CHARACTERIZATIONS OF CONDITIONAL EXPECTATIONS

BY J. PFANZAGL

1. Introduction. Let $(X, \mathfrak{A}, P)$ be a probability space and $\mathfrak{L}_r(X, \mathfrak{A}, P)$ the family of all $\mathfrak{A}$-measurable functions $f: X \to \mathfrak{R}$ such that $|f|^r$ is P-integrable. By $\|f\|_r$ we mean $P[|f|^r]$. (Here $P[g]$ denotes $\int g\, dP$.)

Let $\mathfrak{f}$ be a subfamily of $\mathfrak{L}_1(X, \mathfrak{A}, P)$ and $T: \mathfrak{f} \to \mathfrak{f}$ a given operator. This paper is concerned with the characterization of such operators as conditional expectations. The aim is to give conditions on T assuring that T is the restriction to $\mathfrak{f}$ of a conditional expectation with respect to some σ-algebra $\mathfrak{A}_0 \subset \mathfrak{A}$.

One such characterization was given by Moy (1954), p. 61, Theorem 2.2., :

M1: $T: \mathfrak{L}_1(X, \mathfrak{A}, P) \to \mathfrak{L}_1(X, \mathfrak{A}, P)$

M2: T is linear: $T(af + bg) = aTf + bTg$ for all $f, g \in \mathfrak{L}_1(X, \mathfrak{A}, P)$ and all $a, b \in \mathfrak{R}$,

M3: $\|T\| \leqq 1$

M4: T carries bounded functions into bounded functions,

M5: $T(fTg) = (Tf)(Tg)$ for all bounded $f, g \in \mathfrak{L}_1(X, \mathfrak{A}, P)$,

M6: T is constant preserving: $T1 = 1$.

This result was generalized by Rota (1960) and Olson (1965), p. 979, Theorem 3, mainly by eliminating condition M4 and using $\mathfrak{L}_r(X, \mathfrak{A}, P)$ $(r \geqq 1)$ in M1. The most interesting point in Olson's paper is to consider conditional expectations as integrals with respect to vector-valued measures with values in $\mathfrak{L}_1(X, \mathfrak{A}, P)$.

A similar characterization was given by Bahadur (1955), p. 566, Corollary 2:

B1: $T: \mathfrak{L}_2(X, \mathfrak{A}, P) \to \mathfrak{L}_2(X, \mathfrak{A}, P)$

B2: T is linear,

B3: T is positive: $f \geqq 0$ P-a.e. implies $Tf \geqq 0$ P-a.e.,

B4: T is idempotent: $T^2 = T$ P-a.e.

B5: T is self-adjoint: $P[f(Tg)] = P[(Tf)g]$,

B6: T is constant preserving.

This result was somewhat strengthened by Šidák (1957), p. 269, Theorem 4. Another characterization given by Šidák (p. 271, Theorem 6) uses the condition $T(Tf \vee Tg) = Tf \vee Tg$ P-a.e. instead of B3 (where $\vee$ denotes the pointwise maximum).

Finally, Douglas (1965), p. 453, Corollary 1 has given the following characterization:

D1: $T: \mathfrak{L}_1(X, \mathfrak{A}, P) \to \mathfrak{L}_1(X, \mathfrak{A}, P)$

D2: T is linear,

D3: $\|T\| = 1$,

D4: T is idempotent,

D5: T is constant preserving.

In the papers of Moy, Rota, Olson and Douglas the underlying probability

Received 2 May 1966; revised 21 September 1966.

measure enters only through the domain of definition of T, $\mathfrak{L}_r(X, \mathfrak{A}, P)$, and through the definition of $\| \ \|_r$. In the papers of Bahadur and Šidák a more explicit use of the probability measure is made by the condition of self-adjointness. The author thinks that some conditions on the operator may be weakened if a more effective use is made of the underlying probability measure, for instance by requiring expectation invariance. Considering the use made of conditional expectations in statistical theory, the condition of expectation invariance seems quite natural. It is needed e.g. in connection with the concepts 'power function' and 'unbiased estimation'.

