CONDITIONAL EXPECTATION AND CONVEX FUNCTIONS

E. W. BARANKIN
UNIVERSITY OF CALIFORNIA

1. Introduction

The *integral* inequality (6) of [1], which extends Blackwell's theorem 2, [2], has in turn been generalized by Hodges and Lehmann, [3, lemma 3.1], who consider a k-dimensional, vector valued function f and replace the absolute s-th power with a real valued, convex function ψ on $\mathcal{E}^k$ (Euclidean k-space). Now, both Blackwell and the present author showed in [1, theorem, p. 281] and in the paragraph following lemma [2, p. 107] that their integral inequalities are consequences of the fact that the *integrands satisfy the same inequalities almost everywhere.* On the other hand, Hodges and Lehmann prove their integral inequality directly (except when conditional probabilities exist almost everywhere as measures), and so there is raised the question of whether or not that inequality is likewise a consequence of an almost everywhere inequality between the integrands. It is the purpose of this note to prove, in the theorem below, that this is indeed the case. Thus, in particular, it is shown that the assertions in [3, section 3] concerning risk functions follow because the same assertions can be made almost everywhere concerning "convex" loss functions.

2. Preliminaries

Let Ω be a space of points x; $\mathcal{Q}$ a σ-field of subsets of Ω; and μ, a probability measure on $\mathcal{Q}$. Let t be a function on Ω to a space Γ of points τ; $\mathcal{T}$, a σ-field of subsets of Γ; and $\mathcal{T}^{-1}$—a sub-σ-field of $\mathcal{Q}$—the inverse of $\mathcal{T}$ under t. Let ν denote the measure on $\mathcal{T}$ defined by $\nu(A) = \mu[t^{-1}(A)]$.

Let f be an $\mathcal{Q}$-measurable, μ-integrable function on Ω to Euclidean k-space, $\mathcal{E}^k$. Finally, let ψ be a (finite) real valued, convex function on $\mathcal{E}^k$. It follows that ψ is continuous [4, p. 19]; we assume that $E[\psi(f)]$ exists.

Conditional expectation with respect to the function t will be indicated in the usual way; for example, $E(f|\cdot)$. Also, let us agree to denote by $[h]_A$ the range of a function h over a subset A of its domain.

The heart of the theorem below is the following

LEMMA. *Let ψ be as above, and let h be a $\mathcal{T}$-measurable function on Γ to $\mathcal{E}^k$. Let $A \in \mathcal{T}$ be a set of positive ν-measure such that for $\tau \in A$ we have $\psi[h(\tau)] > a$.*

Then there exists a $\mathcal{T}$-subset $B \subsetneq A$, with $\nu(B) > 0$, such that for all y in the convex hull of $[h]_B$ we have $\psi(y) > a$.

This note was prepared under contract with the Office of Naval Research.

The convex hull of a set in $\mathcal{E}^k$ is the smallest convex set containing the given set.

To prove the lemma, consider the set

$$R = \{y \in \mathcal{E}^k \mid \psi(y) > a\}.$$

Since ψ is continuous, R is open and so is representable as the union of denumerably many, nonoverlapping cells:

$$R = \bigcup_{i=1}^{\infty} J_i.$$

The sets $h^{-1}(J_i)$ are all $\mathcal{T}$-sets and nonoverlapping, and since A is a subset of their union, we have

$$A = \bigcup_{i=1}^{\infty} [A \cap h^{-1}(J_i)];$$

consequently,

$$\nu(A) = \sum_{i=1}^{\infty} \nu[A \cap h^{-1}(J_i)].$$

Since $\nu(A) > 0$, at least one of the terms in the sum on the right must be positive, say the first. Thus, the subset $B = A \cap h^{-1}(J_1)$ of A is a $\mathcal{T}$-set and has positive ν-measure. Moreover, $[h]_B \subseteq J_1$, and since J_1 is convex, the convex hull of $[h]_B$ is likewise a subset of J_1. But J_1 lies entirely in R, so that for all $y \in J_1$ we have $\psi(y) > a$. *A fortiori* the same is true for the convex hull of $[h]_B$, and the proof is complete.

