APPROXIMATIONS OF BAYES DECISION PROBLEMS: THE EPIGRAPHICAL APPROACH

by

J. O. Berger and G. Salinetti
Purdue University
Universitá di Roma "La Sapienza"

Technical Report #94-21C

Department of Statistics
Purdue University

August 1994

APPROXIMATIONS OF BAYES DECISION PROBLEMS: THE EPIGRAPHICAL APPROACH

by

J. O. Berger and G. Salinetti

Purdue University — Universitá di Roma "La Sapienza"

West Lafayette, IN, USA — Rome, Italy

Summary

Solving Bayesian decision problems usually requires approximation procedures, all leading to study the convergence of the approximating infima. This aspect is analysed in the context of epigraphical convergence of integral functionals, as minimal context for convergence of infima. The results, applied to the Monte Carlo importance sampling, give a necessary and sufficient condition for convergence of the approximations of Bayes decision problems and sufficient conditions for a large class of Bayesian statistical decision problems.

Keywords: Bayesian decisions, epi-convergence, integral functionals, Monte Carlo importance sampling.

1. Introduction and Problem Setting

Decision problems under uncertainty from a Bayesian perspective require to solve the minimization problem.

$$\inf_{a\in A}\int L(a,\theta)\Pi(d\theta) \tag{1.1}$$

where A is the space of possible *actions* or decisions, θ is the "unknown" quantity affecting the decision process, commonly called *state of nature*, taking values in a given space Θ; the function $(a,\theta)\to L(a,\theta)$ expresses the *loss* incurred when the chosen action is a and the true state of nature is θ. The function Π is a probability measure on the class of events $\mathcal{B}(\Theta)$ of Θ; it can be the prior probability measure on $\mathcal{B}(\Theta)$ or, in statistical decision problems, the posterior probability measure on $\mathcal{B}(\Theta)$ after seeing the data: in the last case Π then combines the prior distribution on Θ with the likelihood function according the Bayes theorem; in view of the approximation procedures to solve (1.1) discussed here, it is not relevant to distinguish between Π representing a prior or a posterior distribution.

Even if approximation and convergence results presented here can be extended to more general situations, we assume that Θ is a finite dimensional euclidean space, $\mathcal{B}(\Theta)$ the Borel field on Θ and Π a probability measure on $\mathcal{B}(\Theta)$.

The decision space A is metric separable and, as in all problems of interest, the loss function $(a,\theta) \to L(a,\Theta)$ is

A1 lower bounded on $A \times \Theta$, without loss of generality: $L(a,\theta) \geq 0$, $\forall (a,\theta) \in A \times \Theta$,

A2 lower semicontinuous on $A \times \Theta$,

A3 measurable in θ for every $a \in A$.

Many problems in Bayesian analysis, and in particular the minimization problem (1.1), rarely can be solved explicitly. In fact, in all but very specific problems, solving the problem requires analytic or numerical approximations. The developing of suitable approximation techniques has focused on the Monte Carlo approach. This can be pursued estimating $\int L(a,\theta)\Pi(d\theta)$ by using samples $\theta_1, \theta_2, \dots, \theta_n, \dots$ drown from the probability measure Π when possible and not expensive; or by using reweighted samples drawn from some other appropriately chosen distributions P such that $\int L(a,\theta)\Pi(d\theta) = \int L(a,\theta)w(\theta)P(d\theta)$, as in the versions of the Monte Carlo importance sampling.

In these approximation procedures we have a sequence of random vectors $\{\theta_1, \theta_2, \dots\}$ all defined on the same probability space, say $(\Omega, \mathcal{A}, \mu)$; $f(a) = \int L(a,\theta)\Pi(d\theta)$ is approximated by

$$f_n(a,\omega) = \frac{1}{n}\sum_{i=1}^{n} g(a, \theta_i(\omega)) \tag{1.2}$$

and the minimization problem (1.1) by

$$\min_{a \in A} f_n(a,\omega). \tag{1.3}$$

In (1.2) $g(a,\theta)$ can be simply $L(a,\theta)$ as in the direct sampling from Π; or, as in the Monte Carlo importance sampling, $g(a,\theta) = L(a,\theta)w(\theta)$.