The purpose of this paper is to give two characterizations of conditional expectations based on expectation invariance. The first characterization uses, except for expectation invariance and monotonicity, only properties of the range of T. The second characterization is closely related to the one given by Bahadur. It is, however, more general in so far as it is not restricted to functions in $\mathfrak{L}_2(X, \mathfrak{A}, P)$. The use of expectation invariance enables us to eliminate self-adjointness (B5) and to replace linearity (B2) by the weaker conditions of homogeneity and translation invariance. Finally, our proof is more elementary in so far as it avoids the Weierstrass-approximation theorem. This is done by working with $f \vee g$ instead of $f \cdot g$, an idea going back to Šidák (1957, p. 268).

An assumption like translation invariance might seem artificial at first sight. An operator missing this property is, however, hardly of use in general statistical theory: If we consider test functions, for instance, we want to conclude that $T(1 - \varphi) = 1 - T\varphi$. In the theory of estimation, we want to retain the Rao-Blackwell theorem on convex loss-functions C: $P[C(Tf)] \leqq P[C(f)]$. As the proof of this theorem is based on the envelopment of the convex function by straight lines we have to conclude $T(b + af) = b + aTf$. Monotonicity of the operator is needed in the theory of testing hypotheses ($0 \leqq \varphi \leqq 1$ implies $0 \leqq T\varphi \leqq 1$) as well as in estimation ($C(f) \geqq b + af$ implies $TC(f) \geqq T(b + af)$). Expectation invariance is needed in connection with the concepts 'power-function' and 'unbiased estimation'. Hence a set of properties indispensable for any operator useful in general statistical theory is: expectation invariance–monotonicity–homogeneity–translation invariance. We remark that for special problems more general operators might be of interest (see Brunk 1963).

In Theorem 3 it is shown that any expectation invariant, monotone, homogeneous, translation invariant and idempotent operator is a conditional expectation. Contrary to the other assumptions, idempotency is not required for the application. Generalizations of conditional expectations useful in general statistical theory might therefore be found in the domain of nonidempotent operators. However, to each constant preserving, monotone, expectation invariant and linear operator T, there exists an operator T_0 with the same properties which is stronger than T, (i.e. $T_0T = T$) and which is a conditional expectation operator (see *Le Cam*, Proposition 9, p. 1435). Hence useful generalizations have to be sought in the domain of *nonlinear* and *nonidempotent* operators.

2. First characterization. Let $\mathfrak{f} \subset \mathfrak{L}_1(X, \mathfrak{A}, P)$ be an arbitrary family and $T: \mathfrak{f} \to \mathfrak{f}$. For this and the next section, the following properties of an operator will be of basic importance:

Expectation invariance. $T \mid \mathfrak{f}$ is expectation invariant, if $P[Tf] = P[f]$ for all $f \in \mathfrak{f}$.

Monotonicity. $T \mid \mathfrak{f}$ is monotone if for all $f, g \in \mathfrak{f}, f \leqq g$ P-a.e. implies $Tf \leqq Tg$ P-a.e.

LEMMA 1. *A monotone and expectation invariant operator is continuous on pointwise convergent monotone sequences.*

PROOF. Let $f_n \uparrow f$ P-a.e., $f_n, f \in \mathfrak{f}$. As T is monotone, $(Tf_n)_{n=1,2,\cdots}$ is nondecreasing (P-a.e.) and $Tf_1 \leqq Tf_n \leqq Tf$ P-a.e. for $n = 1, 2, \cdots$. Hence $\lim_{n\to\infty} Tf_n$ exists P-a.e. and we have $\lim_{n\to\infty} Tf_n \leqq Tf$ P-a.e. As $f_1, f, Tf_1, Tf \in \mathfrak{L}_1(X, \mathfrak{A}, P)$ we obtain from the *Lebesgue* convergence theorem

$$P[\lim_{n\to\infty} Tf_n] = \lim_{n\to\infty} P[Tf_n] = \lim_{n\to\infty} P[f_n] = P[f] = P[Tf].$$

Hence $\lim_{n\to\infty} Tf_n = Tf$ P-a.e.

Let now $\mathfrak{f} \subset \mathfrak{L}_1(X, \mathfrak{A}, P)$ be a family with the following properties:

$\alpha 1$: $f \equiv a \in \mathfrak{f}$ for all $a \in \mathfrak{R}$

$\alpha 2$: $f, g \in \mathfrak{f}$, $A_1, A_2 \in \mathfrak{A}$ with $A_1 \cap A_2 = \emptyset$ implies $\chi_{A_1} f + \chi_{A_2} g \in \mathfrak{f}$.

