

## A GENERALIZED RADON-NIKODYM DERIVATIVE

H. D. BRUNK AND S. JOHANSEN

Let  $\{\nu_a, a \in R\}$  be a family of signed measures on a  $\sigma$ -field  $\mathcal{A}$  of subsets of an abstract space  $\Omega$ . Let  $\mathcal{M}$  be a sub-lattice of  $\mathcal{A}$ . Under certain conditions we associate with the family of measures and  $\mathcal{M}$  a function  $f$ , which we call the Lebesgue-Radon-Nikodym (LRN) function. The function  $f$  is measurable  $\mathcal{M}$  and satisfies the relations

$$\begin{aligned} \nu_a(B \cap [f < a]) &\leq 0, \quad a \in R, \quad B \in \mathcal{M}, \\ \nu_b(C \cap [f > b]) &\geq 0, \quad b \in R, \quad C \in \mathcal{M}^c. \end{aligned}$$

This paper contains a construction of  $f$  by means of a Jordan-Hahn decomposition for  $\sigma$ -lattices, and gives various characterizations and representations of  $f$ .

Special cases are: the derivative of a signed measure with respect to a nonnegative measure, conditional expectation given a  $\sigma$ -field, and conditional expectation given a  $\sigma$ -lattice. The LRN function also provides a conditional generalized mean whose relationship to the generalized mean parallels the relationship of the conditional expectation to the expectation.

The paper also contains a convergence theorem for LRN functions with respect to an increasing sequence of  $\sigma$ -lattices, thus generalizing the martingale convergence theorem.

Finally it is proved that  $f$  is the solution to a minimization problem, generalizing known minimizing properties of conditional expectation and of conditional expectation given a  $\sigma$ -lattice. These properties exhibit the latter as solution of various problems of restricted maximum likelihood estimation.

Section 1 of this paper establishes the existence and uniqueness of the LRN function of a family  $\{\nu_a, a \in R\}$  satisfying conditions (1.1), (1.2), and (1.3) (Theorems 1.5 and 1.7). We thereby generalize the classical case  $\nu_a = \varphi - a\mu$ ,  $\mathcal{M} = \mathcal{M}^c$ , where the LRN function  $f$  is the ordinary Radon-Nikodym function of  $\varphi$  with respect to  $\mu$  (Example 1.15). We then prove a representation theorem for  $f$  (Theorem 1.12) in terms of mean values of sets. This theorem gives a precise way of expressing how- in the classical case  $-\varphi(A)/\mu(A)$  converges to the Radon-Nikodym function  $f$  at  $\omega$  as  $A \downarrow \{\omega\}$ . We show how the LRN function in the general case can be found if we assume the case  $\nu_a = \varphi - a\mu$  known. The function  $f$  is essentially the solution of the equation  $d\nu_a/d\mu = 0$ . Finally we consider some examples.

Section 2 generalizes the result that in the classical case ( $\nu_a = \varphi - a\mu$ ,  $\mathcal{M} = \mathcal{M}^c$ ) we can characterize a Radon-Nikodym derivative as an  $\mathcal{M}$  measurable function whose indefinite  $\mu$ -integral is  $\varphi$  (Theorem

2.5). In this section we have also included a convergence theorem for LRN functions on increasing  $\sigma$ -lattices thus generalizing the martingale convergence theorem.

The last section contains the application to minimization problems. For a given family  $\{\nu_a\}$  of measures we define a minimization problem to which the LRN function associated with  $\{\nu_a\}$  is the solution. We then discuss how the family  $\{\nu_a\}$  is constructed when a family of quasi convex functions is given.

The LRN function furnishes the maximum likelihood estimate of ordered parameters in sampling from the generalized exponential families introduced in [6]. For maximum likelihood estimation of ordered parameters in sampling from more general unimodal distributions, studied by Eeden [12] and by Robertson and Waltman [18], the solution is shown to be a LRN function; and for each such LRN function a wide class of such problems is exhibited for which it furnishes the solution.

1. The Lebesgue-Radon-Nikodym function. The fundamental tool in this section is the Jordan-Hahn decomposition theorem proved in Johansen [15]. We follow the notation of that paper. Let  $\Omega$  be an abstract space, and let  $\mathcal{A}$  be a  $\sigma$ -field of subsets of  $\Omega$ . Let  $\mathcal{M}$  be a  $\sigma$ -lattice of sets from  $\mathcal{A}$ , such that  $\emptyset \in \mathcal{M}, \Omega \in \mathcal{M}$ . Define  $\mathcal{F} = \{A: A = B \cap C, B \in \mathcal{M}, C \in \mathcal{M}^c\}$ . We use the term measure for a  $\sigma$ -additive extended real valued function on  $\mathcal{F}$  taking at most one of the values  $+\infty, -\infty$ .

Let  $\nu$  be a measure on  $\mathcal{F}$ . A set  $B \in \mathcal{M}$  is called positive, if for all  $C \in \mathcal{M}^c$  we have  $\nu(B \cap C) \geq 0$ . A set  $C \in \mathcal{M}^c$  is called negative if for all  $B \in \mathcal{M}$  we have  $\nu(B \cap C) \leq 0$ . Let  $\mathcal{P}$  be the family of positive sets, and let  $\mathcal{N}$  be the family of negative sets, then  $\emptyset \in \mathcal{N} \cap \mathcal{P}$ , and  $\mathcal{P}$  and  $\mathcal{N}$  are closed under countable unions. A set  $C_0 \in \mathcal{N}$  is called minimal if  $\nu(C_0) = \inf_{C \in \mathcal{N}} \nu(C)$  and  $B_0 \in \mathcal{P}$  is called maximal if  $\nu(B_0) = \sup_{B \in \mathcal{P}} \nu(B)$ . It is seen that both maximal and minimal sets exist. The following theorem and corollary are given in Johansen [5].

**THEOREM 1.1.** *If  $\nu$  is a measure on  $\mathcal{F}$ , if  $\nu < \infty$ , and if  $A$  is a maximal set for  $\nu$ , then  $A^c$  is negative, and in fact minimal.*

**COROLLARY 1.2.** *(Jordan-Hahn decomposition). If  $\nu$  is a measure on  $\mathcal{F}$ , it admits a maximal set  $A^+ \in \mathcal{M}$  and a minimal set  $A^- \in \mathcal{M}^c$ , such that  $A^+ = (A^-)^c$ .*

Let  $R$  denote the set of real numbers, and  $R'$  the set of rationals.

Let  $\nu. (\cdot)$  be a mapping from  $R \times \mathcal{F}$  into the extended real numbers with the following properties:

- (1.1) For each  $a \in R, \nu_a(\cdot)$  is a measure on  $\mathcal{F}$ .
- (1.2) If  $A \in \mathcal{F}$  and  $a < b$ , then

$$\nu_b(A) \geq 0 \Rightarrow \nu_a(A) \geq 0 ,$$

and

$$\nu_a(A) \leq 0 \Rightarrow \nu_b(A) \leq 0 .$$

- (1.3) There exists an real number  $\alpha$ , such that for  $A \in \mathcal{F}$

$$\nu_a(A) < \infty , \quad a > \alpha ,$$

and

$$\nu_a(A) > -\infty , \quad a < \alpha .$$

In this paper  $\{\nu_a, a \in R\}$  will-unless otherwise stated-denote a family of measures satisfying (1.2) and (1.3).

Notice that if  $\{\nu_a, a \in R\}$  satisfy (1.1) and (1.2) then there exists a nonnegative measure  $\mu$  which dominates  $\nu_a, a \in R$ . We can choose  $\mu = \sum_{r \in R} |\nu_r|$ . If  $\nu_a$  is  $\sigma$ -finite for  $a$  in a countable set dense in  $R$ , we can choose  $\mu$  finite.

**LEMMA 1.3.** *Let  $\{\nu_a, a \in R\}$  be given. Then there exists for each  $r, r$  rational, a  $\nu_r$  positive set  $M_r^+$  such that  $M_r^- = (M_r^+)^c$  is  $\nu_r$  negative, and such that  $M_r^+$  is decreasing in  $r$ .*

*Proof.* If  $-\infty < \alpha < \infty$ , choose  $A_\alpha^+$  positive for  $\nu_\alpha$ , such that  $A_\alpha^- = (A_\alpha^+)^c$  is negative for  $\nu_\alpha$ . If  $\alpha = -\infty$  let  $A_\alpha^+ = \Omega, A_\alpha^- = \emptyset$ , if  $\alpha = +\infty$ , let  $A_\alpha^+ = \emptyset, A_\alpha^- = \Omega$ . Consider then set functions

$$\begin{aligned} \nu_\alpha^+(A) &= \nu_\alpha(A_\alpha^+ \cap A) \\ \nu_\alpha^-(A) &= -\nu_\alpha(A_\alpha^- \cap A) \end{aligned}$$

and the families of sets

$$\begin{aligned} \mathcal{M}^+ &= A_\alpha^+ \cap \mathcal{M}, \mathcal{F}^+ = A_\alpha^+ \cap \mathcal{F}, (\mathcal{M}^+)^c = A_\alpha^+ \cap \mathcal{M}^c, \\ \mathcal{M}^- &= A_\alpha^- \cap \mathcal{M}, \mathcal{F}^- = A_\alpha^- \cap \mathcal{F}, (\mathcal{M}^-)^c = A_\alpha^- \cap \mathcal{M}. \end{aligned}$$

Then  $\{\nu_\alpha^+, A_\alpha^+, \mathcal{F}^+, \mathcal{M}^+\}$  and  $\{\nu_\alpha^-, A_\alpha^-, \mathcal{F}^-, \mathcal{M}^-\}$  satisfy (1.1), (1.2), and (1.3) with  $\alpha^+ = \alpha, \alpha^- = -\alpha$ .

In the following  $r$  and  $s$  are rationals. For  $r > \alpha^+$  let  $B_r^+ \in \mathcal{M}^+$  be maximal positive for  $\nu_r^+$ , and let  $C_r^+ \in \mathcal{M}^-$  be maximal positive for  $\nu_r^-, r > -\alpha$ . Then  $B_r^+ \subset A_\alpha^+$  and  $C_r^+ \subset A_\alpha^-$ .

Define

$$M_r^+ = \bigcup_{s \geq r, s \in R'} B_s^+, \quad r > \alpha$$

and

$$M_r^- = \bigcup_{s \leq r, s \in R'} C_{-s}^+, \quad r < \alpha,$$

and let  $M_r^+ = (M_r^-)^c$ .

Since  $B_s^+$  is positive for  $\nu_s^+$  it follows that it is positive for  $\nu_s$ , since  $B_s^+ \in \mathcal{M}$ , and for  $C \in \mathcal{M}^c$ , we have

$$\nu_s(B_s^+ \cap C) = \nu_s^+(B_s^+ \cap (C \cap A_\alpha^+)) \geq 0.$$

Also  $B_s^+$  is maximal for  $\nu_s$ . To see this, note first that for  $B \in \mathcal{M}$ ,  $\nu_s(B) = \nu_s(B \cap A_\alpha^+) + \nu_s(B \cap A_\alpha^-)$ . Now  $A_\alpha^-$  is negative for  $\nu_\alpha$  and therefore by (1.2) for  $\nu_s$ , thus

$$\nu_s(B) \leq \nu_s(B \cap A_\alpha^+) = \nu_s^+(B \cap A_\alpha^+) \leq \nu_s^+(B_s^+) = \nu_s(B_s^+).$$

We now want to prove that  $M_r^+$  is positive, and that  $M_r^-$  is negative for  $\nu_r$ . For  $s > r > \alpha$  we have that  $B_s^+$  is positive for  $\nu_s$  and hence for  $\nu_r$  by (1.2). Therefore  $M_r^+$  is positive for  $\nu_r$  and since it contains a maximal set, it is maximal, and by Theorem 1.1  $M_r^-$  is negative. Similarly it is seen that  $M_r^-$  is negative and minimal for  $\nu_r$ ,  $r < \alpha$  and hence  $M_r^+$  is positive. Clearly  $M_r^+$  is decreasing  $r > \alpha$  and for  $r < \alpha$ . But since  $M_r^+ \subset A_\alpha^+ \subset M_s^+$ ,  $r > \alpha > s$ , we get that  $M_r^+$  is decreasing in  $r$ .

