The Annals of Mathematical Statistics
1970, Vol. 41, No. 2, 698–701

# A CHARACTERIZATION OF A CONDITIONAL EXPECTATION WITH RESPECT TO A $\sigma$-LATTICE

By Richard L. Dykstra

*University of Missouri*

**0. Summary.** Several authors [1], [2], $\cdots$, [6], have derived characterizations of a conditional expectation operator. That is, if $T$ is a transformation which maps a particular set of functions into the same set, then necessary and sufficient conditions are specified so that $T$ is a conditional expectation operator. It is shown in the present paper that a similar sort of characterization can be found in the more general case when $T$ is a conditional expectation with respect to a $\sigma$-lattice operator even though $T$ need not be linear.

**1. Introduction.** Let $(X, \mathscr{A}, \mu)$ denote a measure space. $L_2$ denotes the set of equivalence classes of square-integrable functions defined by $f \sim g$ if $f = g$ a.s. $\mathscr{L}$ is said to be a sub $\sigma$-lattice of $\mathscr{A}$ if $\mathscr{L}$ is a subset of $\mathscr{A}$ containing $\phi$ and $X$, which is closed under countable unions and intersections. A real-valued function defined on $X$ is $\mathscr{L}$-measurable if $[f \geqq a] \in \mathscr{L}$ for all real $a$. The set of $\mathscr{L}$-measurable functions in $L_2$ will be denoted by $L_2(\mathscr{L})$. For an arbitrary function $f$ belonging to $L_2$, the conditional expectation of $f$ with respect to a $\sigma$-lattice $\mathscr{L}$, i.e. $E(f \mid \mathscr{L})$, is defined to be the function in $L_2(\mathscr{L})$ which minimizes $\int (f-g)^2 \, d\mu \; \forall \, g \in L_2(\mathscr{L})$. Brunk [7] shows that $E(f \mid \mathscr{L})$ exists uniquely, and that $E(f \mid \mathscr{L})$ is the usual conditional expectation when $\mathscr{L}$ is also a sub $\sigma$-field of $\mathscr{A}$. Brunk also shows that $E(f \mid \mathscr{L})$ is characterized by (i) $\int [f - E(f \mid \mathscr{L})] E(f \mid \mathscr{L}) \, d\mu = 0$ and (ii) $\int [f - E(f \mid \mathscr{L})] h \, d\mu \leqq 0$ $\forall \, h \in L_2(\mathscr{L})$.

**2. First characterization.** Let us first consider the case where $\mu$ is strictly finite, and $T: L_2 \to L_2$. The following terminology will be used.

*Idempotency.* $T$ is idempotent if $T(Tf) = Tf \; \forall \, f \in L_2$.

*Scale invariance.* $T$ is scale invariant if $T(af) = aTf \; \forall \, f \in L_2$ and $\forall \, a \geqq 0$.

*Monotonicity.* $T$ is monotone if $f \geqq g$ implies $Tf \geqq Tg$.

*Expectation invariance.* $T$ is expectation invariant if

$$\int f \, d\mu = \int Tf \, d\mu \qquad \forall f \in L_2.$$

*Distance Reducing.* $T$ is distance reducing if

$$\int (f-g)^2 \, d\mu \geqq \int (Tf - Tg)^2 \, d\mu \qquad \forall f, g \in L_2.$$

Let $F = \{f : Tf = f\}$ denote the fixed points in $L_2$. Observe that if $T$ is scale invariant, then $f \in F$ implies that $af \in F \; \forall \, a \geqq 0$. The following remark is a well-known fact.

Remark 2.1. If $T$ is distance reducing then $F$ is closed in the $L_2$ norm.

Received December 8, 1968.