LEMMA 2. *If $\mathfrak{f} \subset \mathfrak{L}_1(X, \mathfrak{A}, P)$ fulfils conditions $\alpha 1$ and $\alpha 2$, it contains all $\mathfrak{A}$-measurable simple functions.*

PROOF. $\alpha 2$ immediately extends to any finite number of functions and sets. Together with $\alpha 1$ the assertion follows.

THEOREM 1. *Let $\mathfrak{f} \subset \mathfrak{L}_1(X, \mathfrak{A}, P)$ be a family of functions fulfilling $\alpha 1$ and $\alpha 2$. Let $\mathfrak{A}_0 \subset \mathfrak{A}$ be a sub-σ-algebra and let $\mathfrak{f}_0$ be the system of all $\mathfrak{A}_0$-measurable functions in $\mathfrak{f}$.*

Assume that T is a monotone and expectation invariant operator on $\mathfrak{f}$ such that

(i) $T\mathfrak{f} \subset \mathfrak{f}_0$

(ii) $T\mathfrak{f} = \mathfrak{f}$ *P-a.e. for all $f \in \mathfrak{f}_0$.*

Then

$$T\chi_{A_0} f = \chi_{A_0} Tf \ P\text{-a.e.} \quad \text{for all} \quad f \in \mathfrak{f}, A_0 \in \mathfrak{A}_0, \tag{1}$$

i.e. T is the restriction to $\mathfrak{f}$ of a conditional expectation, given $\mathfrak{A}_0$.

PROOF. (a) Assume that (1) holds for all bounded functions in $\mathfrak{f}$. Let f be a function bounded from below. Let $A_- := \{x: f(x) < 0\}$, $A_+ := \{x: f(x) > 0\}$. We have $A_-, A_+ \in \mathfrak{A}$. Hence $f^+ := \chi_{A_+} f \in \mathfrak{f}$ by $\alpha 2$. Let $(e_n)_{n=1,2,\cdots}$ be a sequence of nonnegative simple functions approximating f^+ from below. Then $f_n := \chi_{A_-} f + \chi_{A_+} e_n$ is a sequence of bounded functions in $\mathfrak{f}$. Hence by assumption: $T\chi_{A_0} f_n = \chi_{A_0} Tf_n$ P-a.e. for all $A_0 \in \mathfrak{A}_0$ and all n. As $f_n \uparrow f$ and $\chi_{A_0} f_n \uparrow \chi_{A_0} f$, continuity of T on monotone sequences (Lemma 1) implies $T\chi_{A_0} f = \chi_{A_0} Tf$ P-a.e. Hence (1) holds for all functions in $\mathfrak{f}$ which are bounded from below. If $f \in \mathfrak{f}$ is arbitrary, then $f_{-n} := \chi_{A_{-n}} f + (-n)\chi_{\bar{A}_{-n}} \in \mathfrak{f}$ where $A_{-n} := \{x: f(x) \geqq -n\}$ for $n = 1, 2, \cdots$. As f_{-n} is bounded from below and as $f_{-n} \downarrow f$ and $\chi_{A_0} f_{-n} \downarrow \chi_{A_0} f$, continuity of T on monotone sequences again implies $T\chi_{A_0}f = \chi_{A_0}Tf$ P-a.e. Therefore (1) holds for all $f \varepsilon \mathfrak{f}$.

It remains to show that (1) holds for bounded functions in $\mathfrak{f}$.

(b) Let $f \varepsilon \mathfrak{f}$ be a bounded function ($c_1 \leqq f \leqq c_2$ P-a.e.) and let $A_0 \varepsilon \mathcal{A}_0$ be such that $\chi_{\bar{A}_0}f \varepsilon \mathfrak{f}_0$. Then

$$\chi_{\bar{A}_0}Tf = \chi_{\bar{A}_0}f \ P\text{-a.e.} \tag{2}$$

and

$$P[\chi_{A_0}Tf] = P[\chi_{A_0}f]. \tag{3}$$

We have $f_i := \chi_{A_0}c_i + \chi_{\bar{A}_0}f \varepsilon \mathfrak{f}_0$ and $f_1 \leqq f \leqq f_2$ P-a.e. Hence $f_1 = Tf_1 \leqq Tf \leqq Tf_2 = f_2$ P-a.e. Furthermore $\chi_{\bar{A}_0}f_i = \chi_{\bar{A}_0}f$ for $i = 1, 2$. Hence $\chi_{\bar{A}_0}Tf = \chi_{\bar{A}_0}f$ P-a.e. Finally, we have $P[\chi_{A_0}Tf] + P[\chi_{\bar{A}_0}Tf] = P[Tf] = P[f] = P[\chi_{A_0}f] + P[\chi_{\bar{A}_0}f]$. Together with (2) this implies (3).