DEFINITION 1.4. An extended real valued function  $f$  on  $\Omega$  is a Lebesgue-Radon-Nikodym (LRN) function of  $\{\nu_a, a \in R\}$  and  $\mathcal{M}$ , if

(1.4)  $f$  is measurable  $\mathcal{M}$ ,

(1.5)  $\nu_b(B \cap [f < b]) \leq 0, b \in R, B \in \mathcal{M}$ ,

(1.6)  $\nu_a(C \cap [f > a]) \geq 0, a \in R, C \in \mathcal{M}^c$ .

We note that it is sufficient that (1.4) holds and that (1.5) and (1.6) hold for  $a$  and  $b$  in a dense set  $Q$ . For suppose  $c \in Q$ , and  $a > c, a \in Q$ , then by (1.6)  $\nu_a(C \cap [f > a]) \geq 0$ , and by (1.2) we get  $\nu_c(C \cap [f > a]) \geq 0$ . For  $a \downarrow c$ , we get  $\nu_c(C \cap [f > c]) \geq 0$ . The proof of (1.5) is similar.

THEOREM 1.5. Let  $\{\nu_a, a \in R\}$  satisfy (1.1), (1.2), and (1.3). Let  $M_r^+, r \in R'$ , be chosen as in Lemma 1.3, then the function

$$\begin{aligned} f &= \sup_{r \in R'} (rIM_r^+ - \infty IM_r^-) \\ &= \inf_{r \in R'} (rIM_r^+ + \infty IM_r^+) \end{aligned}$$

is a LRN function of  $\{\nu_a\}$  and  $\mathcal{M}$ . Here  $IM$  denotes the function which is 1 on  $M$  and 0 on  $M^c$ .

*Proof.* To prove (1.4) we remark that

$$[f > t] = \bigcup_{r>t, r \in R'} M_r^+ \in \mathcal{M}.$$

Since  $M_r^+$  is  $\nu_r$  positive, it follows from (1.2) that  $M_r^+$  is  $\nu_t$  positive, and hence that  $[f > t]$  is  $\nu_t$  positive which proves (1.6). Similarly

$$[f < t] = \bigcup_{r>t, r \in R'} M_r^- \in \mathcal{M}^c$$

is  $\nu_t$  negative, which proves (1.5).

We now turn to the uniqueness of the LRN function.

**DEFINITION 1.6.** We call the family  $\{\nu_a, a \in R\}$  *decreasing at zero with respect to  $\mu$*  if, for all  $A \in \mathcal{F}$ , the relations

$$a < b, \quad \nu_a(A) = \nu_b(A) = 0$$

imply that  $\mu(A) = 0$ .

We note that if  $\nu_a$  is dominated by  $\mu$  for all  $a$ , then  $\{\nu_a\}$  satisfies (1.2) and is decreasing at zero with respect to  $\mu$  if and only if  $\{\nu_a\}$  satisfies a strict form of (1.2)

(1.2)' If  $A \in \mathcal{F}$ ,  $\mu(A) > 0$ ,  $a > b$ , then

$$\nu_a(A) \geq 0 \Rightarrow \nu_b(A) > 0.$$

**THEOREM 1.7.** Let  $\{\nu_a, a \in R\}$  be decreasing at zero with respect to  $\mu$ . Let  $\mu$  be defined on  $\mathcal{F}$ . Then any two derivatives coincide  $\mu$  almost surely.

*Proof.* Let  $f$  and  $g$  be two derivatives. We want to prove that  $\mu([f \neq g]) = 0$ . It is enough to verify that  $A = [f > a > b > g]$  is a  $\mu$ -null set.

By (1.6) we get  $\nu_a(A) \geq 0$ , but then by (1.2)  $\nu_b(A) \geq 0$ . By (1.5) and (1.2) we get  $\nu_b(A) \leq 0$  and  $\nu_a(A) \leq 0$ . Hence

$$\nu_a(A) = \nu_b(A) = 0$$

which by assumption implies that  $\mu(A) = 0$ .

It is seen that the definition of a LRN function  $f$  requires that  $[f > a]$  is positive for  $\nu_a$ . It is not in general true that  $[f > a]$  is maximal or that  $[f \leq a]$  is negative. But we now formulate two theorems which illustrate this point. These theorems will be applied in §3.

**THEOREM 1.8.** *Assume that  $\nu_a, a \in R$ , is  $\sigma$ -finite. Then  $\nu_a$  is dominated by some finite  $\mu$ , and  $f$  is a LRN function associated with  $\{\nu_a, a \in R\}$  and  $\mathcal{M}$  if and only if there exists a countable set  $D$  such that*

- (1)  $f$  measurable  $\mathcal{M}$ ,
- (2)  $\nu_b(B \cap [f \leq b]) \leq 0, B \in \mathcal{M}, b \notin D$ ,
- (3)  $\nu_a(C \cap [f \geq a]) \geq 0, C \in \mathcal{M}^c, a \notin D$ .

*Proof.* Let  $f$  be a LRN function associated with  $\{\nu_a, a \in R\}$  on  $\mathcal{M}$ , and let  $D = \{a \mid \mu([f = a]) > 0\}$ , then  $D$  is countable, and if  $b \notin D$ , then  $[f = b]$  is a  $\mu$ -null set, and hence a  $\nu_b$ -null set. Hence

$$\nu_b(B \cap [f \leq b]) = \nu_b(B \cap [f < b]) \leq 0,$$

by (1.5). Inequality (3) is proved similarly.

If on the other hand  $f$  satisfies (1), (2), and (3) for some countable set  $D_1$ , then we extend  $D_1$  by  $D$  defined above, and we get that (1.5) and (1.6) are satisfied for  $a$  and  $b$  in the dense set  $(D_1 \cup D)^c$ , which by the remark following the definition is enough to ensure that  $f$  is a LRN function.

**THEOREM 1.9.** *If  $\nu_a$  is finite and if  $\nu \cdot (A)$  is continuous,  $A \in \mathcal{F}$ , then  $f$  is a LRN function if and only if*

- (1)  $f$  measurable  $\mathcal{M}$
- (2)  $\nu_b(B \cap [f \leq b]) \leq 0, B \in \mathcal{M}$ ,
- (3)  $\nu_a(C \cap [f \geq a]) \geq 0, C \in \mathcal{M}^c$ .

*Proof.* If  $f$  is a LRN function then for  $b < c < d$  we get from (1.5)

$$\nu_c(B \cap [f < c]) \leq 0, \quad B \in \mathcal{M},$$

which by (1.2) implies that

$$\nu_d(B \cap [f < c]) \leq 0, \quad B \in \mathcal{M}.$$

If we let  $c \downarrow b$ , and then  $d \downarrow b$ , we get (2). Inequality (3) is proved similarly. The fact that (1), (2), and (3) imply (1.4), (1.5), and (1.6) was established in Theorem 1.8.

Before we proceed to give a different representation of the LRN function we state a corollary of Theorem 1.5 which is just a reformulation of the construction given there.

Let  $\mu$  be a finite nonnegative measure on  $(\Omega, \mathcal{A})$ . Let  $\{f_i, i \in I\}$  be any family of real functions on  $\Omega$ . The function  $f = \text{ess sup}_{i \in I} f_i$  is defined to within  $\mu$ -equivalence by the relations:

$$f_i \leq g, \mu.\text{a.s. } i \in I \iff f \leq g, \mu.\text{a.s.}$$

For a reference see Neveu [16]. It can be proved that there exists a countable set  $Q \subset I$  such that  $f = \sup_{i \in Q} f_i, \mu.\text{a.s.}$

**COROLLARY 1.10.** *Let  $\{\nu_a, a \in R\}$  be dominated by  $\mu$ . For each  $a > \alpha$  choose  $A_a^+ \subset A_\alpha^+$  maximal for  $\nu_a$ , and for  $a < \alpha$  choose  $A_a^- \subset A_\alpha^-$  minimal for  $\nu_a$ . Let  $A_\alpha^+ = (A_\alpha^-)^c, \alpha \in R$ . Then the function*

$$f = \begin{cases} \text{ess sup}_{a > \alpha} (aIA_a^+ + \alpha IA_a^-)IA_\alpha^+ \\ + \text{ess inf}_{a < \alpha} (aIA_a^- + \alpha IA_a^+)IA_\alpha^- \end{cases}$$

is  $\mu$  almost surely equal to a LRN function of  $\{\nu_a, a \in R\}$  and  $\mathcal{M}$ .

*Proof.* There exists a countable dense set  $Q$ , such that  $f = g, \mu.\text{a.s.}$ ,

$$g = \begin{cases} \sup_{r > \alpha, r \in Q} (rIA_r^+ + \alpha IA_r^-)IA_\alpha^+ \\ + \inf_{r < \alpha, r \in Q} (rIA_r^- + \alpha IA_r^+)IA_\alpha^- \end{cases}$$

We now want to prove that  $g$  is a LRN function.

For notational convenience  $r$  and  $s$  will in the following denote elements of  $Q$  and the unions below are taken over those  $r$  or  $s$  in  $Q$ , for which the conditions indicated in the unions are satisfied. For  $t \geq \alpha$  we have

$$[g > t] = \bigcup_{r > t} A_r^+ \in \mathcal{M},$$

and this set is positive for  $\nu_t$ . For  $t < \alpha$ , we get

$$\begin{aligned} [g > t] &= A_\alpha^+ \cup \bigcup_{\alpha > r > t} [g < r]^c \\ &= A_\alpha^+ \cup \bigcup_{\alpha > r > t} (\bigcup_{s < r} A_s^-)^c \\ &= A_\alpha^+ \cup \bigcup_{\alpha > r > t} (\bigcup_{s \leq r} A_s^-)^c. \end{aligned}$$

But now  $\bigcup_{s \leq r} A_s^-$  is negative and minimal for  $\nu_r$ , and hence, since  $r < \alpha$  implies that  $\nu_r > -\infty$ , we get by Theorem 1.1 that  $(\bigcup_{s \leq r} A_s^-)^c$  is positive for  $\nu_r$ , and therefore by (1.2) for  $\nu_t$ . Finally the set  $\bigcup_{\alpha > r > t} (\bigcup_{s \leq r} A_s^-)^c$  is positive for  $\nu_t$ , and hence  $[g > t]$  is positive for  $\nu_t$ . Thus (1.6) is proved, and (1.5) is proved similarly.

**DEFINITION 1.11.** For  $A \in \mathcal{F}$  we define

$$\begin{aligned} \bar{M}(A) &= \sup \{a: \nu_a(A) \geq 0\} \\ &= \inf \{a: \nu_a(A) < 0\} \end{aligned}$$

and

$$\begin{aligned} \underline{M}(A) &= \sup \{a: \nu_a(A) > 0\} \\ &= \inf \{a: \nu_a(A) \leq 0\} . \end{aligned}$$

If  $\bar{M}(A) = \underline{M}(A)$  for some set  $A \in \mathcal{A}$ , we denote the common value by  $M(A)$  and call it the mean of  $A$  with respect to the family  $\{\nu_a\}$ . Note that (1.2) ensures that the different expressions for  $\bar{M}$  and  $\underline{M}$  are consistent. Notice also that the function  $\nu_{\cdot}(A)$  is positive in the interval  $]-\infty, M(A)[$ , zero in the interval  $]\underline{M}(A), \bar{M}(A)[$  and negative in  $]\bar{M}(A), \infty[$ . If further  $\{\nu_a\}$  is decreasing at zero with respect to  $\mu$  then for any set  $A$ , for which  $\mu(A) > 0$ , we have that

$$\underline{M}(A) = \bar{M}(A) ,$$

and hence the function  $\nu_{\cdot}(A)$  has at most one point of sign change.