The major convergence question entailed by solving (1.1) through the approximations (1.3) concerns the almost sure (a.s.) convergence of the stochastic infima

$$(1.4) \qquad \inf f_n(a,\cdot) \to \inf f(a) \quad \text{a.s., as } n \to \infty$$

and the related convergence of the optimal solutions.

Under mild conditions the approximation scheme (1.2) provides the a.s. convergences

$$(1.5) \qquad f_n(a,\cdot) \to f(a) \text{ a.s.,} \qquad \forall\, a \in A.$$

From (1.5) it is easy to derive

$$(1.6) \qquad \limsup_{n\to\infty} \inf_{a\in A} f_n(a,\cdot) \le \inf_{a\in A} f(a) \text{ a.s..}$$

Thus the major mathematical aspect for answering the convergence question (1.4) is to state the minimal set of conditions which guarantee

$$(1.7) \qquad \inf_{a\in A} f(a) \le \liminf_{n\to\infty} \inf_{a\in A} f_n(a,\cdot) \text{ a.s..}$$

In view of that observe that for every realization, say $\{\theta_1(\omega), \theta_2(\omega), \ldots\}$ of the sequence $\{\theta_n,\ n = 1,2,\ldots\}$ we have

$$f_n(a,\omega) = \int g(a,\theta) P_n(d\theta,\omega)$$

where $P_n(\cdot,\omega)$ is the empirical probability measure on $\mathcal{B}(\Theta)$ determined by $(\theta_1(\omega), \ldots, \theta_n(\omega))$.

Under rather general conditions, in the Monte Carlo approximation schemes, a.s. on Ω, i.e. for all $\omega \in \Omega \backslash N$, $\mu(N) = 0$, the sequence of probability measures $\{P_n(\cdot,\omega)\}$ weakly converges to P:

$$(1.8) \qquad P_n(\cdot,\omega) \xrightarrow{w} P.$$

In view of all the above, in particular taking into account of (1.5) and (1.8), the key question about a.s. convergence of the stochastic infima (1.4) can be rephrased in the following setting: *Let* $\inf_{a\in A} f(a) = \inf_{a\in A} \int g(a,\theta) P(d\theta)$ *be a minimization problem approximated* *by the sequence* $\inf_{a\in A} f_n(a,\omega) = \inf_{a\in A} \int g(a,\theta)P_n(d\theta,\omega)$, $n = 1,2,\ldots$, with $f_n(a,\cdot) \to f(a)$ a.s. *for all* $a \in A$ *and* $P_n(\cdot) \xrightarrow{w} P$ a.s.; *under which minimal conditions we have* $\inf_{a\in A} f(a) \leq \liminf_{n\to\infty} \inf_{a\in A} f_n(a,\cdot)$ a.s.?

This question is fully answered in section 2: a necessary and sufficient condition for the convergence of stochastic infima is given and sufficient conditions are also provided. In Section 3 the result is applied to the Monte Carlo importance sampling for Bayesian decision problems. This problem has been considered in the Bayesian literature under particular assumptions [7]. The result of Section 2 enlarges the applicability of Monte Carlo importance sampling to Bayesian decision problems.

2. Convergence of Infima of Integral Functionals

For a family $\{f;\ f_n,\ n = 1,2,\ldots\}$ of (possibly extended) real valued functions defined on a metric separable space A, the convergence of infima

$$\inf_{a\in A} f_n(a) \to \inf_{a\in A} f(a) \tag{2.1}$$

is strictly related to the epigraphical convergence of the sequence $\{f_n\}$ to f; (for an extended presentation of epigraphical convergence see [1] and references therein).

Epigraphical convergence, or *epi-convergence*, of the sequence $\{f_n\}$ to f, denoted $f_n \xrightarrow{\text{epi}} f$, means that the following two conditions are satisfied at every $a \in A$:

$\forall\, a \in A,\ \exists\, a_n \to a$ such that $\limsup_{n\to\infty} f_n(a_n) \leq f(a)$,

$\forall\, a \in A,\ \forall\, a_m \to a,\ \forall$ subsequence $\{n_m\} \subset \{n\}$, it is $f(a) \leq \liminf_{m\to\infty} f_{n_m}(a_m)$.

