The Annals of Mathematical Statistics
1970, Vol. 41, No. 3, 1083–1085

CONTINUITY OF THE BAYES RISK[1]

BY ROBERT A. WIJSMAN

University of Illinois

In the theory of statistical decision functions it is sometimes desired to prove that the Bayes risk as a function of the prior distribution is concave and continuous on its domain. The property of concavity is immediate and this implies continuity on the *interior* of the domain if the parameter space is finite. This has been used e.g. by Lehmann [3], Lemma 3.12.5 and by Ferguson [2], Lemma 7.6.1. Continuity on the whole domain does not follow immediately. A proof applicable to a rather special sequential problem (finite parameter space, i.i.d. observations and constant cost per observation) has been given by Blackwell and Girshick [1], Theorem 9.4.2. Another continuity theorem, valid under certain restrictive conditions, can be found in Wald [4], Theorem 4.6. It is desirable to find conditions implying continuity that are both simpler and more widely applicable. The hope is to make strong use of the concavity of the Bayes risk. It should be realized, however, that even on a convex Euclidean domain a concave function is not necessarily continuous on the boundary. On the other hand, it will be shown in this note that continuity is implied by the combination of concavity and a very simple property of the geometry of the domain. This property is satisfied, for instance, by a polyhedron in finite dimensional space. Thus, it can be concluded that in any statistical problem with finite parameter space (plus a mild assumption on the risk functions) the Bayes risk is continuous.

In any given statistical problem let Θ denote the parameter space, whose points θ index the distributions on the sample space; and let D be any class of decision functions δ. We assume that to each $\delta \in D$ there corresponds a *risk function* R_δ on Θ satisfying $0 \leqq R_\delta(\theta) < \infty$ for all $\theta \in \Theta$, $\delta \in D$ (the uniform lower bound 0 could be replaced by any other). We further assume that there is given a sigma-field B on Θ such that every R_δ is B-measurable. Let Λ be a class of probability distributions λ on (Θ, B). The only requirement on Λ is that it be convex (e.g. Λ could be all probability distributions). Let $r_\delta(\lambda) = \int R_\delta(\theta)\, \lambda(d\theta)$ be the average risk of δ when the prior λ is used. Clearly, r_δ is linear on Λ. Define the *Bayes risk* $\rho(\lambda) = \inf_{\delta \in D} r_\delta(\lambda)$. ρ is concave on Λ since it is the infimum of a family of concave functions.

Suppose Λ carries a topology in which all r_δ are upper semicontinuous (often it will be possible to assert that the r_δ are continuous by virtue of their linearity). Then ρ, being the infimum of the r_δ, is also upper semicontinuous on Λ. We would also like to be able to conclude lower semicontinuity of ρ, but this seems impossible without further assumptions. Indeed, it is easy to give an example of a concave function f on a convex set in the plane that is not lower semicontinuous: let f equal 1 on the interior of a disk and in one point x of the boundary, and 0 elsewhere on

Received September 19, 1969.

[1] Research partially supported by National Science Foundation Grant GP-14768.

the boundary. Then $\liminf_{y\to x} f(y) = 0 < f(x)$. However, such a counter-example fails if instead of a disk we take a polygon. This points to a certain assumption on the domain of the function, and we shall make the following

DEFINITION. A subset S of a linear metric space is said to be *polyhedral* if for every $x \in S$ there is $\varepsilon(x) > 0$ such that for every ray r emanating from x the set $\{y: y \in r, 0 < \|y-x\| \leqq \varepsilon(x)\}$ lies entirely either in S or outside S.

In other words, S is polyhedral if a ray from x lies entirely in S for a distance at least $\varepsilon(x)$ or leaves S for at least a distance $\varepsilon(x)$. A polyhedron in n-space furnishes an example. We shall be concerned only with convex S. Clearly, a halfspace in n-space is polyhedral. If $S = \bigcap_{i=1}^m S_i$, and each S_i is polyhedral, then S is polyhedral with $\varepsilon(x) = \min_{1 \leqq i \leqq m} \varepsilon_i(x)$. Thus, the intersection of a finite number of halfspaces in n-space, i.e., a convex polyhedron, is polyhedral.

THEOREM 1. *Let an extended real-valued function f on a subset S of a linear metric space be concave and bounded below. If S is polyhedral then f is lower semicontinuous.*

PROOF. Let $x \in S$. In taking $\liminf_{y\to x} f(y)$ we may restrict y to $\|y-x\| < \varepsilon(x)$, where $\varepsilon(x)$ is given by the definition. Let $y \in S$, $0 < \|y-x\| < \varepsilon(x)$, and consider the ray from x through y. Since S is polyhedral, there exists z on this ray and in S such that $\|z-x\| = \varepsilon(x)$. Putting $\|y-x\|/\varepsilon(x) = \alpha$, we may write $y = \alpha z + (1-\alpha)x$. Let f be bounded below by b, say. Since f is concave, $f(y) \geqq \alpha f(z) + (1-\alpha) f(x) \geqq \alpha b + (1-\alpha) f(x)$. As $y \to x$, $\alpha \to 0$ so that $\liminf_{y\to x} f(y) \geqq f(x)$ and hence f is lower semicontinuous.

An immediate application to statistics is

THEOREM 2. *In a statistical decision problem with $0 \leqq R_\delta(\theta) < \infty$ for every $\theta \in \Theta$, $\delta \in D$, suppose Θ is finite and Λ consists of all probability distributions on Θ (with the obvious topology). Then the Bayes risk ρ is a concave and continuous function on Λ.*

PROOF. Only continuity remains to be shown. Let $\Theta = \{\theta_1, \cdots, \theta_k\}$. A prior $\lambda = (\lambda_1, \cdots, \lambda_k)$ puts probability λ_i on θ_i. Thus, Λ is the simplex $\{\lambda: \lambda_i \geqq 0, i = 1, \cdots, k, \sum \lambda_i = 1\}$, which is a $(k-1)$ dimensional convex polyhedron. Every r_δ is bounded on Λ, therefore continuous, so that ρ is upper semicontinuous. An application of Theorem 1, with $\Lambda = S$, shows that ρ is also lower semicontinuous.

REMARK. In any problem where the r_δ are upper semicontinuous (e.g. when the risk functions also have a uniform upper bound) the conclusion of Theorem 2 applies to any Λ that is polyhedral, using Theorem 1. Here is an example of a Λ that is polyhedral but not a polyhedron in finite dimensional space (unfortunately, the example does not seem to have much statistical interest). Let μ be a sigma-finite measure on (Θ, B). Consider the normed linear space L^∞ of all essentially bounded, B-measurable functions on Θ, with identification of functions equal a.e. μ, and norm equal to the essential supremum. Let $\Lambda \subset L^\infty$ consist of all those functions p that are probability densities with respect to μ, i.e., $\int p \, d\mu = 1$, and have the further property that there exists ε (depending on p) such that $\mu[0 < p < \varepsilon] = 0$. In other words, in almost all points θ, either $p(\theta) = 0$ or $p(\theta) \geqq \varepsilon$.

REFERENCES

[1] Blackwell, D. and Girshick, M. A. (1954). *Theory of Games and Statistical Decisions.* Wiley, New York.

[2] Ferguson, T. S. (1967). *Mathematical Statistics.* Academic Press, New York.

[3] Lehmann, E. L. (1959). *Testing Statistical Hypotheses.* Wiley, New York.

[4] Wald, A., (1950). *Statistical Decision Functions.* Wiley, New York.