LEMMA 2.1. *Suppose that $T$ is distance reducing and monotone. Then if $f$, $g \in F$, $f \vee g$ and $f \wedge g \in F$.*

PROOF. If $T$ is monotone, $T(f \vee g) \geqq Tf = f$, and $T(f \vee g) \geqq Tg = g$, so that $T(f \vee g) \geqq f \vee g \geqq f$. Hence $\int (f - f \vee g)^2 \, d\mu \leqq \int [Tf - T(f \vee g)]^2 \, d\mu$. But if $T$ is distance reducing then $\int (f - f \vee g)^2 \geqq \int [Tf - T(f \vee g)]^2 \, d\mu$, so that $T(f \vee g) = f \vee g \in F$. In a similar manner, one can show that $f \wedge g = T(f \wedge g) \in F$.

LEMMA 2.2. *If $T$ is expectation invariant and distance reducing, then $T(f+c) = Tf + c$, $\forall f \in L_2$ and $\forall$ real $c$.*

PROOF. The result is clear if $c = 0$. Thus we may assume that $c \neq 0$. Then

$$\int T(f+c) - Tf \, d\mu = \int f + c \, d\mu - \int f \, d\mu = c\mu(X).$$

From the Schwarz inequality,

$$\begin{aligned} \int c^2 \, d\mu \cdot \int [T(f+c) - Tf]^2 \, d\mu &\geqq [\int c(T(f+c) - Tf) \, d\mu]^2 \\ &= c^2 [\int T(f+c) - Tf \, d\mu]^2 \\ &= c^2 \cdot c^2 \mu(X)^2 = [\int c^2 \, d\mu]^2. \end{aligned}$$

Divided by $\int c^2 \, d\mu_\mu$, we have

$$\int [T(f+c) - Tf]^2 \, d\mu \geqq \int c^2 \, d\mu = \int (f + c - f)^2 \, d\mu.$$

However, since $T$ is distance reducing, we have equality in the Schwarz inequality, so that $T(f+c) - Tf = a \cdot c$. Moreover, $a = 1$ since $T$ is expectation invariant.

COROLLARY 2.1. *If $T$ is expectation invariant and distance reducing and $f \in F$, then $f + c \in F$ $\forall$ real $c$.*

COROLLARY 2.2. *If $T$ is scale invariant, expectation invariant, and distance reducing, then $F$ contains all constant functions.*

LEMMA 2.3. *Suppose $T$ is scale invariant, monotone, expectation invariant, and distance reducing. Then*

(a) *$\mathscr{L} = \{[f \geqq a]; f \in F$, a real$\}$ is a $\sigma$-lattice containing $\phi$ and $X$, and furthermore*
(b) *$F = L_2(\mathscr{L})$.*

PROOF. (a) Let us show first that $A \in \mathscr{L}$ iff $I_A \in F$. The "if" part is trivial, so we will consider the "only if" part. If $A \in \mathscr{L}$, $A = [f \geqq a]$ for some $f \in F$, and some real $a$. Then $f_1 = [(f - a + 1) \vee 0] \wedge 1$ and $f_n = (nf_1 - n + 1) \vee 0$, $n > 1$, belong to $F$ by Lemmas 2.1 and 2.2. However, one can use the Lebesgue dominated convergence theorem to show $f_n \to I_A$ in $L_2$, since $f_n \to I_A$ pointwise. Thus $I_A \in F$ by Remark 2.1. It is easily shown that if $f \in L_2$, $f_n \in F$, $f_n \leqq f$, $n = 1, 2, \cdots$, then $\bigvee_{n=1}^\infty f_n \in F$. Thus, if $\{A_n\}$ is a sequence of sets in $\mathscr{L}$, $\bigvee_{n=1}^\infty I_{A_n} \in F$, so that $[\bigvee_{n=1}^\infty I_{A_n} \geqq 1] = \bigcup_{n=1}^\infty A_n \in \mathscr{L}$. In a similar manner, one can show $\bigcap_{n=1}^\infty A_n \in \mathscr{L}$, so that $\mathscr{L}$ is a $\sigma$-lattice.