We refer to these relations respectively as

$$\text{epi-}\limsup\ f_n \leq f \quad \text{and} \quad f \leq \text{epi-}\liminf\ f_n.$$

With respect to the convergence of infima (2.1), it is immediate to verify that

$$\text{epi-}\limsup\ f_n \leq f \Rightarrow \limsup\ \inf f_n \leq \inf\ f. \tag{2.2}$$

The opposite relation $\inf\ f \leq \liminf\ \inf\ f_n$ is strictly related to the relation $f \leq$ epi-$\liminf\ f_n$ but in general not implied by it. However it is known, and easy to show, that

$$f \leq \text{epi-}\liminf\ f_n \Rightarrow \inf\ f \leq \liminf\ \inf\ f_n \tag{2.3}$$

if, for any $\varepsilon > 0$, the sequence $\{\inf\ f_n\}$ has a *bounded sequence of ε-optimal solutions*; this means that for any $\varepsilon > 0$ there exist a compact subset $K_\varepsilon \subset A$ and a sequence $\{a_n\}$ such that for all n

$$f_n(a_n) < \inf_{a \in A} f_n(a) + \varepsilon \quad \text{and} \quad a_n \in K_\varepsilon. \tag{2.4}$$

2.5 Remark: It can actually be shown that in presence of epi-convergence $f_n \xrightarrow{\text{epi}} f$, condition (2.4) is also necessary; i.e. if $f_n \xrightarrow{\text{epi}} f$ then $\inf\ f_n \to \inf\ f$ if and only if (2.4) holds.

In the case of interest here, f_n and f are integral functionals defined by

$$f(a) = \int g(a,\theta)P(d\theta);\ f_n(a) = \int g(a,\theta)P_n(d\theta),\ n = 1,2,\ldots$$

where $\{P;\ P_n,\ n = 1,2,\ldots\}$ is a family of probability measures on $(\Theta, \mathcal{B}(\Theta))$ and the sequence $\{P_n\}$ weakly converges to P, $P_n \xrightarrow{w} P$.

The epigraphical convergence of $\{f_n\}$ to f and the related convergence of their infima rely, as natural to expect, on the convergence $P_n \xrightarrow{w} P$ and on the properties of g.

The integrand $(a,\theta) \to g(a,\theta)$ is assumed to be

A1 lower bounded on $A \times \Theta$; without loss of generality: $g(a,\theta) \geq 0,\ \forall\ (a,\theta) \in A \times \Theta$,

A2 lower semicontinuous on $A \times \Theta$,

A3 measurable in θ for each $a \in A$.

The epigraphical convergence of integral functionals is widely analyzed in [4] and [5], in the last with special reference to Bayesian decision problems. In particular [5, Proposition 2.13] we have:

2.6 Proposition: *If g satisfies* A1–A3 *and $P_n \xrightarrow{w} P$ then*

$$f \leq \text{epi-}\liminf_{n \to \infty}\ f_n.$$

2.7 Remark: The opposite relation epi-lim sup $f_n \leq f$ does not follow in general from $P_n \xrightarrow{w} P$; here we simply register that if $f_n \to f$ pointwise on A then $f_n \xrightarrow{\text{epi}} f$. For, just observe that pointwise convergence, and in particular $\limsup f_n(a) \leq f(a)$, $\forall\, a \in A$, implies that epi-$\limsup f_n \leq f$; this, together with Proposition 2.6 gives the epi-convergence $f_n \xrightarrow{\text{epi}} f$.

The convergence of infima of integral functionals, according (2.3), depends on the existence of bounded sequences of ε-optimal solutions. It is obvious that if A is compact then the epigraphical convergence $f_n \xrightarrow{\text{epi}} f$ implies the convergence of the infima $\inf f_n \to \inf f$.

When A is not compact the existence, for every $\varepsilon > 0$, of a bounded sequence of ε-optimal solutions, or more specifically the convergence of infima, depends on the behaviour of g on the tails.

