MINIMIZING INTEGRALS IN CERTAIN CLASSES OF MONOTONE FUNCTIONS

H. D. BRUNK, G. M. EWING AND W. R. UTZ

1. Introduction. This paper is concerned with the existence, uniqueness and representation of minimizing functions. It includes many results of [1] and [2]. Applications are discussed in [3].

The authors are indebted for various ideas to W. T. Reid with whom Brunk and Ewing collaborated in a study [2] of a particular integral (1.4) in the one-variable case. Also, the authors wish to acknowledge the helpful suggestions of the referee.

Extension to n variables and to more general integrands is of interest per se and is motivated by a variety of problems.

For example, let $\boldsymbol{x}$ ($\boldsymbol{y}$) be the random variable, maximum dilution (that is, unity minus concentration) of an insecticide I (J) which is lethal to an insect from a given population. Then

$$p(x, y) = \Pr\{\boldsymbol{x} > x \text{ or } \boldsymbol{y} > y\}$$

is the probability of death for an insect similtaneously dosed with respective dilutions x, y of I, J. Moreover

$$F(x, y) = 1 - p(x, y) = \Pr\{\boldsymbol{x} \leqq x \text{ and } \boldsymbol{y} \leqq y\} , \tag{1.1}$$

is the probability of survival and is a distribution function [5; pp. 78, 260]; hence $p(x, y)$ is nonincreasing in each variable and for each point-pair (x, y), (x', y'),

$$\Delta^2 p = p(x', y') - p(x', y) - p(x, y') + p(x, y) \leqq 0 . \tag{1.2}$$

For each of selected pairs (x_i, y_j) let $\Delta\mu_{ij}$ insects be dosed and let α_{ij} denote the fraction of the sample which is killed. The maximum likelihood estimate $P(x, y)$ of $p(x, y)$ is that function, subject to the restrictions stated above, which maximizes the product

$$\prod p_{ij}^{\alpha_{ij}\Delta\mu_{ij}}(1-p_{ij})^{(1-\alpha_{ij})\Delta\mu_{ij}} , \qquad p_{ij} = p(x_i, y_j) . \tag{1.3}$$

Equivalently, $P(x, y)$ minimizes the integral

$$-\int [\alpha \log p + (1-\alpha)\log(1-p)]d\mu , \tag{1.4}$$

in which μ describes the mass distribution consisting of masses $\Delta\mu_{ij}$ at

Received December 5, 1955, and in revised form March 22, 1956. This research was supported by the United States Air Force, through the Office of Scientific Research of the Air Research and Development Command.

the respective points (x_i, y_j) and no mass elsewhere.

Other problems, for example, [3, p. 610] require only that the function $P(x, y)$ minimizing (1.4) be monotone in each variable and not that it satisfy (1.2). As a further example of this type, suppose that $\alpha(t)$, $t=(t^1, \cdots, t^n)$, not necessarily monotone in any t^i, is a given approximation to $\theta(t)$ a function required to be monotone in each variable. The least squares determination $\Theta(t)$ of $\theta(t)$ minimizes the integral $\int[\theta(t)-\alpha(t)]^2 d\mu(t)$.

2. Formulation and preliminary lemmas. Given a fixed positive integer n and the space R_n with points $t=(t^1, \cdots, t^n)$, let μ be a measure defined on a Borel field $\mathscr{B}$ of subsets of R_n which is *totally finite*, that is, $R_n \in \mathscr{B}$, $\mu(R_n) < \infty$, and *complete*, that is, if $A \subset E \in \mathscr{B}$ and $\mu(E)=0$, then $A \in \mathscr{B}$ and $\mu(A)=0$. The term *measure* will mean μ-measure unless otherwise specified, *measurable set* will mean a set in $\mathscr{B}$, and *measurable function* a μ-measurable function. In particular μ can be a finite Lebesgue-Stieltjes measure.

Let I be a fixed nondegenerate interval of extended real numbers which includes its endpoints $a \geqq -\infty$, $b \leqq \infty$. Let $F(u, v)$ be an extended real-valued function for $u, v \in I$ subject to the following conditions.

(2.1) If $\alpha(t)$, $\theta(t)$ with ranges in I are both measurable then so is $F[\alpha(t), \theta(t)]$.

(2.2) For fixed u in I, either (i) $F(u, v)=\infty$ for $v \neq u$, with $F(u, u) < \infty$ or (ii) $F(u, v)$ is strictly decreasing (increasing) in v for $a \leqq v \leqq u (u \leqq v \leqq b)$ and right (left) continuous in v for $a < v < u (u < v < b)$. (See (5.3) to (5.6) for examples.)

$$\int F[\alpha(t), \alpha(t)] d\mu(t) > -\infty \ . \tag{2.3}$$

For fixed $\alpha(t)$ and arbitrary $\theta(t)$ with ranges in I and both measurable define

$$J[\theta] = \int F[\alpha(t), \theta(t)] d\mu(t) \ . \tag{2.4}$$

Let M denote the class of all measurable functions $\theta(t)$ with ranges in I such that $J[\theta]$ exists finite or infinite and such that $\theta(t)$ is nondecreasing in each coordinate t^i of t. Define M^* as M if $n=1$ and, for $n > 1$, let M^* consist of those $\theta(t)$ in M with the property that the difference $\Delta^2\theta$, defined as in (1.2) for each pair t^i, t^j, (with the other variables fixed for each choice of i, j) shall be nonnegative in the complement of the closure of the set on which $\theta(t)=\infty$ or $\theta(t)=-\infty$. The

principal problems of this paper to minimize J in M and in M^*.

