NOTES

A NOTE ON BAYES ESTIMATES

By Peter J. Bickel[1] and David Blackwell[2]

University of California, Berkeley

0. Introduction and summary. Throughout this paper we are concerned with the problem of estimating a real parameter when the loss function is such that the Bayes estimate exists, is unique, and satisfies a simple Equation, (1.5). If the estimate is unbiased (in the general sense of Lehmann [3]) we show under weak conditions that it must satisfy another Equation, (1.14). The main result of Section 1, Theorem 1.3, shows that, in general, these two equations are incompatible unless the Bayes risk is 0. This extends Theorem 11.2.4 of [1] which states that in estimation with quadratic loss, unbiased Bayes estimates have Bayes risk 0. Some counter-examples at the end of the section indicate the limits of this incompatibility result.

1. Unbiased Bayes estimates. We consider the problem of estimating a real parameter θ on the basis of an observation X distributed according to one of a family of distributions $\{P_\theta\}$, $\theta \varepsilon \Theta$ on some measurable space $(\mathfrak{X}, \alpha)$. We assume P_θ has density p_θ with respect to some σ finite measure μ on $(\mathfrak{X}, \alpha)$. Our decision space $D = R$ and the given loss function $l(\theta, d)$ is such that,

$$l(\theta, d) = \omega(\theta)\tilde{l}(|\theta - d|) \tag{1.1}$$

where $\omega(\theta) > 0$, $\tilde{l}(0) = 0$. We shall assume throughout that $\tilde{l}(t)$ has a continuous derivative $\tilde{l}'(t)$ which is positive for $t > 0$ and 0 for $t = 0$. Since $\tilde{l}$ is only defined for $t \geqq 0$ the derivative at 0 is understood to be one sided. Take Θ to be an open set, endow it with the usual σ field, and suppose that $p_\theta(x)$ is bimeasurable in θ and x. Let π be any Bayes (prior) probability on Θ. We restrict our attention to non-randomized estimates, that is, measurable functions from $\mathfrak{X}$ to R. We assume that a Bayes estimate $\delta_\pi(x)$ exists and minimizes the posterior risk for almost all x. Formally, if we define,

$$h(x, d) = \int l(\theta, d)p_\theta(x)\pi(d\theta), \tag{1.2}$$

then we suppose that we can choose δ_π to be measurable and to satisfy,

$$h(x, \delta_\pi(x)) = \min\{h(x, d), d \varepsilon R\} \tag{1.3}$$

a.s. $[P]$. The probability measure P here is the marginal distribution of X given by,

$$P = \int_\Theta P_\theta\pi(d\theta). \tag{1.4}$$

Received 30, 1966; revised 3 April 1967.

[1] Research supported in part by United States Public Health Service, Grant GM-10525.

[2] Research supported in part by the National Science Foundation, Grant GP-5052.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.* www.jstor.org

The very weak Assumption (1.3) is satisfied if, for instance, $\tilde{l}$ is convex. We shall need the more restrictive assumption that $\delta_\pi(x)$ satisfies a.s. $[P]$ the equation,

$$(1.5)\quad \int_{[\theta>\delta_\pi(x)]} \omega(\theta)\tilde{l}'(|\theta - \delta_\pi(x)|)p_\theta(x)\pi(d\theta) = \int_{[\theta<\delta_\pi(x)]} \omega(\theta)\tilde{l}'(|\theta - \delta_\pi(x)|)p_\theta(x)\pi(d\theta).$$

This equation holds if (1.3) is satisfied and furthermore $h(x, d)$ is finitely differentiable in d and satisfies,

$$(1.6)\quad \partial h(x, d)/\partial d = \int_{[\theta>d]} \omega(\theta)\tilde{l}'(|\theta - d|)p_\theta(x)\pi(d\theta) - \int_{[\theta<d]} \omega(\theta)\tilde{l}'(|\theta - d|)p_\theta(x)\pi(d\theta),$$

for all d, a.s. $[P]$.

For the loss function corresponding to $\tilde{l}(t) = t^2$, Blackwell and Girshick [1] (Theorem 11.2.3) have show that if the Bayes risk is finite and if,

$$(1.7)\quad \int \omega(\theta)\pi(d\theta) < \infty$$

then δ_π satisfies (1.5) or the equivalent equation

$$(1.8)\quad \delta_\pi(x) = \int \theta\omega(\theta)p_\theta(x)\pi(d\theta)/\int p_\theta(x)\omega(\theta)\pi(d\theta).$$

More generally, we can state

THEOREM 1.1. *Suppose that* (1.7) *holds. Let l be given by* (1.1) *and assume that there exists an $r < \infty$ such that, $\tilde{l}'(t) = 0(t^{r-1})$ as $t \to \infty$, and that,*

$$(1.9)\quad \int\int |\theta|^r\omega(\theta)p_\theta(x)\pi(d\theta) < \infty.$$

Then (1.5) *holds, for all d,* a.s. $[P]$.