2.8 Proposition: *Suppose that g, in addition to* A1–A3, *satisfies the following condition: for every compact subset T of Θ there exists a compact subset K of A such that for all $a \notin K$*

$$g(a,\theta) \geq \inf_{a \in A} f, \quad \forall\, \theta \in T. \tag{2.9}$$

If $P_n \xrightarrow{w} P$ then

$$\inf_{a \in A} f \leq \liminf_{n \to \infty} \inf_{a \in A} f_n.$$

Proof: Arguing by contradiction, assume that $\liminf \inf f_n < \inf f - \varepsilon$, for some $\varepsilon > 0$. Then there exist a subsequence $\{n_m\}$ and a corresponding sequence $\{a_m\}$ of $\frac{\varepsilon}{2}$-optimal solutions such that

$$f_{n_m}(a_m) < \inf f_{n_m} + \frac{\varepsilon}{2} < \inf f - \frac{\varepsilon}{2}, \ \forall\, m. \tag{2.10}$$

Let $\delta > 0$ be such that $\delta \cdot \inf f < \frac{\varepsilon}{4}$ and let T be a compact subset of Θ such that $P_n(T) > 1 - \delta$ for all n; let K be the corresponding compact subset of A according (2.9).

If $a_m \in K$ for all but finitely many m then the sequence $\{a_m\}$ possesses a convergent subsequence, say $a_{m'} \to a'$. In this case, Proposition 2.6, through (2.10), gives the contradiction

$$\inf f \leq f(a') \leq \liminf f_{n_{m'}}(a_{m'}) \leq \inf f - \frac{\varepsilon}{2}.$$

On the other hand if there exists a subsequence of $\{a_m\}$, say $\{a_{m'}\}$, such that $a_{m'} \notin K$, $\forall\, m'$, then by (2.10) and (2.9) we have

$$\inf\ f - \frac{\varepsilon}{2} > f_{n_{m'}}(a_{m'}) \geq \inf\ f \cdot P_{n_{m'}}(T) \geq \inf\ f - \frac{\varepsilon}{4}.$$

This is a contradiction and it completes the proof.

As a consequence of Proposition 2.8, more interesting from the applications viewpoint, we have the following result valid for A normed space.

2.11 Proposition: *Suppose that g, in addition to* A1–A3, *satisfies the following condition: for every compact subset T of Θ*

(2.12) $$\lim_{||a||\to\infty} g(a,\theta) > \inf_{a\in A}\ f(a) \quad \text{uniformly on } T;$$

if $P_n \xrightarrow{w} P$ then

$$\inf_{a\in A} f(a) \leq \liminf_{n\to\infty}\ \inf_{a\in A}\ f_n(a).$$

Proof: Just observe that condition (2.12) states that there exists $M > 0$ such that for all a with $||a|| > M$ we have $g(a,\theta) > \inf\ f(a)$, $\forall\, \theta \in T$. The result follows then from the Proposition 2.8 with $K = \{a \in A\colon\ ||a|| \leq M\}$.

2.13 Remark: It can be observed that in Proposition 2.11, if $\limsup\ \inf\ f_n \leq \inf\ f$, condition (2.12) implies that for every $\varepsilon > 0$ there exists K such that for n sufficiently large we have

$$\{a \in A\colon\ f_n(a) < \inf\ f_n + \varepsilon\} \subset K.$$

We consider now the stochastic setting: $(\Omega, \mathcal{A}, \mu)$ is a given probability space; for every n, $\omega \to P_n(\omega)$ maps ω into a probability measure on $\mathcal{B}(\Theta)$, P is a probability measure on $\mathcal{B}(\Theta)$; the integral functionals f_n and f are defined by

$$f(a) = \int g(a,\theta)P(d\theta); f_n(a,\omega) = \int g(a,\omega)P_n(d\theta,\omega), \quad n = 1,2,\dots\ .$$

The objective is the a.s. convergence of the stochastic infima

$$\inf_{a\in A} f_n(a,\cdot) \to \inf_{a\in A} f(a) \text{ a.s., as } n \to \infty.$$