We remark that if  $\{\nu_a\}$  is decreasing at zero with respect to  $\mu$  then  $M(\cdot)$  has the Cauchy mean value property: if  $A = \bigcup_{i=1}^n A_i$ ,  $A_i A_j = \emptyset$  for  $i \neq j$ , and if  $\underline{y} = \bigwedge_{i=1}^n M(A_i)$ ,  $\bar{y} = \bigvee_{i=1}^n M(A_i)$  then  $\underline{y} \leq M(A) < \bar{y}$ . For  $a < \underline{y}$  implies  $\nu_a(A) = \sum_{i=1}^n \nu_a(A_i) > 0$ , whence  $M(A) > a$  and  $M(A) \geq \underline{y}$ ; the proof of the other inequality is similar. We remark further that if in addition  $\nu_{(\cdot)}$  is continuous, if  $\mu(A_i) < 0$ ,  $i = 1, 2, \dots, n$ , and if  $\underline{y} < \bar{y}$  then  $\underline{y} < M(A) < \bar{y}$ .

The following results give the representation of a LRN function in terms of  $\bar{M}$  and  $\underline{M}$ ; see also [3], [12], [10], [13] and [18].

**THEOREM 1.12.** *Let  $\mu$  be finite and dominate the family  $\{\nu_a, a \in R\}$ . Then the function  $f$  defined by*

$$f = \begin{cases} \text{ess sup}_{B \subset A_{\alpha}^+} (\inf_C \bar{M}(BC)IB - \infty IB^c)IA_{\alpha}^+ \\ + \text{ess inf}_{C \subset A_{\alpha}^-} (\sup_B \underline{M}(BC)IC + \infty IC^c)IA_{\alpha}^- \end{cases}$$

is  $\mu$  almost surely equal to a LRN function of  $\{\nu_a, a \in R\}$  and  $\mathcal{M}$ . Here  $B \in \mathcal{M}$  and  $C \in \mathcal{M}^c$ , and  $A_{\alpha}^+$  and  $A_{\alpha}^-$  are chosen as in Lemma 1.3.

*Proof.* There exists a countable family  $\mathcal{B}$  of sets in  $\mathcal{M}$  such that  $f = g$   $\mu$  almost surely, where

$$g = \begin{cases} \sup_{\substack{B \subset A_{\alpha}^+ \\ B \in \mathcal{B}}} \{ \inf_{C \in \mathcal{M}^c} \bar{M}(BC)IB - \infty IB^c \} IA_{\alpha}^+ \\ + \inf_{\substack{C \subset A_{\alpha}^- \\ C \in \mathcal{M}^c}} \{ \sup_{B \in \mathcal{M}} \underline{M}(BC)IC + \infty IC^c \} IA_{\alpha}^- . \end{cases}$$

We can choose  $\mathcal{B}$  to contain the sets  $M_r^+$ ,  $r$  rational, from Lemma



1.3 and  $A_\alpha^+$ . Since for  $-\infty < \alpha < \infty$   $\nu_\alpha(A_\alpha^+ \cap C) \geq 0$ , we get that  $\bar{M}(A_\alpha^+ \cap C) \geq \alpha$  and hence  $\inf_{C \in \mathcal{C}^c} \bar{M}(A_\alpha^+ \cap C) \geq \alpha$ . Thus the first term of  $g$  is always  $\geq \alpha$ . Similarly it is seen that the second is  $\leq \alpha$ . The same conclusion holds for  $\alpha = +\infty$  and  $\alpha = -\infty$ .

For  $t \geq \alpha$

$$[g > t] = \cup B$$

where the union is over  $B \in \mathcal{B}$  such that  $B \subset A_\alpha^+$  and such that  $\inf_{C \in \mathcal{C}^c} \bar{M}(BC) > t$ . This last condition is equivalent to saying that  $B$  is  $\nu_r$  positive for some  $r > t$ , hence for  $\nu_t$ , which means that  $[g > t]$  is  $\nu_t$  positive.

For  $t < \alpha$

$$[g > t] = \bigcup_{\alpha > r > t} (\cup C)^c$$

where the second union is taken over those  $C$  for which  $C^c \in \mathcal{B}$ ,  $C \subset A_\alpha^-$ , and such that  $b(C) = \sup_{B \in \mathcal{B}} \bar{M}(BC) < r$ . The last condition means that  $C$  is  $\nu_s$  negative for some  $s < r$ , hence if  $u > r$ , we have

$$\bigcup_{b(C) < r} C \subset \bigcup_{C: \nu_s \text{ negative}} C \subset \bigcup_{b(C) < u} C$$

which imply that

$$[g > t] = \bigcup_{\alpha > r > t} (\cup C)^c,$$

where the union  $\cup C$  is taken over those  $C$  for which  $C^c \in \mathcal{B}$  and  $C$  is  $\nu_r$  negative. This union is now negative for  $\nu_r$ , and since it contains a minimal element  $M_r^-$  and since  $r < \alpha$ , its complement is positive for  $\nu_r$ , and hence  $[g > t]$  is  $\nu_t$  positive. Similarly it is seen that  $[g < t]$  is  $\nu_t$  negative which proves that  $g$  is a LRN function of  $\{\nu_\alpha\}$  and  $\mathcal{M}$ .

**THEOREM 1.13.** *Let  $\{\nu_\alpha, \alpha \in R\}$  be given, and let  $M_r^+, r$  rational, be chosen as in Lemma 1.3 then the function*

$$f = \begin{cases} \sup_{r > \alpha} \{ \inf_{C \in \mathcal{C}^c} \bar{M}(M_r^+ C) IM_r^+ + \alpha IM_r^- \} IA_\alpha^+ \\ + \inf_{r < \alpha} \{ \sup_{B \in \mathcal{B}} \underline{M}(M_r^- B) IM_r^- + \alpha IM_r^+ \} IA_\alpha^- \end{cases}$$

is a LRN function of  $\{\nu_\alpha, \alpha \in R\}$  and  $\mathcal{M}$ .

*Proof.* The same as for Theorem 1.12.

**THEOREM 1.14.** *Let  $\mathcal{M}$  be generated by sets of the form  $[g > a]$  for some  $g: \Omega \rightarrow R$ . Then the function defined by*

$$f = \begin{cases} \sup_{B \subset A_\alpha^+, B \in \mathcal{M}} (\inf_{C \in \mathcal{M}^c} \bar{M}(BC)IB - \infty IB^c)IA_\alpha^+ \\ + \inf_{C \subset A_\alpha^-, C \in \mathcal{M}^c} (\sup_{B \in \mathcal{M}} \underline{M}(BC)IC + \infty IC^c)IA_\alpha^- \end{cases}$$

is a LRN function of  $\{\nu_a, a \in R\}$  and  $\mathcal{M}$ .

*Proof.* For  $t \geq \alpha$  we have

$$[f > t] = \cup B,$$

where the union is over those  $B$  for which  $B \in \mathcal{M}, B \subset A_\alpha^+$ , and  $\inf_{C \in \mathcal{M}^c} \bar{M}(BC) > t$ . By the special property of  $\mathcal{M}$  which we have assumed we can select a countable number of sets  $B$  satisfying the restrictions above, but having the same union as before. Now the argument proceeds as in the proof of Theorem 1.12.

To get the final representation of the LRN function we consider the following important example.

**EXAMPLE 1.15.** Let  $\varphi$  be a measure and let  $\mu$  be a nonnegative measure on a measurable space  $(\Omega, \mathcal{F})$ . Let either  $\varphi$  or  $\mu$  be finite. Define  $\nu_a = \varphi - a\mu, a \in R$ . Then  $\{\nu_a, a \in R\}$  satisfy (1.1), (1.2), and (1.3) with  $\alpha = 0$  if  $\varphi$  is finite,  $\alpha = +\infty$  or  $\alpha = -\infty$  if  $\mu$  is finite, but not  $\varphi$ . The associated LRN function  $f$  is a generalized LRN derivative of  $\varphi$  with respect to  $\mu$  on  $\mathcal{M}$  as discussed for finite  $\varphi$  by Johansen [15] and defined in Definition 1.16.

**DEFINITION 1.16.** If  $\varphi$  is a measure on  $\mathcal{F}$ , and if  $\mu$  is a nonnegative measure on  $\mathcal{F}$ , then a derivative  $g$  of  $\varphi$  with respect to  $\mu$  is an extended real valued function on  $\Omega$  satisfying

(1.7)  $g$  is measurable  $\mathcal{M}$ ,

(1.8)  $\varphi(B \cap [g < b]) \leq b\mu(B \cap [g < b]), b \in R, B \in \mathcal{M}$ ,

(1.9)  $\varphi(C \cap [g > a]) \geq a\mu(C \cap [g > a]), a \in R, C \in \mathcal{M}^c$ .

The following theorem is contained in [15] for the case in which  $\varphi$  is finite, but also follows from Theorem 1.5 and Theorem 1.7.

**THEOREM 1.17.** *If either  $\varphi$  or  $\mu$  is finite, then  $\varphi$  has a derivative  $g$  with respect to  $\mu$  on  $\mathcal{M}$ . If  $\mu$  is defined on  $\mathcal{S}$ , then any two derivatives coincide a.e.  $[\mu]$ . If  $\varphi$  is finite,  $g$  is finite a.e.  $[\mu]$ .*

*Proof.* The existence follows from Theorem 1.5, the uniqueness from Theorem 1.7, and the finiteness from the relation

$$\varphi([f > a]) \geq a\mu([f > a]) .$$

For  $a \uparrow \infty$  we get that  $\mu([f > a]) \downarrow 0$ , and hence  $\mu([f = \infty]) = 0$ . Similarly it is seen that  $\mu([f = -\infty]) = 0$ .

The following theorem is a direct application of Theorem 1.9, but is also contained in [15].

**THEOREM 1.18.** *Let  $\varphi$  be a finite measure on  $\mathcal{F}$ ,  $\mu$  a nonnegative finite measure on  $\mathcal{F}$ . Then  $g$  is a derivative of  $\varphi$  if and only if*

$$(1.10) \quad g \text{ is measurable } \mathcal{M},$$

$$(1.11) \quad \varphi(B \cap [g \leq b]) \leq b\mu(B \cap [g \leq b]), \quad b \in R, B \in \mathcal{M},$$

$$(1.12) \quad \varphi(C \cap [g \geq a]) \geq a\mu(C \cap [g \geq a]), \quad a \in R, C \in \mathcal{M}^c.$$

We shall now give a relation between the special case considered in Example 1.15 and the general case considered in this section. Indeed, the results of Theorems 1.17 and 1.18 are sufficient to provide a LRN function  $f$  for a family of finite measures  $\{\nu_a\}$ . In effect  $f$  is the solution to the equation  $g_a(\omega) = 0$ , where  $g_a$  is the derivative of  $\nu_a$  with respect to some finite  $\mu$  on  $\mathcal{M}$ .

**THEOREM 1.19.** *Let  $\nu_a$  be finite for all  $a \in R$ . Let  $\mu$  be finite and nonnegative on  $\mathcal{F}$ . Let  $g_a$  be a derivative of  $\nu_a$  with respect to  $\mu$  on  $\mathcal{M}$ . Then the functions*

$$(1.13) \quad \bar{f}_1(\omega) = \sup \{r: g_r(\omega) \geq 0\},$$

$$(1.14) \quad \bar{f}_2(\omega) = \inf \{r: g_r(\omega) < 0\},$$

$$(1.15) \quad \underline{f}_1(\omega) = \inf \{r: g_r(\omega) \leq 0\},$$

$$(1.16) \quad \underline{f}_2(\omega) = \sup \{r: g_r(\omega) > 0\},$$

are all LRN functions of  $\{\nu_a\}$  on  $\mathcal{M}$ .