The methods apply, with suitable small changes, to problems like that of § 1 in which (1.2) is required with $\leqq$ instead of $\geqq$ and in which admissible functions are nonincreasing in the separate variables.

The relation $t_1 < t_2$ means that $t_1^i < t_2^i$, $i=1, \cdots, n$ while $t_1 \leqq t_2$ means that $t_1^i \leqq t_2^i$ for each i. Given a point v consider intervals of the types $(t : t < v)$ and $(t : t \leqq v)$. A measurable set L which is a union of intervals of the first and (or) second of these types is termed a *lower layer*. A measurable set L is then a lower layer if and only if $v \in L$ and $t \leqq v$ imply $t \in L$. An *upper layer* U is similarly defined. The complement $\tilde{L}$ of a lower layer L is an upper layer. If L is not void the common boundary of L and $\tilde{L}$ is called a *monotone graph*. Given a lower layer L and an upper layer U the measurable set $UL=L-\tilde{U}$ is termed a *layer*. For $n=1$, a layer is an interval of the reals which may be void, degenerate or of positive length and, in the latter case, may include either, neither, or both of its endpoints. The layer is the natural extension (for the purposes of this study) of the notion of interval. A monotone graph is connected and is a layer but, for $n > 1$, a layer need not be connected.

LEMMA 2.1. *Let $\mathcal{O}$ denote the union of all open sets of measure* 0. *Then $\mu(\mathcal{O})=0$ and given $t \notin \mathcal{O}$, every layer containing a neighborhood of t has positive measure.*

LEMMA 2.2. *If $\theta(t)$ is measurable and monotone nondecreasing in each variable t^i, then the set of points t, for which $\theta(t)$ is on a given finite or infinite interval of the reals, is a layer.*

The proofs of these lemmas are easy.

LEMMA 2.3. *If $\theta(t)$ is monotone nondecreasing in each variable t^i, the discontinuities of $\theta(t)$ lie on a countable set of monotone graphs.*

This result is Theorem 7 in [4].

LEMMA 2.4. *A monotone graph is of Lebesgue measure* 0.

Proof. The metric density is less than unity at each point of a monotone graph. Alternatively, observe that a line with direction numbers $(1, \cdots, 1)$ cuts a monotone graph in exactly one point and use Fubini's Theorem.

3. Existence theorems. Denote the respective infima of J in M, M^* by γ, γ^*.

LEMMA 3.1. *If $\theta(t)$, $q=1, 2, \cdots$, is a sequence in M^*, there exists a subsequence $\theta_q^*(t)$ of $\theta_q(t)$ and $\theta(t) \in M^*$ such that $\lim \theta_q^*(t)=\theta(t)$ except at most on two monotone graphs.*

Theorem 3 of [4] establishes this result. The exceptional sets, denoted by A and Ω, are respectively boundaries of layers on which $\theta(t)=-\infty, \infty$.

LEMMA 3.2. *If $n=1$, then any sequence in $M=M^*$ contains a subsequence converging everywhere on R_1 to a function $\theta(t)$ in $M=M^*$.*

Proof. If $n=1$ each of the sets A, Ω is either void or consists of a single point; hence the sequence $\theta_q^*(t)$ can be further refined to yield convergence (possibly to ∞ or $-\infty$) everywhere on R_1.

THEOREM 3.1. EXISTENCE THEOREM FOR M^*. *There exists a function $\Theta(t)$ in M^* such that $J[\Theta]=\gamma^*$.*

Proof. Attention is confined to the nontrivial case $\gamma^* < \infty$. Let $\theta_q(t)$ be a sequence in M^* such that $\lim J[\theta_q]=\gamma^*$. By Lemma 3.1 we may suppose that $\theta_q(t)$ converges, for $t \in R_n - A \cup \Omega$, to $\Theta(t) \in M^*$. Let $\theta_*(t)=\liminf \theta_q(t)$, $\theta^*(t)=\limsup \theta_q(t)$. Extend $\Theta(t)$ to $A \cup \Omega$ by the definition,

$$\Theta(t)\begin{cases} = \alpha(t), & \text{if} \quad \theta_*(t) \leqq \alpha(t) \leqq \theta^*(t), \\ = \theta_*(t), & \text{if} \quad \alpha(t) < \theta_*(t), \\ = \theta^*(t), & \text{if} \quad \alpha(t) > \theta^*(t). \end{cases} \tag{3.1}$$

Clearly $\Theta(t)$ is measurable. One varifies that $\Theta(t)$ is in M^*.