2.14 Proposition: *Let g satisfy conditions* A1–A3. *If*

$$(2.15) \qquad f_n(a,\cdot) \to f(a) \ a.s., \ \forall\, a \in A$$

and

$$(2.16) \qquad P_n(\cdot) \xrightarrow{w} P \ a.s.$$

then

$$\inf_{a\in A} f_n(a,\cdot) \to \inf_{a\in A} f(a) \ a.s.$$

if and only if a.s., *i.e. at every* $\omega \in \Omega \backslash N$, $\mu(N) = 0$, *for every* $\varepsilon > 0$, *the sequence* $\{\inf\ f_n(a,\omega),\ n = 1,2,\ldots\}$ *has a bounded sequence of* ε*-optimal decisions.*

Proof: For the direct part observe first, as stated in Section 1, that (2.15) implies

$$(2.17) \qquad \limsup_{n\to\infty}\ \inf\ f_n(a,\omega) \le \inf\ f(a), \quad \forall\, \omega \in \Omega\backslash N_1,\ \mu(N_1) = 0;$$

condition (2.16) and Proposition 2.6 imply

$$(2.18) \qquad f \le \text{ epi-}\liminf\ f_n(\cdot,\omega),\ \forall\, \omega \in \Omega\backslash N_2,\ \mu(N_2) = 0.$$

Let N_3, $\mu(N_3) = 0$, be such that at each $\omega \in \Omega\backslash N_3$, for every $\varepsilon > 0$ the sequence $\{\inf_{a\in A} f_n(a,\omega),\ n = 1,2,\ldots\}$ has a bounded sequence of ε-optimal solutions. Then by (2.3) we have

$$\inf_{a\in A}\ f(a) \le \liminf_{n\to\infty}\ f_n(a,\omega),\ \forall\, \omega \in \Omega\backslash N_2 \cup N_3.$$

This relation together with (2.17), gives the convergence of infima on $\Omega\backslash(N_1 \cup N_2 \cup N_3)$ and completes the direct part.

For the vice versa let N_0, $\mu(n_0) = 0$, be such that

$$(2.19) \qquad \inf_{a\in A} f_n(a,\omega) \to \inf_{a\in A} f(a),\ \omega \in \Omega\backslash N_0.$$

Let $\delta_k \downarrow 0$ as $k \to \infty$ and for each k let a_k be δ_k-optimal for $\inf_{a\in A}\ f(a)$, i.e.

$$(2.20) \qquad f(a_k) < \inf_{a\in A} f(a) + \delta_k.$$

Let N_k, $\mu(N_k) = 0$, be such that, according (2.15),

$$\forall\, \omega \in \Omega\backslash N_k,\ f_n(a_k,\omega) \to f(a_k) \ \text{ as } n \to \infty; \tag{2.21}$$

Set $N = \bigcup_{k=0}^{\infty} N_k$ so that $\mu(N) = 0$. Let $\omega \in \Omega\backslash N$; for any fixed $\varepsilon > 0$ let k_ε be such that $\delta_{k_\varepsilon} < \frac{\varepsilon}{3}$. For n sufficiently large, by (2.19), (2.20) and (2.21), we have

$$f_n(a_{k_\varepsilon},\omega) < f(a_{k_\varepsilon}) + \frac{\varepsilon}{3} < \inf_{a\in A} f(a) + \frac{2}{3}\varepsilon < \inf_{a\in A} f_n(a,\omega) + \varepsilon.$$

The argument shows that, at every $\omega \in \Omega\backslash N$, for any $\varepsilon > 0$ there exists a_{k_ε} which is ε-optimal for $\inf f_n(a,\omega)$ for all n sufficiently large and completes the proof.

The theorem just proved states the minimal conditions for convergence of stochastic infima. From the operational viewpoint however the result could be not immediately applicable. The existence of bounded sequences of ε-optimal solutions is obviously guaranteed when A itself is compact. Out of this case, in the conditions of Proposition 2.14, condition (2.9) in Proposition 2.8 or condition (2.12) in Proposition 2.11 are sufficient conditions for a.s. convergence of the stochastic infima. The observation wants to point out that conditions (2.9) and (2.12) concern only the integrand g, not the approximations, and they are satisfied by a large class of loss functions in Bayesian decision problems as it will be seen in the next section.