(b) Clearly every function in $F$ is in $L_2(\mathscr{L})$, so let us show the converse. Suppose

$f \in L_2(\mathscr{L})$ then $[f \geqq -n+i/2^n] \in \mathscr{L}$ $\forall n$, $i = 1, 2, \cdots, n2^{n+1}$, so that $I_{[f \geqq -n+i/2^n]} \in F$. By Lemmas 2.1 and 2.2 and the properties of scale invariance,

$$f_n = \bigvee_{i=1}^{n2^{n+1}} i/2^n I_{[f \geqq -n+i/2^n]} - n \quad \text{must belong to} \quad F.$$

However, it can be shown that $f_n \to f$ in $L_2$ as $n \to \infty$, so that $f \in F$ by Remark 2.1.

Let us now show that $Tf = E(f \mid \mathscr{L})$ if we add two other restrictions.

THEOREM 2.1. *If $T$ is an expectation invariant, scale invariant, monotone, idempotent, distance reducing transformation such that $\int f^2 \, d\mu \geqq \int (f - Tf)^2 \, d\mu$ for all $f \in L_2$, then $T$ is a conditional expectation* w.r.t. *a $\sigma$-lattice operator.*

PROOF. First we will show that $\int (f - Tf) h \, d\mu \leqq 0$ $\forall f \in L_2$, $\forall h \in F$. We may assume that $0 < \int (f - Tf)^2 \, d\mu$, since otherwise the result is obvious. Suppose $\exists h \in F \ni \int (f - Tf) h \, d\mu > 0$. Then

$$0 < \int (f - Tf)^2 \, d\mu = \int f^2 \, d\mu + \int (Tf)^2 \, d\mu - 2 \int f \, Tf \, d\mu \leqq 2 \int (f - Tf) f \, d\mu$$

since $0 \in F$ and $T$ is distance reducing. Thus

$$h^* = [\int (f - Tf) f \, d\mu][\int (f - Tf) h \, d\mu]^{-1} h \in F.$$

Then $Th^* = h^*$, so that

$$\begin{aligned} \int (Th^* - Tf)^2 \, d\mu &= \int (Th^* - f + f - Tf)^2 \, d\mu \\ &= \int (h^* - f)^2 \, d\mu + \int (f - Tf)^2 \, d\mu + 2 \int (f - Tf)(h^* - f) \, d\mu. \end{aligned}$$

However, the last term is

$$2[\int (f - Tf) f \, d\mu][\int (f - Tf) h \, d\mu]^{-1} \int (f - Tf) h \, d\mu - 2 \int (f - Tf) f \, d\mu = 0.$$

Thus $\int (Th^* - Tf)^2 \, d\mu > \int (h^* - f)^2 \, d\mu$ which contradicts the fact that $T$ is distance reducing.

Observe now that since $\mathscr{L} = \{[f \geqq a], f \in F, a \text{ real}\}$ is a $\sigma$-lattice containing $\phi$ and $X, E(\cdot \mid \mathscr{L})$ is a well defined operator whose range is $F$. Moreover, since $T$ is idempotent, $Tf \in F \, \forall f \in L_2$. Suppose now that $f$ is such that $E(f \mid \mathscr{L}) = 0$. Then since $Tf \in F, 0 \geqq \int [f - E(f \mid \mathscr{L})] Tf \, d\mu = \int f \, Tf \, d\mu$, and because $\int f^2 \, d\mu \geqq \int (f - Tf)^2 d\mu$, we may say $0 \geqq -2 \int f \, Tf \, d\mu + \int (Tf)^2 \, d\mu$, so that $Tf = 0$.

Now let us show that (i) $\int E(f \mid \mathscr{L})^2 \, d\mu \leqq \int (Tf)^2 \, d\mu$, and that (ii) equality holds iff $E(f \mid \mathscr{L}) = Tf$. The result is clear by the previous paragraph if $E(f \mid \mathscr{L}) = 0$. Thus let us assume that that is not the case.