For fixed t, it follows from the definition of $\Theta(t)$ and property (2.2) of F that $F[\alpha(t), \Theta(t)] \leqq F[\alpha(t), v]$ for $\theta_*(t) \leqq v \leqq \theta^*(t)$. Since each point of accumulation of the sequence $\theta_q(t)$ lies in the interval $[\theta_*(t), \theta^*(t)]$ we have

$$F[\alpha(t), \Theta(t)] \leqq \liminf F[\alpha(t), \theta_q(t)] .$$

From Fatou's Lemma [6, p. 113; 7, p. 167] it then follows that

$$J[\Theta]=\int F[\alpha(t), \Theta(t)]d\mu(t) \leqq \liminf \int F[\alpha(t), \theta_q(t)]d\mu(t) ;$$

hence $J[\Theta]=\gamma^*$.

THEOREM 3.2. EXISTENCE THEOREM FOR M. *If $n=1$ or 2 there exists a function $\Theta(t)$ in M such that $J[\Theta]=\gamma$.*

Proof. If $n=1$ the conclusion is contained in that of the preceding theorem. For $n=2$, let $\theta_q(t)$ be a sequence in M such that $\lim J[\theta_q]=\gamma$.

By Lemma 3.2 and the usual diagonalization process there exists a subsequence converging at all points with at least one rational coordinate.

Define $\theta_*(t)$, $\theta^*(t)$ as in the proof of Theorem 3.2 noting that $\theta_*(t)=\theta^*(t)$ if t^1 or t^2 is rational. From the density of these points in R_2, $\theta_*(t)=\theta^*(t)$ at any point t at which both functions are continuous; hence by Lemma 2.3 everywhere except on a countable set of monotone graphs. Define $\Theta(t)$ on the space R_2 as in (3.1).

If $t_1 \leqq t_2 \neq t_1$, the segment with endpoints t_1, t_2 is cut by at least one line on which t^1 is rational or on which t^2 is rational in a point t_0. One sees that

$$\Theta(t_1) \leqq \theta^*(t_1) \leqq \theta^*(t_0)=\theta_*(t_0) \leqq \theta_*(t_2) \leqq \Theta(t_2)\,;$$

hence that Θ is nondecreasing in t^1 and in t^2 and is in M. The proof can be completed by following that of Theorem 3.1.

The point t_0 essential to the last proof need not exist for $n>2$.

A function $t(\tau)$, τ on a finite or infinite interval will be termed a *monotone nondecreasing vector-function* if $\tau_2>\tau_1$ implies that $t(\tau_2) \geqq t(\tau_1)$. If $\theta(t)$ is nondecreasing in each variable t^i and $t(\tau)$ has the above property, then $\theta[t(\tau)]$ is nondecreasing in the real variable τ. *Monotone nonincreasing vector-functions* are similar. The graph of a monotone vector-function is a monotone graph in the sense of § 2 only for certain cases when $n=1$ or 2.

In the following theorem we suppose the class of measurable sets is contained in the class of Lebesgue measurable subsets of R_n. These is then a Lebesgue decomposition [6, p. 134] of μ; that is, μ is the sum of a measure α absolutely continuous with respect to Lebesgue measure λ and a measure σ singular with respect to λ. Thus if $\lambda(E)=0$, then $\alpha(E)=0$ and there is a decomposition of R_n into complementary sets A, $\tilde{A}$ such that $\lambda(A)=0$ and $\sigma(\tilde{A})=0$.

THEOREM 3.3. EXISTENCE THEOREM FOR M. *If $\mu=\alpha+\sigma$ is the Lebesgue decomposition of the given measure μ, and if the mass in R_n described by the singular part σ all lies on the graphs of a counterable set of monotone vector-functions, then there exists a function $\Theta(t)$ in M such that $J[\Theta]=\gamma$.*

This theorem applies in particular if σ describes a discrete mass distribution or if μ is Lebesgue measure. The proof, along lines similar to those followed in preceding theorems, is omitted.

4. Integrands generated by convex functions. The class of problems for which we are able to give more complete results in this section is more restricted than that of § 3. We are moreover primarily interested in cases in which the minimum of J in M is finite. It is convenient

to introduce $\mathscr{M}$ to denote the subset of M consisting of those θ in M such that $J[\theta] < \infty$. Results of this section are for the minimum problem in $\mathscr{M}$ only.

Let I again denote a fixed nondegenerate interval of extended reals which includes its endpoints $a \geqq -\infty$ and $b \leqq \infty$. Let $T(z)$ be a continuous convex function of the real variable z on the interior I_0 of I. The derivative $T'(z)$ exists except on a countable subset of I_0 and it seems convenient to extend $T'(z)$ to I_0 by assigning it the value of the left derivative at each point of I_0, thereby making $T'(z)$ left-continuous on I_0. The extended real-valued function $F(u, v)$ is defined as follows:

$$F(u, v) = T(u) - T(v) - (u - v)T'(v) , \qquad u, v \in I_0 . \tag{4.1}$$

The right member of (4.1) has an obvious interpretation in terms of the tangent to the graph of $T(z)$. $F(u, v)$ is extended to $I \times I$ by the additional definitions

$$\begin{aligned} F(a, v) &= \lim_{u \to a} F(u, v) , && v \in (a, b) , \\ F(b, v) &= \lim_{u \to b} F(u, v) , && v \in (a, b) , \\ F(u, a) &= \lim_{v \to a} F(u, v) , && u \in (a, b] , \\ F(u, b) &= \lim_{v \to b} F(u, v) , && u \in [a, b) , \end{aligned} \tag{4.2}$$

One verifies that, for $u, v \in I_0$,

$$\begin{aligned} F(u, v) &= \int_{[u, v)} (z - u) dT'(z) , && \text{if } u < v , \\ &= -\int_{[v, u)} (z - u) dT'(z) , && \text{if } u > v , \\ &= 0 && \text{if } u = v . \end{aligned} \tag{4.3}$$

Essentially such functions F generalizing the particular integrand of [2] have been suggested independently by Reid.