(c) Let $f_1, f_2 \varepsilon \mathfrak{f}$ be bounded functions ($c_1 \leqq f_1 \leqq c_2$, $c_1 \leqq f_2 \leqq c_2$ P-a.e.) and $A_0 \varepsilon \mathcal{A}_0$ be such that $\chi_{A_0}f_1 = \chi_{A_0}f_2$ P-a.e., $\chi_{\bar{A}_0}f_1 \leqq \chi_{\bar{A}_0}f_2$ P-a.e. and $\chi_{\bar{A}_0}f_i \varepsilon \mathfrak{f}_0$ for $i = 1, 2$. Then

$$\chi_{A_0}Tf_1 = \chi_{A_0}Tf_2 \ P\text{-a.e.} \tag{4}$$

$f_1 \leqq f_2$ P-a.e. implies $Tf_1 \leqq Tf_2$ P-a.e. Together with (3) applied to f_1 and f_2 we obtain: $\chi_{A_0}Tf_1 = \chi_{A_0}Tf_2$ P-a.e.

(d) Let $f, g \varepsilon \mathfrak{f}$ be bounded functions ($c_1 \leqq f \leqq c_2$, $c_1 \leqq g \leqq c_2$ P-a.e.) and $A_0 \varepsilon \mathcal{A}_0$ be such that $\chi_{A_0}f = \chi_{A_0}g$ P-a.e. Then

$$\chi_{A_0}Tf = \chi_{A_0}Tg \ P\text{-a.e.} \tag{5}$$

Let $f_i := \chi_{A_0}f + c_i\chi_{\bar{A}_0}$. We have $f_1 \leqq f, g \leqq f_2$ P-a.e. which implies $Tf_1 \leqq Tf$, $Tg \leqq Tf_2$ P-a.e. Using (4) we obtain (5).

(e) Let $f \varepsilon \mathfrak{f}$ be bounded. Let $A_0 \varepsilon \mathcal{A}_0$ be arbitrary. Then (5) applied to f and $g := \chi_{A_0}f$ yields $\chi_{A_0}Tf = \chi_{A_0}T\chi_{A_0}f$ P-a.e. (2) applied to $\chi_{A_0}f$ instead of f yields $\chi_{\bar{A}_0}T\chi_{A_0}f = 0$ P-a.e. Hence $\chi_{A_0}Tf = T\chi_{A_0}f$ P-a.e.

REMARK. If $\mathfrak{f}$ contains nonnegative functions only, the proof of Theorem 1 works if $\mathfrak{f}$ contains all $\mathcal{A}$-measurable simple functions. (The steps b–e are then concerned with simple functions instead of bounded functions).

In order to give sufficient conditions for $\alpha 1$ and $\alpha 2$ we consider systems of functions $\mathfrak{f} \subset \mathfrak{L}_1(X, \mathcal{A}, P)$ with the following properties:

$\beta 1$: $f \varepsilon \mathfrak{f}$ implies $af \varepsilon \mathfrak{f}$ for all $a \varepsilon \mathcal{R}$,

$\beta 2$: $f \varepsilon \mathfrak{f}$ implies $1 + f \varepsilon \mathfrak{f}$,

$\beta 3$: $f, g \varepsilon \mathfrak{f}$ implies $f \wedge g \varepsilon \mathfrak{f}$ (where $\wedge$ denotes the pointwise minimum),

$\beta 4$: $f_n \downarrow f$ pointwise, $f_n \varepsilon \mathfrak{f}$ for $n = 1, 2, \cdots, f \varepsilon \mathfrak{L}_1(X, \mathcal{A}, P)$ implies $f \varepsilon \mathfrak{f}$.

REMARK. $\beta 1$ and $\beta 2$ together imply $b + af \varepsilon \mathfrak{f}$ for $f \varepsilon \mathfrak{f}$ and a, $b \varepsilon \mathcal{R}$. Furthermore $\beta 1$ and $\beta 3$ together imply $f \vee g \varepsilon \mathfrak{f}$ for $f, g \varepsilon \mathfrak{f}$.