3. Monte Carlo Approximations of Bayesian Decision Problems

Consider the Bayesian decision problem

$$\inf_{a\in A} \int L(a,\theta)\pi(\theta)d\theta \tag{3.1}$$

with L satisfying the assumptions of Section 1 and π density function on Θ of the probability measure Π.

A Monte Carlo method for solving this minimization problem based on the *importance sampling* can be described as follows. Let h be a density function, the *importance function*, with $h(\theta) > 0$ and support including the support of π. Note that

$$f(a) = \int L(a,\theta)\pi(\theta)d\theta = \int L(a,\theta)w(\theta)P(d\theta)$$

where P is the probability measure with density h and $w(\theta) = \pi(\theta)/h(\theta)$.

Suppose that it is possible, in the sense that it is easy and not expensive, to generate a sequence $\{\theta_n,\ n = 1, 2, \dots\}$ of random variables, independent and identically distributed (iid) with common density h. Let $(\Omega, \mathcal{A}, \mu)$ be the underlying probability space to which the sequence $\{\theta_n\}$ is referred. For a fixed n and $\omega \in \Omega$ let

$$f_n(a, \omega) = \frac{1}{n} \sum_{i=1}^{n} L(a, \theta_i(\omega)) w(\theta_i(\omega)) = \int L(a, \theta) w(\theta) P_n(d\theta, \omega)$$

where $P_n(\cdot, \omega)$ is the empirical probability measure on $\mathcal{B}(\Theta)$ determined by $(\theta_1(\omega), \dots, \theta_n(\omega))$. The minimization problems

$$\inf_{a \in A} f_n(a, \omega), \quad n = 1, 2, \dots \tag{3.2}$$

approximate the original problem (3.1) and the key convergence question concerns the convergence of infima

$$\inf_{a \in A} f_n(a, \cdot) \to \inf_{a \in A} f(a) \text{a.s. } . \tag{3.3}$$

The way to choose the importance function h for the approximation procedure is obviously quite relevant and extensive discussions on the choice of h in Bayesian computations can be found for example in [3].

Here we limit ourselves to assume that $w(\theta)$ is lower semicontinuous and bounded on Θ, a condition satisfied in all the problems of interest. Thus the function $(a, \theta) \to g(a, \theta) = L(a, \theta) w(\theta)$ satisfies conditions A1–A3 of Section 2.

In view of the a.s. convergence (3.3) observe first that since the $\{\theta_n\}$ are iid, for each $a \in A$, the random variables $\{g(a, \theta_n) = L(a, \theta_n) w(\theta_n),\ n = 1, 2, \dots\}$ are also iid with expectation $f(a)$; the strong law of large numbers then implies

$$f_n(a, \cdot) \to f(a) \text{ a.s. } \forall\, a \in A. \tag{3.4}$$

Moreover, by Glivenko-Cantelli theorem we have a.s. on Ω

$$P_n(\cdot, \omega) \xrightarrow{w} P. \tag{3.5}$$

It follows from above that the Monte Carlo approximations (3.2) satisfy all the assumptions of Proposition 2.14 and we can conclude:

3.6 Theorem: *The Monte Carlo approximations* $\{\inf_{a\in A} f_n(a,\cdot),\ n=1,2,\dots\}$ a.s. *converge to* $\inf_{a\in A} f(a)$ *if and only if a.s. on* Ω, *i.e. at all* $\omega \in \Omega\backslash N$, $\mu(N)=0$, *for every* $\varepsilon>0$, *the sequence* $\{\inf_{a\in A} f_n(a,\omega),\ n=1,2,\dots\}$ *has a bounded sequence of* ε*-optimal solutions.*

This result and the following sufficient conditions are here obtained as particular case of the more general approximation scheme described in Section 2. However it is relevant to observe that the convergence of the Monte Carlo approximation procedures can also be approached through the epigraphical law of large numbers for random lower semicontinuous functions developed in [2] and [6], once stated that $\{L(\cdot,\theta_n(\cdot))w(\theta_n),\ n=1,2,\dots\}$ is a sequence of random lower semicontinuous functions, independent and identically distributed in the appropriate epigraphical setting.