Such functions F arise in connection with the applications (cf. examples in § 1, also [3], where $\exp\{-F[g(x), \theta]\}$ is the density function, with respect to a measure, of a random variable whose distribution belongs to the exponential family). F as defined above is nonnegative, and has properties (2.1) and (2.2) (except that F need not be right-continuous in v for $a < v < u$ and F is strictly monotone in v for $v \leqq u$ and for $v \geqq u$ only if T is strictly convex).

We again let t denote the generic point in R_n, let μ denote a totally finite complete measure on the given Borel field $\mathscr{B}$, and let $\alpha(t)$ denote a given integrable function with range in I such that $T[\alpha(t)]$ is integrable. It follows that $J[\theta] < \infty$ when $\theta(t) \equiv \theta_0$, θ_0 a constant in I_0, so

that $\theta(t) \equiv \theta_0 \in \mathscr{M}$.

For a measurable subset A or R_1 we define

$$J[\theta\,;\,A]=\int_A F[\alpha(t),\,\theta(t)]d\mu(t)\,, \tag{4.4}$$

$$M(A)=\int_A \alpha(t)d\mu(t)/\mu(A)\,, \qquad \text{if} \quad \mu(A)>0\,. \tag{4.5}$$

THEOREM 4.1. *Let ν be a finite signed measure on the class of measurable subsets of R_n, absolutely continuous with respect to μ. Then there exist an upper layer P and a lower layer N such that*

(i) $\nu(PL)>0$ *for every lower layer L such that* $\mu(PL)>0$;

(ii) $\nu(U\tilde{P})\leqq 0$ *for every upper layer U;*

(iii) $\nu(UN)<0$ *for every upper layer U such that* $\mu(UN)>0$;

(iv) $\nu(\tilde{N}L)\geqq 0$ *for every lower layer L.*

Proof. The proof is an adaptation of that of the Hahn-Jordan decomposition theorem [6, p. 121] and will simply be sketched here in broad outline. Let $\mathscr{N}$ denote the class (a class of sets having a nonpositive property) of lower layers L such that $\nu(UL)\leqq 0$ for every upper layer U. Choose a sequence of lower layers in $\mathscr{N}$ whose measures approach $\beta=\sup_{L\in\mathscr{N}}\mu(L)$; one readily verifies that their union, $\tilde{P}$, is a maximal element of $\mathscr{N}$; that is, $\tilde{P}$ belongs to $\mathscr{N}$ and has measure β. Thus the lower layer $\tilde{P}$ has the nonpositive property (ii). It is possible that the void set is the only element of $\mathscr{N}$, in which event $\tilde{P}=\phi$. We shall now show that P, the complement of $\tilde{P}$, has also the positive property (i). Suppose the contrary. Then there is a lower layer $T\supset\tilde{P}$ such that $\nu(PT)\leqq 0$, while $\mu(PT)>0$, so that $T\notin\mathscr{N}$ (since $\tilde{P}$ is maximal). Hence there is an upper layer $U\subset P$, $U\supset\tilde{T}$, such that $\nu(UT)>0$. One may then determine an expanding sequence (as in the proof in [6], pp. 121–122, of the existence of a Hahn decomposition) U_i, $i=1, 2, \cdots$, of upper layers, contained in P and containing $\tilde{T}$, whose limit, U^*, has a complement, $\tilde{U}^*$, belonging to $\mathscr{N}$, while

$$\nu(U_iU_{i-1})=\nu(U_i-U_{i-1})>0\,, \qquad i=1, 2, \cdots\,; \quad U_0=\tilde{T}\,.$$

From the maximality of $\tilde{P}$ it follows that $\mu(P\tilde{U}^*)=0$, whence $\nu(PU^*)=0$. On the other hand,

$$\nu(U^*T)=\sum_{i=1}^{\infty}\nu(U_i-U_{i-1})>0\,,$$

so that $\nu(PT) > 0$, a contradiction. Thus P does indeed have the positive property (i). The determination of a lower layer N, possibly void, with the desired properties follows similarly on the introduction of a class $\mathscr{P}$ of upper layers U such that $\nu(UL) \geqq 0$ for every lower layer L. For each real x, define a lower layer N_x and an upper layer P_x as the lower and upper layers N and P given by Theorem 4.1 corresponding to the signed measure

$$\nu(A) = \int_A [\alpha(t) - x] d\mu(t) .$$

As a consequence of Theorem 4.1 applied to this signed measure, we have for upper and lower layers U and L,

$$\begin{aligned} &M(P_x L) > x \quad \text{if} \quad \mu(P_x L) > 0 , \\ &M(U\tilde{P}_x) \leqq x \quad \text{if} \quad \mu(U\tilde{P}_x) > 0 , \\ &M(UN_x) < x \quad \text{if} \quad \mu(UN_x) > 0 , \\ &M(\tilde{N}_x L) \geqq x \quad \text{if} \quad \mu(N_x L) > 0 , \end{aligned} \tag{4.6}$$