LEMMA 3. *If $\mathfrak{f} \neq \varnothing$ fulfills $\beta 1$–$\beta 4$, $\mathcal{A}_* := \{A \varepsilon \mathcal{A}: \chi_A \varepsilon \mathfrak{f}\}$ is a σ-algebra and $\mathfrak{f} = \mathfrak{L}_1(X, \mathcal{A}_*, P)$.*

PROOF. As $\mathfrak{f} \neq \varnothing$, $\beta 1$ and $\beta 2$ together imply $1 \varepsilon \mathfrak{f}$ whence $X \varepsilon \mathcal{A}_*$. Hence

$\mathfrak{A}_* \neq \emptyset$. If $A \in \mathfrak{A}_*$, then $\chi_A \in \mathfrak{f}$ whence $1 - \chi_A \in \mathfrak{f}$ whence $\bar{A} \in \mathfrak{A}_*$. If $A_n \in \mathfrak{A}_*$, then $\chi_{A_n} \in \mathfrak{f}$ whence $f_n := \chi_{A_1} \wedge \cdots \wedge \chi_{A_n} \in \mathfrak{f}$ whence by $\beta 4$, $\chi_{\bigcap_1^\infty A_n} = \lim_{n\to\infty} f_n \in \mathfrak{f}$ whence $\bigcap_1^\infty A_n \in \mathfrak{A}_*$. Hence $\mathfrak{A}_*$ is a σ-algebra. To show $\mathfrak{A}_*$-measurability for $f \in \mathfrak{f}$, let $\mathcal{S} := \{f^{-1}[a, \infty): f \in \mathfrak{f}, a \in \mathfrak{R}\}$. Because $f \in \mathfrak{f}$ implies $f - (a - 1) \in \mathfrak{f}$, we have $\mathcal{S} = \{f^{-1}[1, \infty): f \in \mathfrak{f}\}$. Because $f \in \mathfrak{f}$ implies $(0 \vee (1 \wedge f)) \in \mathfrak{f}$, we have: $\mathcal{S} = \{f^{-1}\{1\}: f \in \mathfrak{f}, 0 \leqq f \leqq 1\}$. Finally $f \in \mathfrak{f}, 0 \leqq f \leqq 1$, implies $f_n := 1 - (1 \wedge n(1 - f)) \in \mathfrak{f}$ and $f_0 = \lim_{n\to\infty} f_n$ is an indicator function such that $f_0^{-1}\{1\} = f^{-1}\{1\}$. Hence $\mathcal{S} = \{f^{-1}\{1\}: f \in \mathfrak{f}, f$ indicator function$\}$. Therefore, $\mathcal{S}$ is identical to $\mathfrak{A}_*$. As the system $\{[a, \infty): a \in \mathfrak{R}\}$ generates the Borel-algebra, all functions of $\mathfrak{f}$ are measurable with respect to the σ-algebra generated by $\mathcal{S}$. As $\mathcal{S}$ itself is a σ-algebra, namely $\mathfrak{A}_*$, all functions of $\mathfrak{f}$ are $\mathfrak{A}_*$-measurable.

Finally we show that any $\mathfrak{A}_*$-measurable function in $\mathfrak{L}_1(X, \mathfrak{A}, P)$ belongs to $\mathfrak{f}$. It is immediately clear that any nonnegative $\mathfrak{A}_*$-measurable simple function belongs to $\mathfrak{f}$, for $\sum_1^n a_i\chi_{A_i} = \bigvee_1^n a_i\chi_{A_i}$ ($A_1, \cdots, A_n$ pairwise disjoint). If $f \in \mathfrak{L}_1(X, \mathfrak{A}, P)$ is $\mathfrak{A}_*$-measurable and nonnegative, it is the limit of a nondecreasing sequence of nonnegative $\mathfrak{A}_*$-measurable simple functions and hence (by $\beta 1$ and $\beta 4$) belongs to $\mathfrak{f}$. Let finally $f \in \mathfrak{L}_1(X, \mathfrak{A}, P)$ be an arbitrary $\mathfrak{A}_*$-measurable function. Then $(f + n)^+$ is $\mathfrak{A}_*$-measurable and nonnegative. As $(f + n)^+ \in \mathfrak{L}_1(X, \mathfrak{A}, P)$ we have therefore $(f + n)^+ \in \mathfrak{f}$. Hence $(f + n)^+ - n \in \mathfrak{f}$ for $n = 1, 2, \cdots$. As $(f + n)^+ - n \downarrow f$ and $f \in \mathfrak{L}_1(X, \mathfrak{A}, P)$, we have $f \in \mathfrak{f}$.

