrt{2\pi}\ \Gamma(\nu/2) 2^{\nu/2}\} du \\ &= (1/(C_n^2+1))^{\frac{\nu+1}{2}}\ \Gamma(\frac{\nu+1}{2})\sqrt{2}/\ \Gamma(\nu/2)\sqrt{2\pi}.\end{aligned} \tag{4.5}$$

As C_n ranges from $-\infty$ to 0 we note the left hand side decreases while the right hand side increases so a unique solution exists provided $B/\sqrt{n}$ is sufficiently small.

Acknowledgement

This research was supported by NSF Grant DMS-9310228 and DMS-9112784.

Appendix

Theorem 2.4. *For loss function (1.1), the collection of e-admissible confidence set estimators is the minimal e-complete class, i.e. every confidence set estimator not in the class is dominated by an e-admissible procedure.*

Proof. Since D, the set of decision rules can be regarded as a closed subset of L_∞ under the weak* topology, D is compact.

Now, for every $-\infty < \gamma_1 < \gamma_2 < \infty$, define

$$R^*((\gamma_1,\gamma_2),\psi) = [1/(\gamma_2-\gamma_1)]\int_{\gamma_1}^{\gamma_2} R(\theta,\psi)d\theta. \tag{A.1}$$

We say the procedure ψ is e^*-better than ψ' if $R^*((\gamma_1,\gamma_2),\psi) \le R^*((\gamma_1,\gamma_2),\psi')$ for all $(\gamma_1,\gamma_2) \in \mathbb{R}^2$ and if strict inequality holds for some (γ_1,γ_2). Clearly, ψ is e^*-admissible if and only if it is e-admissible. Hence an e^*-complete class is e-complete and conversely.

For each (γ_1,γ_2) consider the map $R^*((\gamma_1,\gamma_2),\cdot) : D \to [0,\infty]$, defined by $R^*((\gamma_1,\gamma_2),\psi)$. This is a lower semi-continuous map. See Brown (1977). Hence, by standard reasoning, (see for example Ferguson (1967), p. 87), the e^*-admissible procedures are an e^*-complete class.

Corollary 2.5. The limits (in D) of sequences of Bayes confidence sets are an e-essentially complete class.

Proof. The "standard" conclusion is that limits of sequences of Bayes procedures for simple (i.e. finitely supported) priors over $\{(\gamma_1,\gamma_2) \in \mathbb{R}^2\}$ are an e^*-essentially complete class. Hence they are also an e-essentially complete class. However, if a prior G^* gives mass α_i to $(\gamma_{1i},\gamma_{2i})$, then it is equivalent to consider a prior in the original problem having density $\Sigma(\alpha_i/(\gamma_{2i}-\gamma_{1i}))\chi_{(\gamma_{1i},\gamma_{2i})}(\theta)$, i.e. the two priors have the same Bayes procedures (as well as the same Bayes risk). Hence the corollary follows.

References

Brown, L. D. (1977). Closure theorems for sequential-design processes. In Statistical Decision Theory and Related Topics II, 57-91, Academic Press.

Brown, L. D. and Purves, R. (1973). Measurable selections of extrema. Ann. Statist. **1**, 902-912.

Brown, L. D., Johnstone, I. M. and MacGibbon, K. B. (1981). Variation diminishing transformations: A direct approach to total positivity and its statistical applications. J. Amer. Statist. Assoc. **76**, 824-832.

Brown, L. D., Cohen, Arthur and Strawderman, W. E. (1976). A complete class theorem for strict monotone likelihood ratio with applications. Ann. Statist. **4**, 712-722.

Casella, G. and Berger, R. L. (1990). Statistical Inference. Wadsworth, Pacific Grove, California.

Cohen, A. and Sackrowitz, H. B. (1984). Decision theory results for vector risks with applications. Statistics and Decisions. Supplement Issue No. 1, 159-176.

Cohen, A. and Strawderman, W. E. (1973a). Admissibility implications for different criteria in confidence estimation. Ann. Statist. **1**, 363-366.

Cohen, A. and Strawderman, W. E. (1973b). Admissible confidence interval and point estimation for translation or scale parameters. Ann. Statist. **1**, 545-550.

Ferguson, T. S. (1967). Mathematical Statistics: A Decision Theoretic Approach. Academic Press, New York.

Joshi, V. M. (1970). Admissibility of invariant confidence procedures for estimating a location parameter. Ann. Math. Statist. **41**, 1568-1581.

Karlin, S. (1956). Decision theory for Polya type distributions. Case of two actions, I. Proc. 3rd Berkeley Symp. Probab. Statist. **1**, 115-129, University of California Press.

Karlin, S. (1968). Total Positivity. Stanford University Press, Stanford, California.

Robbins, H. E. (1970). Statistical methods related to the law of the iterated logarithm, **41**, 1397-1409.

Winkler, R. L. (1972). A decision-theoretic approach to interval estimation. J. Amer. Statist. Assoc. **67**, 187-191.

Statistics Department, University of Pennsylvania, Philadelphia, PA 19104-6302, U.S.A.

Department of Statistics, Rutgers University, New Brunswick, NJ 08903, U.S.A.

(Received November 1992; accepted August 1994)