LEMMA 4.1. *If Λ is an index set and A_λ, $\lambda \in \Lambda$, is a family of measurable subsets of R_n such that $\mu(A_\lambda A_\sigma) = 0$ for $\Lambda \ni \sigma \neq \lambda \in \Lambda$, then $\mu(A_\lambda) = 0$ except for at most a countable subset of Λ.*

Proof. If the lemma is false then there is a positive number ϵ and a sequence of sets of the family $\{A_\lambda, \lambda \in \Lambda\}$, each having measure greater than ϵ. It follows from the hypothesis

$$\mu(A_\lambda A_\sigma) = 0 \quad \text{for} \quad \Lambda \ni \sigma \neq \lambda \in \Lambda$$

that the usual technique of replacing the sets of a sequence by mutually disjoint sets while preserving their union yields a sequence of disjoint sets each having measure greater than ϵ, so that their union has infinite measure, contradicting the property of μ of being totally finite.

COROLLARY 4.1. *$\mu(P_x N_x) = 0$ for every real x and $\mu(\tilde{N}_x \tilde{P}_x) = 0$ for all but a countable set of real numbers x.*

Proof. If $\mu(P_x N_x) > 0$, the first and third relations (4.6) yield the contradiction $x < x$. It can be seen as follows that the second conclusion is a consequence of Lemma 4.1. Since

$$\tilde{N}_x \tilde{P}_x \cap \tilde{N}_y \tilde{P}_y \subset \tilde{N}_y \tilde{P}_x ,$$

it follows that when $x < y$ and $\mu(\tilde{N}_y \tilde{P}_x) > 0$, then $y \leqq M(N_y P_x) \leqq x$, which is a contradiction. It follows that

$$\mu(\tilde{N}_x\tilde{P}_x \cap \tilde{N}_y\tilde{P}_y)=0$$

for $x \neq y$.

It is convenient to determine the upper and lower layers P_x and N_x so that

$$N_x \subset \tilde{P}_x \subset N_y \quad \text{for} \quad x<y \quad (\text{or } P_x \supset \tilde{N}_y \supset P_y \text{ for } x<y)\,, \tag{4.7}$$

$$N_x=\bigcup_{y<x} N_y\,, \qquad P_x=\bigcup_{y>x} P_y\,. \tag{4.8}$$

Let E denote the countable set consisting of reals r which are rational or for which $\mu(\tilde{N}_r\tilde{P}_r)>0$. It can be shown that

$$N_x^*=\bigcup_{E\ni r<x}(N_r \cup \tilde{P}_r)\,, \qquad P_x^*=\bigcup_{y>x}\bigcap_{E\ni r\leqq y}(\tilde{N}_r \cap P_r)\,,$$

have properties (4.7) and (4.8) and that relations (4.6) hold with N_x^*, P_x^* in place of N_x, P_x. We shall understand from here on that this replacement has been made, but shall omit the asterisks.

Let us define $\Theta(t)$ as the infimum of those x such that $t \in \tilde{P}_x$.

LEMMA 4.2.

$$\Theta(t)>x \quad \textit{if and only if} \quad t \in P_x\,. \tag{4.9}$$

$$\Theta(t)<x \quad \textit{if and only if} \quad t \in N_x\,. \tag{4.10}$$

$$\Theta(t)=x \quad \textit{if and only if} \quad t \in (\tilde{N}_x\tilde{P}_x)\,. \tag{4.11}$$

$$\Theta(t)=\sup_{t\notin N_x} x\,. \tag{4.12}$$

Proof of (4.9). From its definition, $\Theta(t)\leqq x$ if $t \notin P_x$. If $\Theta(t)=x_0>x$, then $t \in P_y$ for $y<x_0$; hence $t \in \bigcup_{y>x} P_y=P_x$.

Proof of (4.10). If $t \notin N_x$, then $t \in P_y$ for each $y<x$; hence

$$\Theta(t)=\inf_{t\notin P_y} y \geqq x\,.$$

If $t \in N_x=\bigcup_{y<x} N_y$, there exists $y_0<x$ such that $t \in N_y$ for $y \geqq y_0$; hence $t \in \tilde{P}_y$ for $y>y_0$; hence

$$\Theta(t)=\inf_{t\notin P_y} y \leqq y_0<x\,.$$

Relation (4.11) follows from (4.9) and (4.10).

Proof of (4.12). Set $\Theta_1(t)=\sup_{t\notin N_x} x$. Arguments similar to the above show that $\Theta_1(t)$ satisfies (4.9), (4.10), (4.11); hence that $\Theta(t)\equiv\Theta_1(t)$.

We remark that, for $\theta\in\mathscr{M}$,

$$M\{t:\ \Theta(t)\leqq z<\theta(t)\}\leqq z\ ,\qquad M\{t:\ \Theta(t)\geqq z>\theta(t)\}\geqq z\ ,$$

provided that the measures of these sets are positive. Each strict inequality between z and $\theta(t)$ in these statements may be replaced by the corresponding weak inequality.