THEOREM 2. *Let $T \mid \mathfrak{f}$ be a monotone and expectation invariant operator on a system of functions $\mathfrak{f} \neq \emptyset$ fulfilling $\beta 1$–$\beta 4$. Let $\mathfrak{f}_0 \subset \mathfrak{f}$ be a subsystem of functions fulfilling $\beta 1$–$\beta 4$ such that*

(i) $T\mathfrak{f} \subset \mathfrak{f}_0$

(ii) *$Tf = f$ P-a.e. for all $f \in \mathfrak{f}_0$.*

Then $\mathfrak{A}_0 := \{A \in \mathfrak{A}: \chi_A \in \mathfrak{f}_0\}$ is a σ-algebra and $T \mid \mathfrak{f}$ is the restriction to $\mathfrak{f}$ of a conditional expectation, given $\mathfrak{A}_0$.

PROOF. According to Lemma 3: $\mathfrak{f} = \mathfrak{L}_1(X, \mathfrak{A}_*, P)$ and $\mathfrak{f}_0 = \mathfrak{L}_1(X, \mathfrak{A}_0, P)$ with $\mathfrak{A}_0 \subset \mathfrak{A}_*$. Hence $\mathfrak{f}$ fulfills $\alpha 1$ and $\alpha 2$ and $\mathfrak{f}_0$ is the system of all $\mathfrak{A}_0$-measurable functions in $\mathfrak{f}$. Hence the assertion follows immediately from Theorem 1.

3. Second characterization. Theorem 1 depends on the rather strong assumption (ii) that T leaves the functions in $\mathfrak{f}_0$ invariant. The purpose of this section is to replace this assumption by other assumptions on T. In addition to the assumption of monotonicity and expectation invariance we will assume that T is idempotent. (Idempotency is an immediate consequence of (i) and (ii).) Furthermore, we need the following properties:

Homogeneity. T is homogeneous, if $Taf = aTf$ P-a.e. for all f and all $a \in \mathfrak{R}$ (for which these terms are defined).

Translation invariance. T is translation invariant, if $T(1 + f) = 1 + Tf$ P-a.e. for all f (for which these terms are defined).

The natural domain of definition of an operator with these properties is a family $\mathfrak{f} \subset \mathfrak{L}_1(X, \mathfrak{A}, P)$ which fulfills $\beta 1$–$\beta 4$.

REMARK. Let $T \mid \mathfrak{f}$ be an expectation invariant, monotone and constant preserving operator, defined on a family $\mathfrak{f} \subset \mathfrak{L}_1(X, \mathfrak{A}, P)$ which is closed under $\vee$ and contains all constants. Under these conditions, idempotency of T is equivalent to the following property (used by Šidák (1957), p. 271, Theorem 6):

$$T(Tf \vee Tg) = Tf \vee Tg \ P\text{-a.e.}$$

PROOF. (i) Monotonicity of T implies $T(f \vee g) \geqq Tf \vee Tg$ P-a.e. Applied to Tf, Tg (instead of f, g), this relation together with idempotency yields $T(Tf \vee Tg) \geqq Tf \vee Tg$ P-a.e. Using expectation invariance, we obtain $T(Tf \vee Tg) = Tf \vee Tg$ P-a.e.

(ii) As T is constant preserving, Šidák's condition implies $T((Tf) \vee (-n)) = (Tf) \vee (-n)$ P-a.e. for all integers n. According to Lemma 1, T is continuous on monotone limits. Hence for $n \to \infty$ we obtain: $T(Tf) = Tf$ P-a.e.