PROOF. By (1.7) and (1.9) $h(x, d)$ is finite for all d. Of course,

$$(1.10)\quad \lim_{h\to 0} \tilde{l}(\theta, d + h) - \tilde{l}(\theta, d)/h = \tilde{l}'(|\theta - d|) \operatorname{sgn} (\theta - d)$$

where $\operatorname{sgn} t = 1$ if $t \geqq 0$, -1 if $t < 0$. By assumption we have,

$$(1.11)\quad \tilde{l}'(t) \leqq C + Mt^{r-1}$$

for some C, M, finite. Now, (1.7), (1.9) and (1.11) yield,

$$(1.12)\quad \int \sup \{\tilde{l}'(|\theta + d|): |d| \leqq A\}\omega(\theta)p_\theta(x)\pi(d\theta) < \infty$$

for all $A < \infty$. The result follows from (1.10) (1.12), and (1.6) by a standard argument.

For the loss function l corresponding to $\tilde{l}(t) = t^2$, Blackwell and Girshick have proved the following result.

THEOREM 11.2.4 [1]. *If* (1.7) *holds and the Bayes risk is finite, then the unbiasedness of δ_π implies that the Bayes risk is* 0.

An equivalent statement of this theorem is: "If δ_π is unbiased, $\int_{[\delta_\pi=\theta]} p_\theta(x)\mu\cdot(dx)\pi(d\theta) = 1$". In other words from a Bayesian point of view, given the prior knowledge π, X provides complete information about θ.

In [3] Lehmann introduced a generalization of the concept of unbiasedness which is given by the following definition.

DEFINITION 1.1. δ is unbiased if

$$(1.13) \qquad E_\theta(l(\theta, \delta(x))) = \min \{E_\theta(l(\theta', \delta(x)): \theta' \varepsilon \Theta\}.$$

The most important special case of the above definition is, of course, $l(\theta, d) = (\theta - d)^2$ which leads to the usual definition of unbiasedness. Another example of interest is $l(\theta, d) = |\theta - d|$ which leads to so-called median unbiasedness.

We will be interested in those cases in which δ_π is unbiased for the loss function $l^*(\theta, d) = \tilde{l}(|\theta - d|)$. In fact we will require that δ_π satisfies for almost all $\theta[\pi]$ the equation,

$$(1.14) \quad \int_{[\theta>\delta_\pi(x)]} \omega(\theta)\tilde{l}'(|\theta - \delta_\pi(x)|)p_\theta(x)\mu(dx) = \int_{[\theta<\delta_\pi(x)]} \omega(\theta)\tilde{l}'(|\theta - \delta_\pi(x)|)p_\theta(x)\mu(dx).$$

Let us define,

$$(1.15) \qquad g(\theta, \theta') = \int \tilde{l}(|\theta' - \delta_\pi(x)|)p_\theta(x)\mu(dx).$$

If $g(\theta, \theta')$ is differentiable in θ' for almost all $\theta[\pi]$ and,

$$(1.16) \quad \partial g(\theta, \theta')/\partial\theta' = \int_{[\theta'>\delta_\pi]} \tilde{l}'(|\theta' - \delta_\pi(x)|)p_\theta(x)\mu(dx) - \int_{[\theta'<\delta_\pi]} \tilde{l}'(|\theta' - \delta_\pi(x)|)p_\theta(x)\mu(dx),$$

then δ_π satisfies (1.14) for almost all $\theta[\pi]$. If $\tilde{l}(t) = t^2$ and $E_\theta(\delta_\pi^2) < \infty$ (1.14) is seen to be satisfied and is equivalent to,

$$(1.17) \qquad E_\theta(\delta_\pi(x)) = \theta.$$

More generally, we can state,

THEOREM 1.2. *Suppose* $\tilde{l}'(t) = O(t^{r-1})$ *and,*

$$(1.18) \qquad \int\int |\delta_\pi(x)|^r p_\theta(x)\pi(d\theta)\mu(dx) < \infty$$

Then, if δ_π *is unbiased in Lehmann's sense for* l^*, (1.14) *holds.*

PROOF. Same as that of Theorem 1.1. Of course, the Bayes nature of δ_π is immaterial.

We now prove the main theorem of this section, a generalization of Theorem 11.2.4. of [1]. Our method is essentially a generalization of an argument ascribed to Kinney and Snell in Doob [2] p. 314. We define the Bayes risk R_π of π in the usual fashion by,

$$(1.19) \qquad R_\pi = \int\int l(\theta, \delta_\pi(x))p_\theta(x)\mu(dx)\pi(d\theta).$$

THEOREM 1.3. *Suppose*

$$(1.20) \qquad \int\int \tilde{l}'(|\theta - \delta_\pi(x)|)\omega(\theta)p_\theta(x)\pi(d\theta)\mu(dx) < \infty$$

and δ_π *satisfies* (1.6) *for almost all* $x[P]$ *and* (1.14) *for almost all* $\theta[\pi]$. *Then the Bayes risk is* 0.

The proof hinges on the following lemma.