LEMMA 4.3. *If $\theta(t)\in\mathscr{M}$, if E is a measurable set, if a_n is a sequence of real numbers strictly decreasing to a (b_n a sequence strictly increasing to b), if $\theta_n(t)=\max[\theta(t),\ a_n]$ ($\min[\theta(t),\ b_n]$), then $\theta_n(t)\in\mathscr{M}$, $n=1,2,\cdots$, and $\lim J[\theta_n;\ E]=J[\theta;\ E]$.*

Proof We recall that the function of t assuming the constant value a_n is in $\mathscr{M}$, and that, as a function of v, $F(u,\ v)$ is nondecreasing for $v>u$ and nonincreasing for $v<u$. Since

$$\theta(t)\leqq\theta_n(t)\leqq\theta_1(t)=\max[\theta(t),\ a_1]\ ,$$

we have

$$\begin{aligned}0\leqq F[\alpha(t),\ \theta_n(t)]&\leqq\max\{F[\alpha(t),\ \theta(t)],\ F[\alpha(t),\ \theta_1(t)]\}\\&\leqq\max\{F[\alpha(t),\ \theta(t)],\ F[\alpha(t),\ a_1]\}\ .\end{aligned}$$

The functions $F[\alpha(t),\ \theta(t)]$, $F[\alpha(t),\ a_1]$ are integrable; so then is the function $\max\{F[\alpha(t),\ \theta(t)],\ F[\alpha(t),\ a_1]\}$. Also

$$\lim\theta_n(t)=\theta(t),\ \lim F[\alpha(t),\ \theta_n(t)]=F[\alpha(t),\ \theta(t)]\ ,$$

and by the dominated convergence theorem, $\lim J[\theta_n;\ E]=J[\theta;\ E]$.

LEMMA 4.4. *Given $\theta',\ \theta''\in\mathscr{M}$, let*

$$E=\{t:\ \theta'(t)<\theta''(t)\}\quad\textit{and}\quad E(z)=\{t:\ \theta'(t)\leqq z<\theta''(t)\}\ .$$

Then

$$J[\theta'';\ E]-J[\theta';\ E]=\int_{(a,\,b)}\{z-M[E(z)]\}\,\mu[E(z)]dT'(z)\ .$$

Proof. Let a_n and b_n, $n=1,2,\cdots$, be sequences strictly decreasing and increasing to the endpoints a and b of I repectively. Set

$$\theta'_n(t)=\max[\theta'(t),\ a_n]\ ,\quad\text{and}\quad\theta''_n(t)=\min[\theta''(t),\ b_n]\ ,$$

$n=1,2,\cdots$. We have

$$J[\theta_n''; E]-J[\theta_n'; E]=\int_E \{F[\alpha(t); \theta_n''(t)]-F[\alpha(t); \theta_n'(t)]\}d\mu$$
$$=\int_E d\mu\int_{[\theta_n'(t),\theta_n''(t))} [z-\alpha(t)]dT'(z) .$$

For fixed n, set

$$A=\{(z, t): \quad \theta'(t)\leqq z<\theta''(t) , \quad a_n\leqq z<b_n\} .$$

Both z and $\alpha(t)$ are integrable over A with respect to the product measure $(d_\mu\times dT')$, so that Fubini's Theorem permits a change in the order of integration. We have that

$$J[\theta_n''; E]-J[\theta_n'; E]=\int_{[a_n,b_n)} dT'(z)\int_{E(z)}[z-\alpha(t)]d\mu$$
$$=\int_{[a_n,b_n)}\{z-M[E(z)]\}\mu[E(z)]dT'(z) .$$

Applying Lemma 4.3 and taking limits as $n\to\infty$ we obtain the desired conclusion.

THEOREM 4.2. *$\Theta(t)$ minimizes J in $\mathscr{M}$.*

Proof. For $\theta(t)$ in $\mathscr{M}$, set

$$B_1=\{t: \Theta(t)<\theta(t)\} ,$$
$$B_2=\{t: \Theta(t)>\theta(t)\} ,$$
$$B_3=\{t: \Theta(t)=\theta(t)\} .$$

Then

$$J[\Theta]=\Sigma J[\Theta; B_i]$$

and similarly for $J[\theta]$. We have $J[\theta; B_3]-J[\Theta; B_3]=0$. In Lemma 4.4 set $\Theta=\theta'$, $\theta=\theta''$ so that E becomes B_1 and $E(z)$ becomes that set $\{t: \Theta(t)\leqq z<\theta(t)\}$. From Lemma 4.2 (see remark preceding Lemma 4.3) it follows that $M[E(z)]\leqq z$ if $\mu[E(z)]>0$; hence from Lemma 4.4 that

$$J[\theta; B_1]-J[\Theta; B_1]\geqq 0 . \tag{4.13}$$

Now set $\Theta=\theta''$, $\theta=\theta'$ in Lemma 4.4 and then $E=B_2$ and

$$E(z)=\{t: \Theta(t)>z\geqq\Theta(t)\} .$$

Again, from Lemma 4.2, $M[E(z)]\geqq z$ if $\mu[E(z)]>0$; hence, from Lemma 4.4,

$$J[\theta;\ B_2]-J[\Theta;\ B_2]\geqq 0\ . \tag{4.14}$$

Adding (4.13) and (4.14) we find that $J[\theta]-J[\Theta]\geqq 0$, completing the proof.