As application of Propositions 2.8 and 2.11 we have

3.7 Theorem: *Suppose that for every compact subset* T *of* Θ *there exists a compact subset* K *of* A *such that for all* $a\notin K$ *we have*

$$L(a,\theta)w(\theta) \geq \inf_{a\in A} f(a), \quad \forall\, \theta \in T \tag{3.8}$$

then

$$\inf_{a\in A} f_n(a,\cdot) \to \inf_{a\in A} f(a) \text{ a.s. .}$$

On the applicability of (3.8) it can be objected that $\inf\ f(a)$ is unknown; however $\inf\ f$ could be replaced by $f(\bar{a})$ for some $\bar{a}$ where $f(\bar{a})$ is easily computed. For $A=\mathbb{R}^q$ (but more generally in a normed space), (3.8) is implied by the assumption

$$\lim_{||a||\to\infty} L(a,\theta)w(\theta) > r \qquad \text{uniformly on } T$$

with $r > \inf_{a\in A} f(a)$.

In fact condition (3.8) can be replaced by the simpler condition

$$L(a,\theta) \geq \inf_{a\in A} f(a), \quad \forall\, \theta\in T.$$

as next theorem shows.

Observe first that since $\theta \to w(\theta)$ is lower semicontinuous and bounded and $P_n(\cdot,\omega) \xrightarrow{w} P(\cdot)$, $\forall\, \omega \in \Omega \backslash N$, $\mu(N) = 0$, then

$$(3.9) \qquad \int w(\theta)P(d\theta) \leq \liminf \int w(\theta)P_n(d\theta,\omega) \qquad \omega \in \Omega\backslash N,\ \mu(N) = 0;$$

moreover, for every $\delta > 0$ there exists a compact subset T of Θ such that for all n

$$1 - \frac{\delta}{2} < \int_T w(\theta)P(d\theta) \quad \text{and} \quad \int_{T^c} w(\theta)P_n(d\theta,\omega) < \frac{\delta}{2},$$

T^c denoting the complement of T, so that (3.9) gives

$$(3.10) \qquad 1 - \frac{\delta}{2} < \int_T w(\theta)P(d\theta) \leq \liminf_{n\to\infty} \int_T w(\theta)P_n(d\theta,\omega) + \frac{\delta}{2}.$$

We have then

3.11 Theorem: *Suppose that for every compact subset T of Θ there exists a compact subset K of A such that for all $a \notin K$ we have*

$$(3.12) \qquad L(a,\theta) \geq \inf_{a\in A} f(a), \quad \forall\, \theta \in T.$$

Then

$$\inf_{a\in A} f_n(a,\cdot) \to \inf_{a\in A} f(a).$$

Proof: It is sufficient to prove that

$$(3.13) \qquad \inf_{a\in A} f(a) \leq \liminf_{n\to\infty} \inf_{a\in A} f_n(a,\cdot) \ \text{ a.s. .}$$

Arguing by contradiction, suppose that (3.13) does not hold for some $\omega \in \Omega$ where $f_n(\cdot,\omega) \xrightarrow{\text{epi}} f(\cdot)$. Then for some $\varepsilon > 0$ there exist a subsequence $\{n_m\}$ of $\{n\}$ and a sequence $\{a_m\}$ of ε-optimal solutions such that

$$(3.14) \qquad f_{n_m}(a_m,\omega) < \inf_{a\in A} f_{n_m}(a,\omega) + \varepsilon < \inf_{a\in A} f(a) - \varepsilon, \quad \forall\, m.$$

Let $\delta > 0$ be such that $\delta \cdot \inf\, f < \frac{\varepsilon}{2}$ and let T be the compact subset of Θ satisfying (3.10); let K be the corresponding subset of A satisfying (3.12). If $a_m \in K$ for all but finitely many n then the same argument used in the proof of Proposition 2.6 gives a contradiction. On the other hand if there exists a subsequence of $\{a_m\}$, say $\{a_{m'}\}$, such that $a_{m'} \notin K$, $\forall\, m'$, then along this subsequence relations (3.10) and (3.14) give

$$\begin{aligned}\inf_{a\in A} f(a) - \varepsilon &\geq \liminf_{m'\to\infty} f_{n_{m'}}(a_{m'},\omega) = \liminf_{m'\to\infty} \int_T L(a_{m'},\theta)w(\theta)P_{n_{m'}}(d\theta,\omega) \\ &\geq \inf f \cdot \liminf_{m'\infty} \int_T w(\theta)P_{n_{m'}}(d\theta,\omega) \geq \inf f \cdot \Big(\int_T w(\theta)P(d\theta) - \frac{\delta}{2}\Big) \\ &\geq \inf f \cdot (1-\delta) > \inf f - \frac{\varepsilon}{2}.\end{aligned}$$