*Proof.* Let  $A_r = \{g_r \geq 0\}$ . Then

$$\bar{f}_1 = \sup_r (rIA_r - \infty IA_r^c) .$$

If  $A_r$  is positive and  $A_r^c$  is negative for  $\nu_r$ , then  $A_r$  is maximal and  $A_r^c$  is minimal, and the proof of Corollary 1.10 carries over without change, taking  $\alpha = -\infty$ . But the content of Theorem 1.18 is precisely that  $A_r$  is positive while  $A_r^c = [g_r < 0]$  is negative by definition of  $g_r$ . Thus  $\bar{f}_1$  is a LRN function. The others are treated similarly.

We shall now illustrate the results of this section by some examples.

**EXAMPLE 1.20.** Let  $(\Omega, \mathcal{A}, \mu)$  be a probability space. Let  $X$  be

a random variable such that  $\int |X| d\mu < \infty$ . If we define

$$\nu_a(A) = \int_A (X - a) d\mu = \int_A X d\mu - a\mu(A),$$

then  $\{\nu_a, a \in R\}$  satisfies (1.1), (1.2), and (1.3). The LRN function associated with this family is nothing but the derivative of the set function  $\varphi(A) = \int_A X d\mu$  with respect to  $\mu$  on  $\mathcal{M}$  (see Definition 1.16). Hence  $f = E(X || \mathcal{M})$ . Notice that if  $\mu(A) > 0$  then  $\bar{M}(A) = \underline{M}(A) = \int_A X d\mu / \mu(A)$ , the mean of the distribution of  $X$  given  $A$ . In this case we may take  $\alpha = \infty, A_\alpha^+ = \Omega$ . Theorem (1.12) gives

$$\begin{aligned} d\varphi \backslash d\mu &= E(X || \mathcal{M}) \\ &= \text{ess sup}_{B \in \mathcal{M}} [\{ \inf_{C \in \mathcal{M}^c} M(BC) \} IB - \infty IB^c] \end{aligned}$$

where  $M(A) = \int_A X d\mu / \mu(A)$  if  $\mu(A) > 0, M(A) = \infty$  if  $\mu(A) = 0$ . In the classical case in which  $\mathcal{S}$  is a sub- $\sigma$ -field of  $\mathcal{A}$ , we have

$$\begin{aligned} d\varphi \backslash d\mu &= E(X || \mathcal{S}) \\ &= \text{ess sup}_{B \in \mathcal{S}} [\{ \inf_{\substack{A \subset B \\ A \in \mathcal{S}}} M(A) \} IB - \infty IB^c]. \end{aligned}$$

Since  $B_1 \supset B_2$  implies  $\inf_{A \subset B_1} M(A) \leq \inf_{A \subset B_2} M(A)$ , this representation may be interpreted as giving the derivative  $d\varphi \backslash d\mu$  as the "limit" of averages of  $X$  over  $A$  as  $A \downarrow \{\omega\}$ .

Instead of  $\mathcal{M}$  we sometimes have a partial order  $\succeq$  defined on  $\Omega$ , i.e., (1)  $\omega \succeq \omega$ , (2)  $\omega_1 \succeq \omega_2$  and  $\omega_2 \succeq \omega_3 \Rightarrow \omega_1 \succeq \omega_3$ , (3)  $\omega_1 \succeq \omega_2$  and  $\omega_2 \succeq \omega_1 \Rightarrow \omega_1 = \omega_2$ . We are interested in functions  $f$  which are isotone, that is

$$\omega_1 \succeq \omega_2 \Rightarrow f(\omega_1) \geq f(\omega_2).$$

In this case we call  $B$  an upper set if  $\omega_1 \in B, \omega_2 \succeq \omega_1 \Rightarrow \omega_2 \in B$ , or if  $IB$  is isotone. If we let  $\tilde{\mathcal{M}}$  be the class of upper sets, then  $\tilde{\mathcal{M}}$  is a  $\sigma$ -lattice, and  $f$  is measurable  $\tilde{\mathcal{M}}$  if and only if  $f$  is isotone; see Brunk [10].

Let us consider two special cases of Example 1.20.

(a)  $\Omega = [0, 1], \mathcal{A} =$  Borel sets,  $\mu$  Lebesgue measure. Let  $\tilde{\mathcal{M}}$  be generated by the intervals  $[a, 1], a \in [0, 1]$ , i.e.,  $\tilde{\mathcal{M}}$  consists of the upper sets generated by the usual order on  $[0, 1]$ . Then  $f$  is measurable  $\tilde{\mathcal{M}}$  if  $f$  is nondecreasing.

In this case we can apply Theorem 1.14, and we get that

$$f(x) = \sup_b \left( \inf_c \frac{\int_b^c X(\omega) d\omega}{c - b} \cdot I_x[b, 1] - \infty I_x[0, b] \right)$$

or

$$f(x) = \sup_{b < x} \inf_{b < c} \frac{\int_b^c X(\omega) d\omega}{c - b}$$

is a derivative. Similarly

$$g(x) = \inf_{x < c} \sup_{b < c} \frac{\int_b^c X(\omega) d\omega}{c - b}$$

is a representation of the conditional mean value of  $X$  given the  $\sigma$ -lattice  $\mathcal{M}$ .

(b)  $\Omega = \{x_1, \dots, x_k\}$ ,  $\mathcal{A} =$  all subsets of  $\Omega$ ,  $\mu$  is normalized counting measure.  $\mathcal{M}$  is the class of sets  $\{x_i, \dots, x_k\}$ ,  $i = 1, \dots, k$ , together with the empty set. Then the representation in Theorem 1.14 gives

$$f(x_k) = \sup_{i < k} \inf_{i < j} \frac{\sum_{t=i+1}^j X(x_t)}{j - i}.$$

It is interesting to see that in the last two examples, if we let  $\tilde{X}(x) = \int_0^x X(t) dt$  (respectively  $\sum_{i=1}^j X(x_i)$ ), then  $f$  is a derivative of the largest concave minorant of  $\tilde{X}$ , a result found independently by Reid [9] and by Grenander [13] (see also Brunk [7]).

As a more general example we shall consider

EXAMPLE 1.21. Let  $(\Omega, \mathcal{A}, \mu)$  be a probability space and  $X$  random variable. Let  $\psi$  be an extended real valued function on  $\Omega \times R$  such that  $\psi(\omega, a)$  is  $\mathcal{A}$ -measurable for  $a \in R$ . Assume that for some  $\alpha$ ,  $-\infty \leq \alpha \leq \infty$ ,  $\psi(\cdot, a)^+ = \psi(\cdot, a) \vee 0$  is integrable,  $a < \alpha$ , and  $\psi(\cdot, a)^- = -\psi(\cdot, a) \wedge 0$  is integrable,  $a \geq \alpha$ . Define

$$\nu_a(A) = \int_A \psi(\omega, a) \mu(d\omega), \quad a \in R, A \in \mathcal{F}.$$

Then  $\{\nu_a, a \in R\}$  satisfies (1.1) and (1.3), and we further assume that it satisfies (1.2). This will in particular be the case if  $\psi(\omega, a)$  is non-decreasing for  $\omega \in \Omega$ . It is seen that the previous example is contained in this if we put  $\psi(\omega, a) = X(\omega) - a$ .

Brøns [5], Brøns, Brunk, Franck and Hanson [6] and Cashwell and Everett [11] have studied special cases in which  $\mathcal{M} = \{\emptyset, \Omega\}$ . For

the case of Cashwell and Everett, we set

$$(1.17) \quad \psi(\omega, a) = \int_a^{X(\omega)} w(\xi, \omega) d\xi ,$$

where  $w$  is positive. For the case studied by Brøns, Brunk, Franck and Hanson we set

$$(1.18) \quad \psi(\omega, a) = -\varphi[X(\omega), a] ,$$

where  $\varphi(x, a)$  is an extended real valued function on  $R \times R$ , which is Borel measurable in  $X$ , nondecreasing in  $a$ , nonnegative for  $a > x$ , and nonpositive for  $a < x$ . For such  $\{\nu_a\}$ , we term the LRN function a conditional generalized mean of  $X$  given  $\mathcal{M}$ . The content of Theorem 1.19 in these cases is, loosely speaking, that the conditional generalized mean of  $X$  given  $\mathcal{M}$  can be obtained by solving, for each  $\omega \in \Omega$ , the equation

$$E(\psi(\cdot, a) | \mathcal{M})(\omega) = 0$$

for  $a$ . If in particular  $\mathcal{M} = \{\emptyset, \Omega\}$  then the conditional generalized mean is constant; in the case of (1.17) it is called in [Cashwell and Everett] a mean of  $X$  relative to the weight function  $w$ ; in the case of (1.18) it is called in [Brøns, Brunk, Franck and Hanson] a  $\varphi$ -mean of  $X$ . (We note that a  $\varphi$ -mean is determined by the function  $\varphi$  and the probability distribution of  $X$ ; whereas for a given weight function  $w$  two random variables on  $(\Omega, \mathcal{A}, \mu)$  may have the same probability distribution but different means relative to  $w$ .)

We now consider the special  $\varphi$ -function defined by

$$\varphi(x, a) = \text{sgn}(a - x) = \begin{cases} 1 & x \leq a , \\ -1 & x > a . \end{cases}$$

Now

$$\nu_a(A) = \mu(A) - 2\mu(A \cap [X \leq a]) , \quad A \in \mathcal{F} .$$

If  $\mathcal{M} = \{\emptyset, \Omega\}$  we want the point of sign change for

$$\nu_a(\Omega) = 1 - 2\mu([X \leq a]) ,$$

which means that the  $\varphi$ -mean of  $X$  is any median in the distribution of  $X$ . It is seen that the condition that  $\{\nu_a\}$  is decreasing at zero with respect to  $\mu$ , in this case just ensures that the median in the distribution of  $X$  is unique. The conditional generalized mean is in this case called the conditional median of  $X$ .

Let  $\Omega = [0, 1]$ ,  $\mu =$  Lebesgue measure on  $\mathcal{A}$  the Borel sets, and let  $\mathcal{M}$  be generated by the sets  $[a, 1]$ ,  $a \in [0, 1]$ . Let  $X$  be a random

variable, and assume  $\mu(A \cap [X \leq a])$  strictly increasing on the range of  $X$  for all  $A \in \mathcal{F}$  such that  $\mu(A) > 0$ . Then

$$\nu_a(A) = \int_A -\operatorname{sgn}(a - X)d\mu$$

is decreasing at zero with respect to  $\mu$ , and the equation

$$\nu_a(A) = 0$$

gives for  $\mu(A) > 0$  that

$$\mu([X \leq a] | A) = \frac{1}{2},$$

i.e.,  $\underline{M}(A) = \bar{M}(A) = M(A)$ , where  $M(A)$  is the median of the distribution of  $X$  given  $A$ . By means of this we get from Theorem 1.10 that the conditional median of  $X$  given  $\mathcal{M}$  is

$$f(x) = \sup_{b < x} \inf_{b < c} M([b, c]),$$

where  $\mathcal{M}(b, c]$  denotes the median of the distribution of  $X$  given the interval  $[b, c]$ .

Another way of finding it is by solving the equation

$$E[\operatorname{sgn}(X - a) | \mathcal{M}] = 0.$$

**2. A characterization of the LRN function.** The martingale theorem. If we consider the Radon-Nikodym derivative  $f$  of a finite  $\sigma$ -additive set function  $\varphi$  with respect to a  $\sigma$ -finite measure  $\mu$  on a  $\sigma$ -field (see Theorem 1.17) then we can characterize it by the fact that its indefinite integral gives the continuous part of  $\varphi$ , and the singular part is concentrated on the set where  $|f| = \infty$  (see Hewitt and Stromberg [14]).

This was discussed for  $\sigma$ -lattices in Johansen [15], where certain inequalities between  $\varphi$  and the indefinite integral of  $f$  were proved. If further  $\mu$  was finite,  $f$  could be characterized by these inequalities. The purpose of this section is to extend the same ideas to a special case of the situation considered in §1, and to prove a martingale convergence theorem for  $\sigma$ -lattices.