(i) By the Schwarz inequality

$$[\int Tf \, E(f \mid \mathscr{L}) \, d\mu]^2 \leqq \int (Tf)^2 \, d\mu \cdot \int E(f \mid \mathscr{L})^2 \, d\mu. \tag{2.1}$$

However, since $E(f \mid \mathscr{L}) \in F$, $\int (f - Tf) E(f \mid \mathscr{L}) \, d\mu \leqq 0$ by the first part of the theorem, and hence

$$0 < \int E(f \mid \mathscr{L})^2 \, d\mu = \int f E(f \mid \mathscr{L}) \, d\mu \leqq \int Tf \cdot E(f \mid \mathscr{L}) \, d\mu$$

so that

$$[\int E(f \mid \mathscr{L})^2 \, d\mu]^2 \leqq [\int Tf E(f \mid \mathscr{L}) \, d\mu]^2. \tag{2.2}$$

Combining (2.1) and (2.2) and dividing by $\int E(f \mid \mathscr{L})^2 \, d\mu$ gives us (i).

The "if" part of (ii) is clear, so let us show the "only if" part. If equality is actually the case, then we have equality in the Schwarz inequality, so that $E(f \mid \mathscr{L})$ is proportional to $Tf$, i.e. $E(f \mid \mathscr{L}) = aTf$. However, $a = \pm 1$ by assumption. If $a = -1$, then from the statement just prior to 2.2, $0 < \int E(f \mid \mathscr{L})^2 d\mu \leqq -\int E(f \mid \mathscr{L})^2 d\mu$, which is a contradiction.

Since $g = 0$ is such that

$$\int (f - E(f \mid \mathscr{L}) - g) g \, d\mu = 0 \quad \text{and} \quad \int (f - E(f \mid \mathscr{L}) - g) h \, d\mu \leqq 0 \; \forall \, h \in L_2(\mathscr{L})$$

we have that $E[[f - E(f \mid \mathscr{L})] \mid \mathscr{L}] = 0$, and hence that $T(f - E(f \mid \mathscr{L})) = 0$. Moreover, $T$ is distance reducing so that

$$\begin{aligned} \int E(f \mid \mathscr{L})^2 \, d\mu &= \int [f - (f - E(f \mid \mathscr{L}))]^2 \, d\mu \geqq \int [Tf - T(f - E(f \mid \mathscr{L}))]^2 \, d\mu \\ &= \int (Tf)^2 \, d\mu. \end{aligned}$$

Thus $\int (Tf)^2 \, d\mu = \int E(f \mid \mathscr{L})^2 \, d\mu$, so that $Tf = E(f \mid \mathscr{L})$.

One can obtain the same result if the property of expectation invariance is weakened to require only that there exist constant functions $c_1 > 0$ and $c_2 < 0$ such that $\int T(c_i) \, d\mu = \int c_i \, d\mu$, $i = 1, 2$. Examples can be found to show that the conditions are independent.

**3. Acknowledgment.** The author would like to thank Professors Timothy Robertson and Jonathan Cryer for their helpful suggestions.

## REFERENCES

[1] Moy, S. C. (1954). Characterization of conditional expectation as a transformation on function spaces. *Pacific J. Math.* **4** 47–63.

[2] Bahadur, R. R. (1955). Measurable subspaces and subalgebras. *Proc. Amer. Math. Soc.* **6** 565–570.

[3] Sidák, Z. (1957). On relations between strict-sense and wide-sense conditional expectations. *Theory Probability Appl.* **2** 267–271.

[4] Olsen, M. P. (1965). A characterization of conditional probability. *Pacific J. Math.* **15** 971–938.

[5] Douglas, R. G. (1965). Contractive projections on an Lp space. *Pacific J. Math.* **15** 443–462.

[6] Pfanzgl, J. (1967). Characterizations of conditional expectations. *Ann. Math. Statist.* **38** 415–421.

[7] Brunk, H. D. (1965). Conditional expectation given a $\sigma$-lattice and applications. *Ann. Math. Statist.* **36** 1339–1350.