This is a contradiction and completes the proof.

It is relevant to observe that condition (3.12) is satisfied by a large class of loss functions of the Bayesian statistical decision theory. When $A = \Theta = \mathbb{R}^q$, as in the point estimation, the typical loss function is a non decreasing function of the distance between a and θ, here denoted $||a-\theta||$. For it we have:

3.15 Theorem: *Let* $L(a,\theta) = \Phi(||a-\theta||)$ *with* $\Phi\colon \mathbb{R}^+ \to \mathbb{R}^+$ *non decreasing in the argument and*

$$\int \Phi(||\theta||)w(\theta)P(d\theta) < \lim_{||\theta||\to\infty} \Phi(||\theta||). \tag{3.16}$$

Then

$$\inf_{a\in A} f_n(a,\cdot) \to \inf_{a\in A} f(a) \text{ a.s. .}$$

Proof: From (3.16), for $\varepsilon < \lim_{||\theta||\to\infty} \Phi(||\theta||) - \int \Phi(||\theta||)\, w(\theta)P(d\theta)$, recalling that $\int \Phi(||\theta||)$ $w(\theta)P(d\theta) = \int \Phi(||\theta||)\Pi(d\theta)$, there exists k_ε such that for all $k > k_\varepsilon$ we have

$$f(0) = \int \Phi(||\theta||)w(\theta)P(d\theta) < \Phi(\frac{1}{2}k)\Pi(||\theta|| \leq \frac{1}{2}k) - \varepsilon. \tag{3.17}$$

Let T be a compact subset of Θ and let k be such that $k > k_\varepsilon$ and $T \subset B(0, \frac{1}{2}k)$, the ball with center at the origin and radius $\frac{1}{2}k$; let $K = B(0,k)$. Then for $a \notin B(0,k)$ and $\theta \in B(0,\frac{1}{2}k)$ we have $||a-\theta|| > \frac{1}{2}k$, and by monotonicity of Φ, $\Phi(||a-\theta||) \geq \Phi(\frac{1}{2}k)$. This last relation, together with (3.17) states that

$$L(a,\theta) = \Phi(||a-\theta||) \geq \Phi(\frac{1}{2}k) \geq \Phi(\frac{1}{2}k)\Pi(||\theta|| \leq \frac{1}{2}k) > f(0) + \varepsilon > \inf f \qquad \forall\, \theta \in T,$$

i.e. condition (3.12) is satisfied. The result follows then from theorem 3.1.

References

1. ATTOUCH, H., *Variational convergence for Functions and Operators*, Pitman, London, 1984.

2. ATTOUCH, H., and WETS, R. J-B, Epigraphical processes: laws of large numbers for random lsc functions, *Séminaires d'Analyse Convexe, Montpellier* 1990, 13.1–13.29.

3. BERGER, J. O., *Statistical Decision Theory and Bayesian Analysis*, Springer Verlag, New York, 1985.

4. LUCCHETTI, R. and WETS, R. J-B, Convergence of Minima of Integral Functionals, with Applications to Optimal Control and Stochastic Optimization, *IIASA WP-90-81*, 1990.

5. SALINETTI, G., Stability of Bayesian Decision, *J. of Statistical Planning and Inference*, **40** (1994), 313–326.

6. WETS, R. J-B, Laws of large numbers for random lsc functions, *Applied Stochastic Analysis & Stochastich Monographs*, **5** (1991), 101–120.

7. SHAO, J., Approximations in Bayesian Decision Theory, Tech. Rep. #88–39, Purdue University, 1988.