Let us therefore consider a family  $\{\nu_a, a \in R\}$  defined by

$$(2.1) \quad \nu_a(A) = \int_A \psi(\omega, a)\mu(d\omega), \quad A \in \mathcal{F}$$

where  $\psi$  maps  $\Omega \times R$  into  $R$  such that  $\psi(\omega, \cdot)$  is nonincreasing and continuous, and  $\psi(\cdot, a)$  is integrable with respect to  $\mu$  where  $\mu$  is a finite measure. This means that  $\nu_a$  is finite, and if we choose  $\psi(\omega, a) = X(\omega) - a$ , we get the situation considered in Johansen [15] for a finite  $\mu$ .

The family can be extended in a natural way by defining

$$\psi(\omega, \pm\infty) = \lim_{a \rightarrow \pm\infty} \psi(\omega, a)$$

and

$$\nu_{\pm\infty}(A) = \lim_{a \rightarrow \pm\infty} \nu_a(A).$$

By the monotone convergence theorem we get that

$$\nu_{\pm\infty}(A) = \int_A \psi(\omega, \pm\infty) \mu(d\omega).$$

LEMMA 2.1. *The function*

$$\psi: \Omega \times [-\infty, \infty] \longrightarrow [-\infty, \infty]$$

*is jointly measurable with respect to the  $\sigma$ -field generated by  $\mathcal{F} \times \mathcal{B}$ .*

Proof omitted.

DEFINITION 2.2. Let  $f$  be an extended real valued  $\mathcal{N}$  measurable function. We define

$$\nu_f(A) = \int_A \psi(\omega, f(\omega)) \mu(d\omega), \quad A \in \mathcal{F},$$

whenever the righthand side is defined. We call  $f$  *integrable on  $A$* , if  $\nu_f(A)$  exists and is finite, and  $f$  is *integrable* if  $\nu_f(\Omega)$  exists and is finite. Notice that  $\nu_f(A \cap [f < a])$  is well defined even if it may be infinite, while  $\nu_f(A \cap [a < f < b])$  is well defined and finite.

In the case where  $\psi(\omega, a) = X(\omega) - a$  then

$$\nu_f(A) = \int_A (X(\omega) - f(\omega)) \mu(d\omega),$$

and in this case  $\nu(A)$  becomes an affine functional on the class of integrable functions.

In general, however, this is not the case.

LEMMA 2.3. *The class  $L$  of integrable functions is a lattice which contains the constants.*

*Proof.* The lemma follows from the inequality

$$|\psi(\omega, (f \wedge g)(\omega))| \vee |\psi(\omega, (f \vee g)(\omega))| \leq |\psi(\omega, f(\omega))| + |\psi(\omega, g(\omega))|.$$

LEMMA 2.4. *The functional  $\nu.(A)$  has the following properties:*



- (1) If  $f \equiv a$ , then  $\nu_f(A) = \nu_a(A)$ .
- (2) If  $f \leq g$  on  $A$ , and if  $\nu_f(A)$  and  $\nu_g(A)$  exist, then  $\nu_f(A) \geq \nu_g(A)$ .
- (3) If  $f \in L, g \in L$ , then  $\nu_{f \vee g}(A) + \nu_{f \wedge g}(A) = \nu_f(A) + \nu_g(A)$ .
- (4) If  $f_n \downarrow f, n \rightarrow \infty$ , and if  $\nu_{f_1}(A) > -\infty$ , then  $\nu_{f_n}(A) \uparrow \nu_f(A)$ ,  $n \rightarrow \infty$ .
- (5) If  $f_n \uparrow f, n \rightarrow \infty$ , and if  $\nu_{f_1}(A) < +\infty$ , then  $\nu_{f_n}(A) \downarrow \nu_f(A)$ ,  $n \rightarrow \infty$ .

The properties of  $\nu$ . ( $\Omega$ ) describe a  $\sigma$ -continuous isotone valuation (see Alfsen [1]).

*Proof.* The proof of (1), (2), and (3) follows directly from the definition while the proof of (4) and (5) follows from the monotone convergence theorem together with the continuity of  $\nu(\omega, \cdot)$ .

**THEOREM 2.5.** *Let  $f$  denote a LRN function associated with the family  $\{\nu_a, a \in R\}$  on  $\mathcal{M}$ , then  $f$  is integrable on the set where it is finite, and*

- (1)  $\nu_f(B \cap [-\infty < f < \infty]) \leq 0, B \in \mathcal{M}$ .
- (2)  $\nu_f(C \cap [-\infty < f < \infty]) \geq 0, C \in \mathcal{M}^c$ .
- (3)  $\nu_f([a \leq f < b]) = 0, a \in R, b \in R$ .
- (4)  $\nu_f(B \cap [f = -\infty]) \leq 0, B \in \mathcal{M}$ .
- (5)  $\nu_f(C \cap [f = +\infty]) \geq 0, C \in \mathcal{M}^c$ .

If, conversely,  $f$  is  $\mathcal{M}$  measurable and integrable on the set where it is finite and satisfies (1) through (5), then  $f$  is a LRN function associated with  $\{\nu_a, a \in R\}$  on  $\mathcal{M}$ .

Let us first remark that if further

$$\nu_\infty \leq 0$$

and

$$\nu_{-\infty} \geq 0,$$

then  $\nu_f$  vanishes at subsets of  $[|f| = \infty]$ , and  $f$  becomes integrable ( $\nu_f(\Omega) = 0$ ). Before proceeding to the proof, let us again consider the situation where

$$\nu_a(A) = \int_A (X(\omega) - a)\mu(d\omega),$$

and let us assume that  $\mathcal{M}$  is a  $\sigma$ -field. Now  $\nu_\infty \leq 0$  and  $\nu_{-\infty} \geq 0$ , and hence  $f$  is integrable, and (1) and (2) reduces to

$$\nu_f(A) = 0, \quad A \in \mathcal{M},$$

or

$$\int_A X d\mu = \int_A f d\mu$$

which is the usual relation between the set function  $\int_A X d\mu$  and its derivative  $f$ .

*Proof of Theorem 2.5.* We split up the set  $A = [a \leq f < b]$  as follows:

$$A = \bigcup_{k=1}^n A_k,$$

where  $A_k = [a_k \leq f < a_{k+1}]$ ,  $a_{k+1} - a_k = (b - a)n^{-1}$ ,  $k = 1, 2, \dots, n$ ,  $a_1 = a$ ,  $a_{n+1} = b$ . We want to prove that

$$\nu_f(A \cap B) \leq 0, \quad B \in \mathcal{M}.$$

On the set  $A_k \cap B$  we know that  $f \geq a_k$  or  $(b - a)n^{-1} + f \geq a_{k+1}$ , hence by (2) of Lemma 2.4

$$\nu_{(b-a)n^{-1}+f}(A_k \cap B) \leq \nu_{a_{k+1}}(B \cap [a_k \leq f < a_{k+1}]) \leq 0.$$

But  $(b - a)n^{-1} + f \leq 2b - a$  on  $A \cap B$ , and hence

$$-\infty < \nu_{2b-a}(A \cap B) \leq \nu_{f+(b-a)n^{-1}}(A \cap B) \leq 0.$$

For  $n \rightarrow \infty$  we get by (4) of Lemma 2.4 that

$$\nu_f(A \cap B) = \nu_f([a \leq f < b] \cap B) \leq 0, \quad B \in \mathcal{M}.$$

Similarly it is seen that

$$\nu_f([a \leq f < b] \cap C) \geq 0, \quad C \in \mathcal{M}^c.$$

If we choose  $B = C = \Omega$  we get (3). We now want to prove integrability of  $f$  on the set  $[-\infty < f < \infty]$ , and we therefore want to let  $b \rightarrow \infty$ ,  $a \rightarrow -\infty$ .

The function

$$\psi(\cdot, a) - \psi(\cdot, f(\cdot))$$

is nonnegative on  $[f \geq a]$ , hence  $\nu_a - \nu_f$  is well defined and nonnegative on  $[f \geq a]$ .

Therefore

$$0 \leq (\nu_a - \nu_f)(B \cap [a \leq f < b]) \uparrow (\nu_a - \nu_f)(B \cap [a \leq f < \infty])$$

for  $b \rightarrow \infty$ . Since  $\nu_a$  is finite we have proved that  $\nu_f(B \cap [a \leq f < \infty])$  is well defined and that

$$\nu_f(B \cap [a \leq f < b]) \rightarrow \nu_f(B \cap [a \leq f < \infty]), \quad b \rightarrow \infty .$$

In particular for  $B = \Omega$  the left hand side is zero and hence

$$\nu_f([a \leq f < \infty]) = 0 .$$

Then clearly  $\nu_f(B \cap [a \leq f < \infty])$  is finite and  $\leq 0$ . This proves the integrability on the set  $f \geq a$ . The integrability on  $f < a$  is proved similarly. Combining these inequalities we obtained (1). The inequality (2) is proved similarly.

From the inequality

$$\nu_b(C \cap [f > a]) \leq \nu_a(C \cap [f > a]) \leq 0, \quad 0 < b < a, \quad C \in \mathcal{M}^c$$

we obtain by letting  $a \uparrow \infty$ , that

$$\nu_b(C \cap [f = +\infty]) \geq 0 ,$$

and hence for  $b \uparrow \infty$ , we get

$$\nu_f(C \cap [f = +\infty]) = \nu_\infty(C \cap [f = \infty]) \geq 0 .$$

Hence (5) is proved and (4) is proved analogously.

To prove the last statement of Theorem 2.5, let us assume that  $f$  is  $\mathcal{M}$  measurable and integrable on the set where it is finite, and that it satisfies (1) through (5). We have to verify (1.5) and (1.6). To prove (1.5) we consider

$$\begin{aligned} \nu_b(B \cap [f < b]) &\geq \nu_f(B \cap [f < b]) \\ &= \nu_f(B \cap [f = -\infty]) + \nu_f(B \cap [-\infty < f < b]) \\ &= \nu_f(B \cap [f = -\infty]) + \nu_f([-\infty < f < b]) - \nu_f(B^c \cap [-\infty < f < b]) . \end{aligned}$$

If we now apply (4), (3), and (2) we get that this sum is  $\leq 0$ , which proves (1.5). (1.6) is proved in the same way.

In proving Theorem 2.5 essential use has been made of the assumptions that the kernel  $\psi(\cdot, \cdot)$  is continuous and nonincreasing in the second (real) argument. We remark that Theorem 2.5 can be applied to situations in which the kernel is neither continuous nor nonincreasing. As an example of this let us consider a fixed positive Borel function  $h$ , and the family defined by

$$\lambda_a(A) = h(a)\nu_a(A) .$$

It is seen that  $\{\lambda_a, a \in R\}$  satisfies (1.1), (1.2), and (1.3), and that if  $f$  is LRN for  $\{\nu_a, \mathcal{M}\}$ , then  $f$  is LRN for  $\{\lambda_a, \mathcal{M}\}$ . The results such as

$$\lambda_f[a \leq f < b] = \int_{[a \leq f < b]} h(f(\omega))\nu_f(d\omega) = 0$$

now follow from Theorem 2.5 since the measure  $\nu_f(\cdot)$  vanishes on the  $\sigma$ -field induced by  $f$  on the set  $[-\infty < f < \infty]$ .

We shall now apply Theorem 2.5 to discuss the uniqueness of the LRN function.

**DEFINITION 2.6.** Let  $f$  and  $g$  be measurable  $\mathcal{M}$ . We call  $f$  and  $g$  *equivalent* if

$$\int \left| \psi(\omega, f(\omega)) - \psi(\omega, g(\omega)) \right| \mu(d\omega) = 0$$

or if

$$(\nu_{f \vee g} - \nu_{f \wedge g})(\Omega) = 0 .$$

Notice that in the case where

$$\psi(\omega, a) = X(\omega) - a ,$$

this notion becomes the usual one of equality almost surely  $[\mu]$ .