By (4.11), the minimizing function $\Theta(t)$ assumes a given value x on the layer $\tilde{N}_x\tilde{P}_x$. In calculating for specific examples it is useful to observe as a consequence of equations (4.6) that if $\mu(\tilde{N}_x\tilde{P}_x)>0$ then $\tilde{N}_x\tilde{P}_x$ is the maximal layer among layers $\tilde{N}_xL$ over which the mean is minimal:

$$M(\tilde{N}_x\tilde{P}_x)=x\leqq M(\tilde{N}_xL)$$

if $\mu(\tilde{N}_xL)>0$; while if $\tilde{N}_xL\supset\tilde{N}_x\tilde{P}_x$ and if $M(\tilde{N}_xL)=M(\tilde{N}_x\tilde{P}_x)$, then $\tilde{N}_xL$ and $\tilde{N}_x\tilde{P}_x$ differ by a set of measure 0. Similarly $\tilde{N}_x\tilde{P}_x$ is the maximal layer among layers $U\tilde{P}_x$ over which the integral mean of $\alpha(t)$ is maximal.

We term the subset of a neighborhood of a point t_0 consisting of points $t>t_0$, an *upper neighborhood* of t_0 and the subset consisting of points $t<t_0$, a *lower neighborhood* of t_0. Let $\mathscr{D}$ denote the set of points each of which has an upper or a lower neighborhood in $\mathscr{O}$ (defined in Lemma 2.1); $\mathscr{D}\supset\mathscr{O}$.

THEOREM 4.3. REPRESENTATION THEOREM. *If, given $\varepsilon>0$, we have $\mu(UN_{c+\varepsilon})>0$ for every upper layer U containing a given point t_0, and $\mu(P_{c-\varepsilon}L)>0$ for every lower layer L containing t_0, where $c=\Theta(t_0)$, then*

$$\Theta(t_0)=\sup_{U\ni t_0}\inf_{L\ni t_0} M(UL)\ , \tag{4.15}$$

$$\Theta(t_0)=\inf_{L\ni t_0}\sup_{U\ni t_0} M(UL)\ , \tag{4.16}$$

$$\Theta(t_0)=\sup_{U\ni t_0}\inf_{L} M(UL)\ , \tag{4.17}$$

$$\Theta(t_0)=\inf_{L\ni t_0}\sup_{U} M(UL)\ . \tag{4.18}$$

In particular (4.15), $\cdots$, (4.18) *hold if t_0 is a mass point of μ or if t_0 is a point of continuity of $\Theta(t)$ not in $\mathscr{D}$.*

We note that the measure of $\mathscr{O}$ is 0, and that the Lebesgue measure of $\mathscr{D}$-$\mathscr{O}$ is 0. Further, since $\Theta(z)\in\mathscr{M}$, its discontinuities lie on a countable set of monotone graphs (Lemma 2.3.). Theorem 4.3 thus gives almost everywhere representations of $\Theta(t)$, provided that μ is absolutely continuous with respect to Lebesgue measure, or provided that the Lebesgue singular part of μ concentrates its mass at a countable number of mass points. In general, these representations need not be valid almost everywhere.

Proof of (4.15) *and* (4.17). Given $U \ni t_0$ and given $\varepsilon > 0$, then $\mu(UN_{c+\varepsilon}) > 0$ by hypothesis, so that by (4.6), $M(UN_{c+\varepsilon}) < c+\varepsilon$. Hence

$$\inf_{L \ni t_0} M(UL) \leqq c \ , \quad \text{if} \quad U \ni t_0 \ , \tag{4.19}$$

and also

$$\inf_{L} M(UL) \leqq c \quad \text{if} \quad U \ni t_0 \ . \tag{4.20}$$

Further, if $\mu(\tilde{N}_c L) > 0$, then by (4.6), $M(\tilde{N}_c L) \geqq c$. But $\tilde{N}_c \ni t_0$, and hence relations (4.15) and (4.17) follow respectively from (4.19) and (4.20).

Relations (4.16) and (4.18) may be proved similarly.

We note that under the hypotheses of Theorem 4.3, if $\alpha(t)$ is monotone nonincreasing in each argument, then the constant function

$$\Theta(t) \equiv \int \alpha d\mu \Big/ \int d\mu$$

minimizes J in M and also in M^*. If $\alpha(t)$ is in the class $M(M^*)$ then clearly $\Theta(t) \equiv \alpha(t)$ minimizes J in $M(M^*)$, even under the less restrictive conditions of F in §§ 2 and 3.

5. Uniqueness theorems. By the relation $\theta(t) \cong \Theta(t)$, we mean that equality holds almost everywhere.

THEOREM 5.1. *Under the conditions of* § 4, *if* $T(z)$ *is strictly convex (that is,* $T'(z)$ *is strictly increasing) on* I_0 *and if* $\theta(t)$ *and* $\Theta(t)$ *both minimize* J *in* $\mathscr{M}$, *then* $\theta(t) \cong \Theta(t)$.