Notice that only the continuity of the function ψ has been used for the lemma. Observe also that since ψ is continuous we can assert that for all y in the closure of the convex hull of $[h]_B$ we have $\psi(y) \geqq a$. This remark will be used below.

3. The result

Let us begin by noting that for a set $B \in \mathcal{T}$, with $\nu(B) > 0$, we have

$$\int_{t^{-1}(B)} f \frac{d\mu}{\mu[t^{-1}(B)]} = \int_B E(f \mid \cdot) \frac{d\nu}{\nu(B)}$$

and

$$\int_{t^{-1}(B)} \psi(f) \frac{d\mu}{\mu[t^{-1}(B)]} = \int_B E[\psi(f) \mid \cdot] \frac{d\nu}{\nu(B)};$$

and therefore, applying Jensen's inequality [5, p. 186, equation (5′)] on the left hand members, we obtain

$$\psi\left(\int_B E(f \mid \cdot) \frac{d\nu}{\nu(B)}\right) \leqq \int_B E[\psi(f) \mid \cdot] \frac{d\nu}{\nu(B)}. \tag{1}$$

We now go to the proof of

THEOREM. *For almost all* (ν) $\tau \in \Gamma$ *we have*

$$\psi[E(f \mid \tau)] \leqq E[\psi(f) \mid \tau]. \tag{2}$$

Suppose the assertion of the theorem were false. Then we should have a $\mathcal{T}$-set A, with $\nu(A) > 0$, and numbers a, b with $a > b$, such that

$$\begin{cases} \psi\,[E(f \mid \tau)] > a\,, \\ E\,[\psi(f) \mid \tau] < b\,, \qquad \tau \in A\,. \end{cases}$$

We apply the lemma and the remark following the lemma, and so obtain a $\mathcal{T}$-set $B \subsetneq A$ with $\nu(B) > 0$ such that

(3) $\qquad \psi(y) \geqq a\,, \qquad y \in$ closure of convex hull of $[E\,(f \mid \cdot)]_B$,

and

(4) $$E\,[\psi\,(f) \mid \tau] < b\,, \qquad \tau \in B\,.$$

If we notice now that the point

$$\int_B E\,(f \mid \cdot)\,\frac{d\nu}{\nu\,(B)}$$

of $\mathcal{E}^k$ belongs to the closure of the convex hull of $[E(f \mid \cdot)]_B$,[1] then using (1), (3) and (4) we get the following string of inequalities:

$$a \leqq \psi\left(\int_B E\,(f \mid \cdot)\,\frac{d\nu}{\nu\,(B)}\right) \leqq \int_B E\,[\psi\,(f) \mid \cdot\,]\,\frac{d\nu}{\nu\,(B)} < b\,.$$

Thus, the fact that $a > b$ is contradicted, and the theorem is established.

If one takes expectations of the two members of (2), one obtains the result of Hodges and Lehmann.

REFERENCES

[1] E. W. BARANKIN, "Extension of a theorem of Blackwell," *Annals of Math. Stat.*, Vol. 21 (1950), pp. 280–284.

[2] D. BLACKWELL, "Conditional expectation and unbiased sequential estimation," *Annals of Math. Stat.*, Vol. 18 (1947), pp. 105–110.

[3] J. L. HODGES, JR. and E. L. LEHMANN, "Some problems in minimax point estimation," *Annals of Math. Stat.*, Vol. 21 (1950), pp. 182–197.

[4] T. BONESSEN and W. FENCHEL, *Theorie der Konvexen Körper, Ergebnisse der Mathematik und ihrer Grenzgebiete*, Vol. 3, No. 1, Springer, Berlin, 1934.

[5] J. L. W. V. JENSEN, "Sur les fonctions convexes et les inegalités entre les valeurs moyennes," *Acta Math.*, Vol. 30 (1906) pp. 175–193.

[1] To assume this not true is to be able to exhibit a convex function which takes on a larger value at the given point than at any point of the closure of the convex hull of $[E(f \mid \cdot)]_B$—thus contradicting Jensen's inequality. Such a convex function is, for example, distance from the closed convex hull of $[E(f \mid \cdot)]_B$.