The following trivial example is also covered by the formulation in this section. Let  $\mu$  be a probability measure, and let  $\mathcal{M} = \{\emptyset, \Omega\}$ . Instead of choosing

$$\psi(\omega, a) = \text{sgn}(X(\omega) - a) ,$$

which is not continuous in  $a$ , we define

$$\nu_a(\Omega) = \psi(\omega, a) = 1 - 2\mu([X \leq a]) .$$

Then  $\nu_a$  has the representation (2.1) for sets in  $\mathcal{M} = \mathcal{F}$ , and  $(\psi(\omega, \cdot))$  is continuous if we assume that the distribution function of  $X$  is continuous. Two constants  $a$  and  $b$  are now equivalent if the interval  $[a, b]$  is a null set for the distribution of  $X$ . In particular, the Theorem 2.7 below tells that any two medians are equivalent in this case.

**THEOREM 2.7.** *If  $\{\nu_a, a \in R\}$  is defined by (2.1), and if  $f$  and  $g$  are two LRN functions associated with the family and  $\mathcal{M}$ , then  $f$  and  $g$  are equivalent.*

*Proof.* We want to prove that

$$(\nu_{f \vee g} - \nu_{f \wedge g})(\Omega) = 0$$

or equivalently

$$(\nu_f - \nu_g)[f < g] = 0 = (\nu_g - \nu_f)[g < f] .$$

It is enough to prove the former; the latter follows on interchanging  $f$  and  $g$ . It clearly suffices to show that

$$(\nu_f - \nu_g)[f < a < b < g] = 0$$

for real  $a$  and  $b$ .

As above we prove that

$$\nu_f(B \cap [f < a]) \leq 0$$

and similarly

$$\nu_g(C \cap [g > b]) \geq 0,$$

therefore

$$0 \geq \nu_f[f < a < b < g] \geq \nu_g[f < a < b < g] \geq 0$$

which was to be proved.

We now want to prove a martingale convergence theorem for  $\sigma$ -lattices. We let  $\mathcal{M}_n$  be an increasing sequence of  $\sigma$ -lattices, generating the  $\sigma$ -lattice  $\tilde{\mathcal{M}}$ . Let  $\tilde{A}_n = \sigma(\tilde{M}_n)$ ,  $n = 1, 2, \dots$  and  $\tilde{A} = \sigma(\tilde{M})$ . Let  $\{\nu_a, a \in R\}$  denote a family of finite measures on  $\tilde{\mathcal{A}}$  which satisfy the condition (1.2). Then there exists a finite measure  $\mu$  such that  $\mu$  dominates  $\nu_a, a \in R$ . We further assume that  $\{\nu_a, a \in R\}$  is decreasing at zero with respect to  $\mu$ , and that  $\nu$ . (A) is continuous for all  $A \in \tilde{\mathcal{A}}$ .

For each  $n$  we let  $\nu_{a,n}$  denote the contraction of  $\nu_a$  to  $\tilde{\mathcal{A}}_n$ , and we let  $f_n$  denote a LRN function associated with  $\{\nu_{a,n}, a \in R\}$  and  $\tilde{\mathcal{M}}_n$ . We call the sequence  $\{f_n, n \geq 1\}$  a martingale.

**THEOREM 2.8.** *If  $\bar{f} = \limsup f_n$  and  $\underline{f} = \liminf f_n$ , then  $\bar{f}$  and  $\underline{f}$  are LRN functions of  $\{\nu_a, a \in R\}$  and  $\tilde{\mathcal{M}}$ , and  $\bar{f}$  equals  $f$   $\mu$ -almost surely.*

*Proof.* The proof is a slight modification of the proof by Andersen and Jessen [2]. We have

$$[\bar{f} \geq b] = \bigcap_{n=1}^{\infty} \bigcup_{p=n}^{\infty} [f_p > b - \epsilon_n] \in \tilde{\mathcal{M}}$$

and

$$[\underline{f} \leq b] = \bigcap_{n=1}^{\infty} \bigcup_{p=n}^{\infty} [f_p < a + \epsilon_n] \in \tilde{\mathcal{M}}^c,$$

where  $\epsilon_n$  decreases to zero. This proves that  $\underline{f}$  and  $\bar{f}$  are measurable  $\tilde{\mathcal{M}}$ . Now let  $H_n = \bigcup_{p=n}^{\infty} H_{n,p}$ , where

$$H_{n,p} = [f_n \leq b - \varepsilon_n, \dots, f_{p-1} \leq b - \varepsilon_n, f_p > b - \varepsilon_n], \quad p \geq n.$$

For  $C \in \mathcal{M}_m^c$  and  $n \leq m$ , we get that  $H_n$  is a positive set for  $\nu_{b-\varepsilon_n}$  on  $\mathcal{S}_m$ , since

$$\begin{aligned} \nu_{b-\varepsilon_n}(C \cap H_n) &= \sum_{p=n}^{\infty} \nu_{b-\varepsilon_n}(C \cap H_{n,p}) \\ &= \sum_{p=n}^{\infty} \nu_{b-\varepsilon_n,p}(C \cap H_{n,p}) \leq 0. \end{aligned}$$

If we let  $n \rightarrow \infty$ , we get by continuity that

$$\nu_b(C \cap [\bar{f} \geq b]) \geq 0, \quad C \in \mathcal{M}_m^c.$$

By the monotone class argument this can be extended to hold for  $C \in \mathcal{M}^c$ .

Similarly it is seen that

$$\nu_a(B \cap [f \leq a]) \leq 0, \quad B \in \mathcal{M}.$$

Hence if  $A = [f \leq a < b \leq \bar{f}]$ , we get  $\nu_a(A) \leq 0$  which by (1.2) implies that  $\nu_b(A) \leq 0$ . On the other hand  $\nu_b(A) \geq 0$  and by (1.2)  $\nu_a(A) \geq 0$ , thus  $\nu_a(A) = \nu_b(A) = 0$ . Since  $\{\nu_a\}$  is decreasing at zero we must have  $\mu(A) = 0$ , which proves that  $[f \neq \bar{f}]$  is a  $\mu$  null set. Since  $\mu$  dominates  $\nu_a$  it is a null set for  $\nu_a$ , and therefore

$$\nu_a(C \cap [f \geq a]) = \nu_a(C \cap [\bar{f} \geq a]) \geq 0$$

which proves that  $f$  and therefore  $\bar{f}$  are LRN functions of  $\{\nu_a, a \in R\}$  on  $\mathcal{M}$ .

3. A minimizing property. One of the useful elementary properties of expectation is that it minimizes the mean square deviation: if  $X$  is a random variable,  $E(X - \theta)^2$  is minimized in the class of real  $\theta$  by  $EX$ . This is a special instance of a minimizing property of conditional expectation given a  $\sigma$ -field: for a given random variable  $X$ ,  $E(X - g)^2$  is minimized in the class of random variables  $g$  measurable with respect to a  $\sigma$ -field  $\mathcal{S}$  by the conditional expectation  $E(X|\mathcal{S})$ . Indeed  $E(X|\mathcal{S})$  minimizes in the same class also

$$E[T(X) - T(g) - (X - g)T'(g)]$$

where  $T$  is convex, the above being the special case  $T(x) = x^2$ . (This latter property plays a role in such theorems as the Rao-Blackwell Theorem.) In the statements above, “ $\sigma$ -field” may be replaced by “ $\sigma$ -lattice”. By virtue of the last-mentioned property of conditional expectation given a  $\sigma$ -lattice, it provides solutions to many problems of maximum likelihood estimation (cf. [8]); in particular, to the problem

of maximum likelihood estimation of ordered means in sampling from distributions belonging to a common exponential family. This latter requirement is relaxed in a study by van Eeden [12] in which the distributions need only be strictly unimodal. An intermediate class is the class of distributions belonging to a generalized exponential family, as defined in [6].

The minimizing property of the LRN function which is developed in this section is sufficiently general essentially to cover all of these special instances, which are detailed in Examples 3.6 through 3.9. It may be regarded as an elaboration of a maximizing property of a maximal set, whose proof is immediate, as follows.

If  $A$  is a positive set for  $\nu$ , and if  $A^c$  is negative, then  $A$  is maximal, and in fact, for any  $B \in \mathcal{M}$ , we have

$$\nu(B) \leq \nu(B) + \nu(AB^c) = \nu(A \cup B) = \nu(A) + \nu(BA^c) \leq \nu(A) .$$

Thus  $\nu$  assumes its maximum value on  $\mathcal{M}$  at  $A$ .

The underlying idea of the following theorems is now that if  $A_u$  is positive for  $\nu_u$ , and  $A_u^c$  is negative, and if  $B_u$  is any set in  $\mathcal{M}$ , then

$$\int \nu_u(A_u)\gamma(du) \leq \int \nu_u(B_u)\gamma(du)$$

provided that both sides are defined. We shall apply this for sets  $A_u$  of the form  $[f > u]$  or  $[f \geq u]$  and for sets  $B_u$  of the form  $[g > u]$  or  $[g \geq u]$ , where  $f$  is a LRN function, and  $g$  is measurable  $\mathcal{M}$ .

Let  $\{\nu_a, a \in R\}$  satisfy (1.1), (1.2), and (1.3). Let further  $\nu_a$  be  $\sigma$ -finite, then they are all dominated by some finite  $\mu$ . For the sake of being definite let us assume that  $\nu_a < \infty$ . We now have the representation

$$(3.1) \quad \nu_a(A) = \int_A \psi(\omega, a)\mu(d\omega), \quad a \in R, A \in \mathcal{F},$$

where we further assume that the function  $\psi(\cdot, \cdot)$  is measurable with respect to the  $\sigma$ -field generated by  $\mathcal{F} \times \mathcal{B}$ .

Let  $\gamma$  be a  $\sigma$ -finite positive measure on  $\mathcal{B}$  the Borel sets of  $R$ .

$$(3.2) \quad k(\omega, a) = \begin{cases} \int_{[a, \infty[} \psi^+(\omega, u)\gamma(du) + \int_{] -\infty, a[} \psi^-(\omega, u)\gamma(du), & a \geq \alpha \\ \int_{] a, \infty[} \psi^+(\omega, u)\gamma(du) + \int_{] -\infty, a[} \psi^-(\omega, u)\gamma(du), & a < \alpha . \end{cases}$$

For an  $\mathcal{M}$ -measurable function  $g$  on  $\Omega$  define

$$(3.3) \quad I(g) = \int k(\omega, g(\omega))\mu(d\omega) .$$

**PROPOSITION 3.1.** *Let  $\{\nu_a, a \in R\}$  satisfy (1.1), (1.2), and (1.3), and let  $\nu_a$  be  $\sigma$ -finite and  $\nu_a < \infty$ . Let  $f$  be a LRN function associated with  $\{\nu_a\}$  on  $\mathcal{M}$ . If  $[f \leq u]$  is negative for  $\nu_u, u \geq \alpha$ , and  $[f \geq u]$  is positive for  $\nu_u, u < \alpha$ , then  $I(g) \geq I(f)$  for all  $g$  measurable  $\mathcal{M}$ . It suffices that these conditions be satisfied  $\gamma$ -almost surely.*

*Proof.* Let us first remark that since  $[f > u]$  is positive for  $\nu_u$ , and  $[f \leq u]$  is negative, then  $[f > u]$  is maximal, and

$$\nu_u([f > u]) \geq \nu_u([g > u]), \quad u \geq \alpha$$

similarly

$$\nu_u([f < u]) \leq \nu_u([g < u]), \quad u < \alpha.$$

We now insert  $a = g(\omega)$  in (3.2) and interchange the order of integration in (3.3). For a fixed  $u \geq \alpha$ , we shall integrate  $\psi^+$  on the set

$$[u \geq g \geq \alpha] \cup [\alpha > g] = [u \geq g],$$

and  $\psi^-$  is integrated on the set

$$[g > u] \cup \phi = [g > u].$$

Hence we get a first term of  $I(g)$ :

$$\int_{[\alpha, \infty[} \gamma(du) \left\{ \int_{[g \leq u]} \psi^+(\omega, u) \mu(d\omega) + \int_{[g > u]} \psi^-(\omega, u) \mu(d\omega) \right\}.$$

By considering a fixed  $u < \alpha$ , we get a second term

$$\int_{]-\infty, \alpha[} \gamma(du) \left\{ \int_{[g < u]} \psi^+(\omega, u) \mu(d\omega) + \int_{[g \geq u]} \psi^-(\omega, u) \mu(d\omega) \right\}.$$

Hence

$$\begin{aligned} I(g) &= \int_{[\alpha, \infty[} \gamma(du) \{ \nu_u^+([g \leq u]) + \nu_u^-([g > u]) \} \\ &\quad + \int_{]-\infty, \alpha[} \gamma(du) \{ \nu_u^+([g < u]) + \nu_u^-([g \geq u]) \}. \end{aligned}$$

Here and in the following we apply the notation

$$\nu_u^+(A) = \int_A \psi^+(\omega, u) \mu(d\omega), \quad \nu_u^-(A) = \int_A \psi^-(\omega, u) \mu(d\omega).$$

Now for  $u \geq \alpha, \nu_u^+ < \infty$ , and

$$\nu_u^+([g \leq u]) + \nu_u^-([g > u]) = \nu_u^+(\Omega) - \nu_u^-([g > u]).$$

By assumption and the remark at the beginning of this section we get that this is



$$\geq \nu_u^+(\Omega) - \nu_u(f > u) .$$

Treating the second term similarly, and reversing the calculations with  $f$  replacing  $g$ , we get that

$$I(g) \geq I(f) .$$

It is clear that the conclusion holds if the conditions are satisfied  $\gamma$ -almost surely.