LEMMA 1.4. *Let* Q *be a finite measure on the plane and* $Q(\cdot \mid y)$, $Q(\cdot \mid z)$ *be the regular conditional probabilities given the first and second co-ordinates respectively. If*

$$Q[y > z \mid z] = Q[y < z \mid z] \tag{1.21}$$

and

$$Q[y > z \mid y] = Q[y < z \mid y]$$

for almost all y, $z[Q]$, *then*

$$Q[y \neq z] = 0 \tag{1.22}$$

PROOF. From (1.21) we can easily see that for any given real A

$$Q[y > z, y \geqq A] = Q[y < z, y \geqq A] \tag{1.23}$$

and

$$Q[y > z, z \geqq A] = Q[y < z, z \geqq A]. \tag{1.24}$$

Equivalently, we have,

$$Q[y > z, y \geqq A, z < A] = Q[y < z, y \geqq A] - Q[y > z, z \geqq A] \tag{1.25}$$

and

$$Q[y < z, z \geqq A, y < A] = Q[y > z, z \geqq A] - Q[y < z, y \geqq A]. \tag{1.26}$$

We conclude that

$$Q[y < z, y < A, z \geqq A] = Q[y > z, y \geqq A, z < A] = 0 \tag{1.27}$$

for all A. The lemma follows.

To prove the theorem consider the probability measure Q^* induced on the plane by the joint distribution of $(\delta_\pi(x), \theta)$. If we use the notation of Lemma 1.4 (1.5) and (1.14) become,

$$\int_{[y>z]} \omega(z)\bar{l}'(|y - z|)Q^*(dy \mid z) = \int_{[y<z]} \bar{l}'(|y - z|)\omega(z)Q^*(dy \mid z) \tag{1.28}$$

and

$$\int_{[y>z]} \omega(z)\bar{l}'(|y - z|)Q^*(dz \mid y) = \int_{[y<z]} \omega(z)\bar{l}'(|y - z|)Q^*(dz \mid y) \tag{1.29}$$

Define Q by

$$Q(A) = \int_A \omega(z)\bar{l}'(|y - z|)Q^*(dy, dz) \tag{1.30}$$

By (1.20) Q is finite and it is easy to see that

$$Q(A \mid y) = \int_A \omega(z)\bar{l}'(|y - z|)Q^*(dz \mid y)/\int \omega(z)\bar{l}'(|y - z|)Q^*(dz \mid y) \tag{1.31}$$

and that a similar relation holds for $Q(\cdot \mid z)$. It readily follows that (1.28) and (1.29) are equivalent to (1.21) holding for the given Q. Since $\bar{l}'(t) > 0$ for $t \neq 0$

and $\omega(\theta) > 0$, $Q(y \neq z) = 0$ implies $Q^*(y \neq z) = 0$ and the theorem follows.

Corollary 1.5. *Suppose* $\tilde{l}'(t) = O(t^{r-1})$, ω *is bounded, and* (1.9) *and* (1.18) *hold. Then if* δ_π *is unbiased for* $l^*(\theta, d) = \tilde{l}(|\theta - d|)$, *the Bayes risk is* 0.

Proof. The conditions of Theorems 1.1–1.3 are easily seen to be satisfied.

Remarks. If $\tilde{l}(t) = t^r$, $r > 1$, (1.9) and (1.18) roughly correspond to requiring finiteness of the Bayes risk. In this instance it is clear that Theorem 1.3 is strictly stronger than its corollary since the Equations (1.5) and (1.14) involve only the $(r - 1)$st moments of δ_π and θ. Thus if $r = 2$ the corollary corresponds to Theorem 11.2.4 while the theorem is equivalent to the statement made in [1] that 11.2.4 holds without requiring second moments of θ and δ_π.

If $\tilde{l}'(0) \neq 0$ Equation (1.5) does not, in general, hold. Theorem 1.3 may fail and, in fact, is false for median unbiased estimates as can be seen from the following counter example.

Let θ be uniformly distributed on $[-1, 1]$, $\mathfrak{X} = R$, P_θ assign mass $\frac{1}{3}$ to 0, and $\frac{2}{3}$ to θ. Then $\delta_\pi(x) = \text{median } (\theta \mid x) = x$ and θ is the unique median of P_θ.

Even if $\tilde{l}(t) = t^2$, δ_π must satisfy Equation (1.5) for the theorem to hold. For instance if (c) does not hold Blackwell and Girshick give an example (binomial estimation with $\omega(\theta) = [\theta(1 - \theta)]^{-1}$) in which the conclusion of Theorem 1.3 fails. But, in this instance (1.5) is not satisfied for $x = 0$ and $x = 1$.

We conclude by remarking that Lemma 1.3 is false if Q is not a finite measure. A classical counterexample is given by,

$$Q(A) = \int \int_A \psi(t - \theta)\, dt\, d\theta$$

where ψ is the standard normal density.

REFERENCES

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

[2] Doob, J. L. (1953). *Stochastic Process.* Wiley, New York.

[3] Lehmann, E. L. (1951). A generalized concept of unbiasedness. *Ann. Math. Statist.* **22** 587–597.