LEMMA 4. *If $T \mid \mathfrak{f}$ is an expectation invariant, monotone, homogeneous, translation invariant and idempotent operator defined on a family $\mathfrak{f} \neq \emptyset$ with properties $\beta 1$–$\beta 4$, the family $\mathfrak{f}_0 := \{f \varepsilon \mathfrak{f}: Tf = f \ P\text{-a.e.}\}$ has properties $\beta 1$–$\beta 4$ too. Furthermore we have $T\mathfrak{f} = \mathfrak{f}_0$ P-a.e.*[1]

PROOF. The relation $T\mathfrak{f} = \mathfrak{f}_0$ P-a.e. is obvious from the idempotency of T. That $\mathfrak{f}_0$ has properties $\beta 1$, $\beta 2$ and $\beta 4$ is obvious from homogeneity, translation invariance and Lemma 1, respectively. It remains to show that $f, g \varepsilon \mathfrak{f}_0$ implies $f \wedge g \varepsilon \mathfrak{f}_0$. Monotonicity of T implies $T(f \wedge g) \leqq (Tf) \wedge (Tg)$ P-a.e. If $f, g \varepsilon \mathfrak{f}_0 : T(f \wedge g) \leqq f \wedge g$ P-a.e. Together with expectation invariance we obtain $T(f \wedge g) = f \wedge g$ P-a.e.

THEOREM 3. *If $T \mid \mathfrak{f}$ is an expectation invariant, monotone, homogeneous, translation invariant and idempotent operator, defined on a family $\mathfrak{f} \neq \emptyset$ with properties $\beta 1$–$\beta 4$, then T is the restriction to $\mathfrak{f}$ of a conditional expectation with respect to the σ-algebra $\mathfrak{A}_0 := \{A \varepsilon \mathfrak{A}: \chi_A \varepsilon \mathfrak{f}_0\}$ where $\mathfrak{f}_0$ is defined as in Lemma 4.*

PROOF. According to Lemma 4, $\mathfrak{f}_0 := \{f \varepsilon \mathfrak{f}: Tf = f \ P\text{-a.e.}\}$ has properties $\beta 1$–$\beta 4$. Condition (ii) of Theorem 2 is fulfilled by definition of $\mathfrak{f}_0$; condition (i) is a consequence of idempotency. Hence the assertion follows immediately from Theorem 2.

We remark that the assumptions on T made in Theoem 3 are less restrictive than the assumptions made by Bahadur. First of all the domain of definition of T is an arbitrary subfamily of $\mathfrak{L}_1(X, \mathfrak{A}, P)$ fulfilling $\beta 1$–$\beta 4$ whereas Bahadur requires $\mathfrak{f} = \mathfrak{L}_2(X, \mathfrak{A}, P)$. Expectation invariance is an immediate consequence of B5 and B6 (taking $f = 1$). Monotonicity follows from B2 and B3. Furthermore B2 implies homogeneity, B2 and B6 together imply translation invariance. Hence Theorem 3 is stronger than Bahadur's characterization (Corollary 2, p. 566).

A number of easy examples show that the properties of the operator required in Theorem 3 (expectation invariance, monotonicity, homogeneity, translation invariance and idempotency) are mutually independent.

Acknowledgment. The author thanks Mr. W. Pierlo and the referees for critical comments on earlier versions of the manuscript.

[1] $\mathfrak{f}_1 \subset \mathfrak{f}_2$ P-a.e. if to each $f_1 \in \mathfrak{f}_1$ there exists $f_2 \in \mathfrak{f}_2$ such that $f_1 = f_2$ P-a.e.; $\mathfrak{f}_1 = \mathfrak{f}_2$ P-a.e. if $\mathfrak{f}_1 \subset \mathfrak{f}_2$ P-a.e. and $\mathfrak{f}_2 \subset \mathfrak{f}_1$ P-a.e.

REFERENCES

BAHADUR, R. R. (1955). Measurable subspaces and subalgebras. *Proc. Amer. Math. Soc.* **6** 565–570.

BRUNK, H. D. (1963). On an extension of the concept conditional expectation. *Proc. Amer. Math. Soc.* **14** 298–304.

DOUGLAS, R. G. (1965). Contractive projections on an L_1 space. *Pacific J. Math.* **15** 443–462.

LE CAM, L. (1964). Sufficiency and approximate sufficiency. *Ann. Math. Statist.* **35** 1419–1455.

MOY, S. C. (1954). Characterizations of conditional expectation as a transformation on function spaces. *Pacific J. Math.* **4** 47–63.

OLSON, M. P. (1965). A characterization of conditional probability. *Pacific J. Math.* **15** 971–983.

ROTA, G. C. (1960). On the representations of averaging operators. *Rendiconti Sem. Mat. Univ. Padova*, **30** 52–64.

ŠIDÁK, Z. (1957). On relations between strict-sense and wide-sense conditional expectations. *Theory Prob. Appl.* **2** 267–271.