We now consider conditions under which the hypothesis of Proposition 3.1 holds.

**THEOREM 3.2.** *Let  $\{\nu_a, a \in R\}$  satisfy (1.1), (1.2), and (1.3). Let  $\nu_a$ , given by (3.1), be  $\sigma$ -finite,  $\nu_a < \infty$  and define  $D = \{a \mid \mu([f = a]) > 0\}$ . Then if  $\gamma(D) = 0$  it follows that  $I(g) \geq I(f)$  for all  $\mathcal{M}$  measurable functions  $g$ .*

*Proof.* The result follows from Proposition 3.1 and the proof of Theorem 1.8, since the set  $D$  can be chosen as indicated above.

**THEOREM 3.3.** *If  $\nu_a$  is finite, and if  $\nu \cdot (A)$  is continuous, then  $I(g) \geq I(f)$ , for all  $g$  measurable  $\mathcal{M}$ , where  $f$  is the LRN function associated with  $\{\nu_a\}$  and  $\mathcal{M}$ .*

*Proof.* Follows immediately from Theorem 1.9 and Proposition 3.1.

We emphasize that these theorems imply that  $f$  provides the solution to a rather wide class of minimum problems: given  $\psi$  and  $\mu$ , the same function minimizes  $I(g)$  in the class of  $\mathcal{M}$  measurable functions, in the first case for all  $\gamma$  such that  $\gamma(D) = 0$ , in particular if  $\gamma$  has no atoms, and in the second case for all  $\gamma$ .

There are thus two ways of getting rid of the exceptional set  $D$ , either by letting  $\gamma$  ignore it or by assuming continuity of  $\nu \cdot (A)$ . This last requirement is relaxed in Proposition 3.5.

**DEFINITION 3.4.** The family  $\{\nu_a\}$  is said to be *right continuous at  $a$* , if

$$\nu_s(A) \rightarrow \nu_a(A) , \quad s \downarrow a ,$$

whenever  $A$  is such that  $\nu_s(A) < \infty$  for  $s$  sufficiently close to  $a$ . *Left continuity* is defined similarly.

**PROPOSITION 3.5.** *Let  $\{\nu_a, a \in R\}$  satisfy (1.1), (1.2), and (1.3).*

Let  $f$  be a LRN function of  $\{\nu_u\}$  and  $\mathcal{N}$ . Let  $\{\nu_u\}$  be right continuous at  $u \geq \alpha$ , and left continuous at  $u < \alpha$ . Let  $g$  be measurable  $\mathcal{N}$ . For  $u \geq \alpha$  let  $u < b < c$  imply that  $\nu_b([g > c])$  is finite and for  $u < \alpha$  let  $c < b < u$  imply that  $\nu_b([g < c])$  is finite. Then  $I(g) \geq I(f)$ .

It suffices that the conditions on  $u$  hold  $\gamma$ -almost surely.

*Proof.* In the proof of Proposition 3.1 we used only the relation

$$\nu_u([g > u]) \leq \nu_u([f > u]), \quad u \geq \alpha$$

and

$$\nu_u([g < u]) > \nu_u([f < u]), \quad u < \alpha.$$

These relations are easily verified provided

$$\nu_u([g > u] \cap [f \leq u]) \leq 0, \quad u \geq \alpha,$$

and

$$\nu_u([g < u] \cap [f \geq u]) \geq 0, \quad u < \alpha.$$

Now let  $\alpha \leq u < a < b < c$ , then

$$\nu_u([g > c] \cap [f < a]) \leq 0$$

which implies by (1.2) that

$$\nu_b([g > c] \cap [f < a]) \leq 0.$$

Since the left hand side is assumed finite, we let  $a \downarrow u$ , and get

$$\nu_b([g > c] \cap [f \leq u]) \leq 0.$$

By right continuity we let  $b \downarrow u$ , and get

$$\nu_u([g > c] \cap [f \leq u]) \leq 0.$$

For  $c \downarrow u$ , we get

$$\nu_u([g > u] \cap [f \leq u]) \leq 0.$$

The proof of the other relation

$$\nu_u([g < u] \cap [f \geq u]) \leq 0 \quad u < \alpha,$$

is similar, and the result now follows as in Proposition 3.1.

The purpose of Proposition 3.5 is to prove

**THEOREM 3.6.** Let  $(\Omega, \mathcal{A}, \mu)$  be a measure space, and let  $X$  and  $g$  be functions in  $L_2(\Omega, \mathcal{A}, \mu)$ . Assume further that  $X^+ \in L_1(\Omega, \mathcal{A}, \mu)$ .

Let  $f = E(X || \mathcal{N})$ . Let  $T$  be convex and let  $t(a)$  be its derivative from left if  $a \geq 0$ , and right for  $a < 0$ . Then

$$\int [T(X) - T(g) - (X - g)t(g)]d\mu \geq \int [T(X) - T(f) - (X - f)t(f)]d\mu .$$

A theorem of this kind appears in Brunk [8]. (Here the hypothesis  $X^+ \in L_1$  replaces the hypothesis  $T(X), T(f)$ , and  $T(g) \in L_1$  and  $t(f)$ , and  $t(g) \in L_2$  from [8]).

*Proof.* If we define  $\psi(\omega, u) = X(\omega) - u$ , then for  $u > 0$ ,

$$\int \psi^+(\omega, u)d\mu = \int (X - u)I[X > u]d\mu < \infty$$

since

$$0 \leq (X - u)I[X > u] \leq \frac{1}{4u} X^2 ,$$

and  $\int X^2 d\mu < \infty$ . Similarly it is seen that

$$\int \psi^-(\omega, u)d\mu < \infty , \quad u < 0 .$$

Finally since  $\int X^+ d\mu < \infty$  we can define  $\{\nu_a\}$  by

$$\nu_a(A) = \int_A (X - a)d\mu ,$$

and  $\{\nu_a\}$  now satisfies (1.1), (1.2), and (1.3) with  $\alpha = 0$ . The function  $t$  determines a measure  $\gamma \geq 0$  by the relation

$$\gamma([a, b]) = t(b + 0) - t(a + 0) .$$

The definition (3.2) now yields

$$k(\omega, u) = T(X(\omega)) - T(u) - (X(\omega) - u)t(u) .$$

In order to prove the theorem we only have to verify the conditions on continuity and finiteness of the family  $\{\nu_a\}$ . Clearly, the function  $\nu_a(A)$  is continuous if finite, and if  $g \in L_2$ , then for  $0 < b < c$  we have that

$$\nu_b([g > c]) = \int XI[g > c]d\mu - b \int I[g > c]d\mu$$

which is finite since  $I[g > c] \in L_1$  and  $L_2$  and  $X \in L_2$ . Similarly, for  $c < b < 0$ , we have  $\nu_b([g < c])$  is finite. Thus the theorem follows from Proposition 3.5.

In the previous sections we have taken our starting point in a family  $\{\nu_\alpha, \alpha \in R\}$  of measures. Under certain conditions we can construct a LRN function associated with this family and a  $\sigma$ -lattice. Under other conditions we have, in §3, constructed a family of functions (3.2) whose mixtures with respect to  $\mu$  are quasi convex functions, and a certain functional  $I$ , which is minimized by the LRN function  $f$ .

When the theory is applied, the starting point is usually a family of functions, whose mixtures are quasi convex, and a minimization problem involving these quasi convex functions. Thus, in order to apply the previous results, we must show how such a family of functions gives rise to a family of measures and how the conditions imposed on the measures arise in a natural way from conditions on the quasi convex functions.

We state briefly a few properties of quasi convex functions: (1) Let  $k$  be an extended real valued function in  $R$ . The function  $k$  is called quasi convex if  $[k \leq a]$  is convex,  $a \in R$ . (2) A quasi convex function  $k$  determines  $m_1$  and  $m_2$ , such that  $-\infty \leq m_1 \leq m_2 \leq \infty$ , and such that  $k$  is decreasing in  $] -\infty, m_2[$ , and increasing in  $]m_1, \infty[$ . (3)  $k$  is called strictly quasi convex if it is quasi convex and strictly decreasing in  $] -\infty, m_2[$  and strictly increasing in  $]m_1, \infty[$ . (3) If  $[k \leq a]$ ,  $a \in R$ , is closed, then  $k$  is lower semi-continuous (l.s.c.), and if  $k$  is also quasi convex, then  $k$  is right continuous in  $] -\infty, m_2[$  and left continuous in  $]m_1, \infty[$ . (4) A quasi convex l.s.c. function  $k$  determines a signed measure  $\gamma$  by the relation

$$\gamma([a, b]) = k(b + 0) - k(a + 0) .$$

The measure  $\gamma$  is nonpositive at subsets of  $] -\infty, m_2[$  and non-negative at subsets of  $]m_1, \infty[$ . (5) A measure with these properties determines a quasi convex lower semi-continuous function by

$$(3.4) \quad k(\theta) = \gamma^+([-\infty, \theta]) + \gamma^-([\theta, \infty]) .$$

Let now  $(\Omega, \mathcal{A}, u)$  be a measure space, and associate with each  $\omega$  a l.s.c. quasi convex function  $k(\omega, \cdot)$  such that  $k(\cdot, \theta)$  is measurable  $\mathcal{A}$ , and such that  $\inf_\theta k(\omega, \theta) = 0$ . Let  $\gamma_\omega$  denote the measure associated with  $k(\omega, \cdot)$ . Let us assume that there exists a  $\sigma$ -finite measure  $\gamma$ , such that

$$(3.5) \quad \gamma_\omega \ll \gamma, \quad \omega \in \Omega .$$

If we denote the Radon-Nikodym derivative of  $\gamma_\omega$  with respect to  $\gamma$  by  $-\psi(\omega, \cdot)$  then combining (3.4) with (3.5) we get

$$(3.6) \quad k(\omega, \theta) = \int_{] \theta, \infty[} \psi^+(\omega, u) \gamma(du) + \int_{] -\infty, \theta[} \psi^-(\omega, u) \gamma(du) .$$

Notice that (3.6) is the l.s.c. version of (3.2). Under the assumption (3.5) we have now constructed the function  $\psi$ , and we further assume that

$$(3.7) \quad \nu_a(A) = \int_A \psi(\omega, a) \mu(d\omega), \quad A \in \mathcal{F},$$

is well defined and satisfies the regularity condition (1.3). We want the function defined by

$$\varphi(\theta) = \int_A k(\omega, \theta) \mu(d\omega) = \int_{] \theta, \infty[} \nu_u^+(A) \gamma(du) + \int_{]-\infty, \theta[} \nu_u^-(A) \gamma(du)$$

to be quasi convex, and this is done by imposing the condition (1.2) on the measures  $\{\nu_a\}$ . We further have that  $m_1(\varphi) = \underline{M}(A)$  and  $m_2(\varphi) = \overline{M}(A)$ . Furthermore  $\varphi$  is strictly increasing in  $] \overline{M}(A), \infty[$ , in the sense that if  $\gamma([a, b]) > 0$ ,  $\overline{M}(A) < a < b$ , then  $\varphi(a) < \varphi(b)$ . Similarly  $\varphi$  is strictly decreasing in  $] -\infty, \underline{M}(A)[$ .