Proof. The set $\{t : \theta(t) \neq \Theta(t)\}$ is the union over all rationals r,

$$\bigcup \{t : \Theta(t) \leqq r < \theta(t)\} \bigcup \{t : \theta(t) \leqq r < \Theta(t)\} \ .$$

It suffices to prove that each of these sets has measure zero. Suppose there is an r_0 such that $\mu\{t : \Theta(t) \leqq r_0 < \theta(t)\} > 0$. Then there exists z_1 such that, for $r_0 < z < z_1$, $\mu\{t : \Theta(t) \leqq z < \theta(t)\} > 0$. As a consequence of Corollary 4.1, $\mu\{t : \Theta(t) = z\} = 0$ for all but a countable set of z, hence

$$M\{t : \Theta(t) \leqq z < \theta(t)\} = M\{t : \Theta(t) < z < \theta(t)\} < z$$

except for a countable set of z between r_0 and z_1. It follows from Lemma 4.4 that

$$J[\theta ; B_1] > J[\Theta ; B_1] \ , \qquad B_1 = \{t : \Theta(t) < \theta(t)\} \ .$$

Similarly, if

$$\mu\{t:\ \theta(t)\leqq r_0<\Theta(t)\}>0\ ,$$

then

$$J[\theta;\ B_2]>J[\Theta;\ B_2]\ ,\qquad B_2=\{t:\ \Theta(t)>\theta(t)\}\ ,$$

and hence $J[\theta]>J[\Theta]$, contradicting the hypothesis that $J[\theta]=J[\Theta]$.

If $T'(z)$ is not strictly increasing on I the above conclusion need not hold. For example if 0 is interior to I and $T(z)=0$ or z according as $z\leqq 0$ or >0, then if $\Theta(t)$ minimizes J, any distinct admissible function $\theta(t)$, agreeing everywhere in sign with $\Theta(t)$, also minimizes J.

The next theorem applies either to problems covered by § 3 or to problems based on an integrand (4.1), and to both the minimum problems in M and in M^*.

If $\Theta(t)$ and $\theta(t)$ are both in M or both in M^*, then

$$\theta_z(t)=\Theta(t)+z[\theta(t)-\Theta(t)]$$

is in M (M^*) for $0\leqq z\leqq 1$. Setting $\mathscr{J}(z)=J[\theta_z]$ we find that

$$\mathscr{J}''(z)=\int(\theta-\Theta)^2F_{vv}(\alpha,\ \theta_z)d\mu\ , \tag{5.1}$$

provided the formal differentiation is valid. Moreover if Θ minimizes J in M (M^*), if $\mathscr{J}'(0)$ exists, and if Taylor's formula is applicable, then

$$J[\theta]-J[\Theta]=\mathscr{J}(1)-\mathscr{J}(0)\geqq\mathscr{J}(1)-\mathscr{J}(0)-\mathscr{J}'(0)=\mathscr{J}''(\zeta)/2\ , \tag{5.2}$$
$$0<\zeta<1\ .$$

THEOREM 5.2. *If* (5.1) *and* (5.2) *are valid, if, for each* z *on the unit interval* $F_{vv}[\alpha(t),\ \theta_z(t)]$ *is positive for almost all* t, *and if* $\theta(t)$ *and* $\Theta(t)$ *both minimize* J *in* M *or both minimize* J *in* M^*, *then* $\theta(t)\cong\Theta(t)$.

The last two theorems apply in particular to integrands given by (4.1) and (4.2) in terms of any one of the convex functions

$$T(z)=z\log z+(1-z)\log(1-z),\qquad z\in I_0=(0,\ 1)\ , \tag{5.3}$$

$$T(z)=z^2\ ,\qquad z\in I_0=(-\infty,\ \infty)\ , \tag{5.4}$$

$$T(z)=z-\log z\ ,\qquad z\in I_0=(0,\ \infty)\ , \tag{5.5}$$

$$T(z)=z\log z\ ,\qquad z\in I_0=(0,\ \infty)\ . \tag{5.6}$$

Applications of these examples in mathematical statistics are discussed in [3]. Each of these examples is covered by the hypotheses of § 3 and of § 4. It is easy to find suitable sufficient conditions for the validity of (5.1) and (5.2) in each case.

REFERENCES

1. Miriam Ayer, H. D. Brunk, G. M. Ewing, W. T. Reid and Edward Silverman, *An empirical distribution function for sampling with incomplete information*, Ann. Math. Stat. **26** (1955), 641–47.
2. H. D. Brunk, G. M. Ewing and W. T. Reid, *The minimum of a certain definite integral suggested by the maximum likelihood estimate of a distribution function*, Bull. Amer. Math. Soc., **60** (1954), 535.
3. H. D. Brunk, *Maximum likelihood estimates of monotone parameters*, Ann. Math. Stat. **26** (1955), 607–16.
4. H. D. Brunk, G. M. Ewing and W. R. Utz, *Some Helly Theorems for monotone functions*, Proc. Amer. Math. Soc. **7** (1956), 776–83.
5. H. Cramér, *Mathematical methods of statistics*, Princeton University Press, 1946.
6. P. R. Halmos, *Measure theory*, Van Nostrand, New York, 1950.
7. E. J. McShane, *Integration*, Princeton University Press, 1944.
8. M. E. Munroe, *Introduction to measure theory and integration*, Addison-Wesley, Cambridge, Mass., 1953.

UNIVERSITY OF MISSOURI