The minimization problem, which we can solve, is now that of minimizing

$$I(h) = \int_{[h \geq \alpha]} k(\omega, h(\omega) - 0) \mu(d\omega) + \int_{[h < \alpha]} k(\omega, h(\omega) + 0) \mu(d\omega).$$

in the class of  $\mathcal{M}$ -measurable functions (see also (3.3)).

Let us consider some examples.

EXAMPLE 3.7. Let  $k(\omega, \cdot)$  be continuous, in which case  $\gamma$  can be chosen without atoms. Let  $\mathcal{M} = \{\emptyset, \Omega\}$  and let  $|\nu_a(\Omega)| < \infty$ . Then an  $\mathcal{M}$ -measurable function is a constant, and any value in  $[\underline{M}(\Omega), \overline{M}(\Omega)]$  is a LRN function associated with  $\{\nu_a\}$  and  $\mathcal{M}$ . It follows from Theorem 3.2 that

$$I(\theta) = \int k(\omega, \theta) \mu(d\omega)$$

is minimized by any LRN function. This just reflects the fact that  $I(\cdot)$  is l.s.c., and continuous when finite (see [6], Th. 2.2).

EXAMPLE 3.8. Let  $\Omega = \{1, \dots, N\}$  and let  $\mathcal{M}$  be any sub- $\sigma$ -lattice of subsets of  $\Omega$ . Let  $\mu_i > 0$ ,  $i = 1, \dots, N$ . Let  $k(i, \cdot)$ ,  $i = 1, \dots, N$ , be left continuous quasi convex functions, and let  $k(i, \cdot)$  have the integral representation (3.6) for some  $\gamma$  and some right continuous finite functions  $\psi(i, \cdot)$ . Let  $\{\nu_a\}$  be defined by

$$\nu_a(A) = \sum_{i \in A} \psi(i, a) \mu_i, \quad A \in \mathcal{M}.$$

Then  $\{\nu_a\}$  satisfy (1.1) and (1.3) for arbitrary  $\alpha$ . Let us define  $\alpha = -\infty$ .

If we also assume that  $\{\nu_a\}$  satisfy (1.2), then the mixtures

$$\sum_{i \in A} k(i, \cdot)$$

are quasi convex. This will in particular be the case if  $k(i, \cdot)$  is convex, since then  $\psi(i, \cdot)$  and  $\nu \cdot(A)$  can be chosen nonincreasing. Then  $\nu \cdot(A)$  is right continuous, and it follows from proposition 3.5 that

$$\sum_{i=1}^N k(i, h(i))\mu_i$$

is minimized in the class of  $\mathcal{M}$ -measurable functions by a LRN function of  $\{\nu_a\}$  and  $\mathcal{M}$ .

The problem of minimizing a sum of the form  $\sum_{i=1}^N k(i, h(i))\mu_i$  in the class of  $\mathcal{M}$ -measurable functions, when the functions  $k(i, \cdot)$  are quasi convex, has been discussed by van Eeden [12] and by Robertson and Waltman [18]. Van Eeden assumes that the functions and their mixtures are strictly quasi convex, and Robertson and Waltman relax this condition but assume continuity. Our contribution here is to identify the solution with the LRN function associated with  $\{\nu_a\}$  and to point out that the same function  $f$  provides the solution to a wide class of minimization problems obtained by varying  $\gamma$ . (cf. Theorem 3.6).

The representation given in Theorem 1.12 gives an explicit formula for  $f$ . Such formulas are also given by van Eeden and by Robertson and Waltman. If further  $\{\nu_a\}$  is decreasing at zero with respect to  $\mu$ , so that  $\mu(A) > 0$  implies  $\bar{M}(A) = \underline{M}(A) = M(A)$ , and if  $\nu_{(\cdot)}$  is continuous, then  $f$  is given by the following algorithm. Determine  $B_1$  as the largest set in  $\mathcal{M}$  maximizing  $M(B)$ ,  $B \in \mathcal{M}$ . Set  $f(i) = M(B_1)$ ,  $i \in B_1$ . Then choose  $B_2 \in \mathcal{M}$ , as the largest set in  $\mathcal{M}$  containing  $B_1$ , which maximizes  $M(B \cap B_1^c)$  in this class. Set  $f(i) = M(B_2 \cap B_1^c)$  for  $i \in B_2 \cap B_1^c$ . Continue until an integer  $k$  is reached such that  $B_k = \Omega$ . A proof can be based on Theorem 1.12 and properties of  $M(\cdot)$  mentioned immediately following Definition 1.11. We illustrate the ideas by showing that the LRN function given by Theorem 1.12 has the value  $M(B_1)$  on  $B_1$ . We note first that  $M(B_1) = \inf_C M(B_1 C)$ . For if there is a set  $C \in \mathcal{M}^c$  such that  $M(B_1 C) < M(B_1)$  then  $M(B_1 C^c) > M(B_1)$ , contradicting the choice of  $B_1$ . Second,  $\sup_B \inf_C M(BC)$  is achieved for  $B = B_1$ . To see this, set  $y_i = M(B_i B_{i-1}^c)$ ,  $i = 1, 2, \dots, k$ , and suppose  $B \in \mathcal{M}$ . We have  $B = BB_1 \cup BB_2 B_1^c \cup \dots \cup BB_k B_{k-1}^c$ . For  $i = 1, 2, \dots, k$ ,  $M(BB_i B_{i-1}^c) \leq y_i \leq y_1$  by the choice of  $B_i$ , hence  $M(B) \leq y_1$ , whence  $\inf_C M(BC) \leq y_1$ . Thus  $\sup_B \inf_C M(BC)$  is achieved for  $B = B_1$ .

**EXAMPLE 3.9.** (Maximum likelihood estimation of ordered parameters in the generalized exponential distribution.)

The following theorem can be found in [6].

**THEOREM.** *If*

(1) *F is a distribution function of a nondegenerate random variable X with range in the finite interval ]a, b[,*

(2) *φ(x, θ) is a bounded Borel function on [a, b] × [a, b] which is nondecreasing in θ for each x ∈ [a, b], nonnegative for θ ≥ x, non-positive for θ < x,*

(3) *P{φ(X, ·) is discontinuous at θ} = 0, then there exists a non-decreasing function θ on R with range in [a, b], such that*

$$\int \exp \left\{ - \int_0^\tau \varphi(x, \theta(u)) du \right\} dF(x) = 1, \quad \tau \in R.$$

We call the family of distributions with densities

$$\exp \left\{ - \int_0^\tau \varphi(x, \theta(u)) du \right\}$$

with respect to  $dF$ , the generalized exponential distributions. If we choose  $\varphi(x, \theta) = \theta - x$ , we get the exponential distributions, and we can then choose

$$\theta(u) = \frac{d}{du} \log \Phi(u),$$

where

$$\Phi(u) = \int \exp (ux) dF(x).$$

Let now  $X_{ij}, j = 1, \dots, n_i, i = 1, \dots, n,$  be independent random variables such that  $X_{ij}$  has density

$$\exp \left\{ - \int_0^{\tau_i} \varphi_i(x, \theta_i(u)) du \right\}$$

with respect to  $dF$ . The maximum likelihood estimate of  $\tau_i$  can be found by minimizing the convex function

$$k(i, \tau_i) = \sum_{j=1}^{n_i} \int_0^{\tau_i} \varphi_i(x_{ij}, \theta_i(u)) du.$$

If, however, the  $\tau_i$ 's are partially ordered we consider that as measurability with respect to  $\mathcal{M}$ , and we have to minimize

$$\sum_{i=1}^n k(i, \tau_i)$$

in the class of  $\mathcal{M}$ -measurable functions  $\tau(\cdot)$ .

It is seen that  $k(i, \cdot) - k(i, 0)$  has the representation (3.6) if we choose

$$\psi(i, \tau) = - \sum_{j=1}^{n_i} \varphi(x_{ij}, \theta_i(u))$$

and  $\gamma$  as Lebesgue measure. We define

$$\nu_\tau(A) = \sum_{i \in A} \psi(i, \tau), \quad \tau \in R,$$

then  $\nu_\tau(A)$  is nonincreasing and finite and the family  $\{\nu_\tau\}$  satisfies (1.1), (1.2), and (1.3) with arbitrary  $\alpha$ . It follows by Theorem 3.2 that

$$\sum_{i=1}^n k(i, \tau_i)$$

is minimized by a LRN function of  $\{\nu_a\}$  in the class of  $\mathcal{M}$ -measurable functions. Thus the solution is given explicitly by the representation theorems in § 1, as well as by the algorithm mentioned in Example 3.8.

#### REFERENCES

1. E. M. Alfsen, *Order theoretic foundations of integration*, Math. Ann. **149** (1963), 419-461.
2. E. S. Andersen and B. Jessen, *Some limit theorems on set functions*, Mat.-Fys. Medd. Danske Vid. Selsk. **25** (1948).
3. M. Ayer, H. D. Brunk, G. M. Ewing, W. T. Reid and E. Silverman, *An empirical distribution function for sampling with incomplete information*, Ann. Math. Statist. **26** (1955), 641-647.
4. H. Brøns, *Mean-values*, Preprint, 1963.
5. ———, *Forelaesningsnoter i matematisk statistik og sand synlighedsregning*, University of Copenhagen, 1965.
6. H. Brøns, H. D. Brunk, W. E. Franck, Jr., and D. L. Hanson, *Generalized means and associated families of distributions*, Ann. Math. Statist. **40** (1969), 339-355.
7. H. D. Brunk, *On an inequality for convex functions*, Proc. Amer. Math. Soc. **7** (1956), 817-824.
8. ———, *Conditional expectation given a  $\sigma$ -lattice and applications*, Ann. Math. Statist. **36** (1965), 1339-1350.
9. H. D. Brunk, G. M. Ewing and W. T. Reid, *The minimum of certain definite integrals suggested by the maximum likelihood estimate of a distribution function*, Bull. Amer. Math. Soc. **60** (1954), 535.
10. H. D. Brunk, G. M. Ewing and W. R. Utz, *Minimizing integrals in certain classes of monotone functions*, Pacific J. Math. **7** (1957), 833-847.
11. E. D. Cashwell and C. J. Everett, *A general mean value theorem*, Bull. Amer. Math. Soc. **74** (1968), 1135-1138.
12. C. van Eeden, *Maximum likelihood estimation of partially or completely ordered parameters, I and II*, Indag. Math. **19** (1957), 128-136, 201-211.
13. U. Grenander, *On the theory of mortality measurements*, Part II, Skand. Aktuarietidsskrift **39** (1956), 125-153.
14. E. Hewitt and K. Stromberg, *Real and abstract analysis*, Springer, New York, 1965.



15. S. Johansen, *The descriptive approach to the derivative of a set function with respect to a  $\sigma$ -lattice*, Pacific J. Math. **21** (1967), 49-58.
16. J. Neveu, *Calcul des probabilités*, Masson et Cie, Paris, 1964.
17. T. Robertson, *A representation for conditional expectations given  $\sigma$ -lattices*, Ann. Math. Statist. **37** (1966), 1279-1283.
18. T. Robertson and P. Waltman, *On estimating monotone parameters*, Ann. Math. Statist. **39** (1968), 1030-1039.

Received May 8, 1969.

UNIVERSITY OF MISSOURI (Now at Oregon State University)  
UNIVERSITY OF COPENHAGEN

