AN INVESTIGATION OF SOLUTIONS TO ALLOCATION PROBLEMS

BERNARD E. CLARK

[redacted]

70

[redacted]

AN INVESTIGATION OF SOLUTIONS TO ALLOCATION PROBLEMS

* * * * *

Bernard E. Clark

[redacted]

AN INVESTIGATION OF

SOLUTIONS TO ALLOCATION PROBLEMS

by

Bernard E. Clark

Captain, United States Marine Corps

Submitted in partial fulfillment of
the requirements for the degree of

MASTER OF SCIENCE
IN
OPERATIONS RESEARCH

United States Naval Postgraduate School
Monterey, California

1 9 6 4

AN INVESTIGATION OF

SOLUTIONS TO ALLOCATION PROBLEMS

by

Bernard E. Clark

This work is accepted as fulfilling

the thesis requirements for the degree of

MASTER OF SCIENCE

IN

OPERATIONS RESEARCH

from the

United States Naval Postgraduate School

ABSTRACT

During the past ten years there has been increasing interest in the solution of allocation-type problems. Published papers on the analytical solution of such problems have in the main been restricted to particular cases.

This paper, while not offering an all-purpose algorithm, addresses the problem in general through the development of theorems giving necessary conditions for a given vector or function to be a solution. These conditions in many cases enable one to deduce the solution.

The paper is presented in two parts. The first addresses the problem of maximizing a sum of n functions over an n-dimensional simplex. The second is concerned with the maximization of the (Lebesgue) integral of a functional over a class of integrable non-negative functions with common L_1 norm.

TABLE OF CONTENTS

1. Introduction.

During the past ten years there has been an increasing interest in the solution of allocation type problems -- problems that are concerned with the distribution of available resources in a manner that will maximize some sort of utility or payoff function. The resources may be defined in terms of men, materials, money, weight, or a myriad of other conceptual products over which some individual has control. The payoff may be in terms of dollars gained, losses to an opponent, expected value of a random variable, probability of the occurrence of an event, or a mixture of these and other factors.

A fairly large literature has grown up around these allocation problems as they apply to many fields of endeavor such as target assignment, theory of search, theory of games, etc. However, most of what has been published has been in the form of particular solutions to particular problems (or classes of problems).

It is the purpose of this paper to set forth certain theorems which appear to be at the heart of nearly all analytical solutions to allocation type problems. The theorems in Section 2 are not by any means original but have been used by mathematicians for years. These and Theorem V of Section 4 were first introduced to the author by Dr. J. M. Danskin of the Institute of Naval Studies in his lectures at the U. S. Naval Postgraduate School in the spring of 1963. Theorems VI, VII and VIII, while apparently never before published (at least not in this form), will be readily seen to be extensions of the basic concepts of Theorems II, III and IV respectively to the case of the Lebesgue integral over E_n.

It is hoped that a study of the theorems and examples of this paper will prepare the reader who is concerned with the analytical solution of allocation type problems to better extend the reasoning of the authors of papers attacking certain special problems, many of which are referenced, to his own particular problem or problem area.

The order of development in this paper will be as follows:

First, four theorems concerning the solution of finite allocation problems will be developed.

Second, two examples of applications of the first four theorems will be worked out and a discussion of appropriate published references will be presented.

Third, four theorems and a lemma concerning the case of the continuum will be developed.

Fourth, two examples and a discussion of published references will be presented.

I would like to express my gratitude for the encouragement, guidance, and inspiration which Professor Kenneth Lucas and Dr. John Danskin provided and without which this paper would not have been possible.

2. The Finite Allocation Problem.

In this section we shall develop four theorems concerning the analytical solution of problems in which one desires to maximize the return from the allocation of a fixed resource among a finite number of activities each of which has associated with it a known return function.

Suppose we have a finite number, n, of activities to which we wish to allocate resources available in a fixed quantity X. Suppose further that associated with each activity there is a function representing our gain upon allocation of resource to that activity. That is $f_i(x)$ is the gain realized from the allocation of an amount x of our resource to the i'th activity. The problem we are concerned with is to maximize our total gain by the proper allocation of our resources. That is we wish to find the n-dimensional vector $\vec{x}^{o} = (x_1^{o}, x_2^{o}, \ldots x_n^{o})$ that maximizes $\sum_{i=1}^{n} f_i(x_i)$ subject to the constraints $x_i \geqq 0$ for all i and $\sum_{i=1}^{n} x_i = X$ where x_i represents the allocation to the i'th activity.

Since the properties derived in the theorems are independent of X, or the x_i's can be normalized, we shall assume henceforth that $X = 1$.

We shall also assume that the functions f_i are twice continuously differentiable (the derivative being taken from the right at 0 and from the left at 1) in what follows. In all but Theorem IV this restriction is excessive. For the reader who encounters problems not meeting the above condition the allowable reduction of the restriction in Theorems I, II and III is readily apparent.

Theorem I. If $\vec{x}^o = (x_1^o, x_2^o, \ldots, x_n^o)$ maximizes

$$\sum_{i=1}^{n} f_i(x_i) \text{ subject to } \sum_{i=1}^{n} x_i = 1 \text{ and } x_i \geqq 0 ,$$

then there is a number λ such that:

$$f_i'(x_i^o) \begin{cases} = \lambda \text{ if } x_i^o > 0 \\ \leqq \lambda \text{ if } x_i^o = 0 \end{cases}$$

(The problem and constraints set forth in the above hypothesis shall be referred to as "the basic problem" throughout the remainder of this section.)

Proof.

Suppose $x_i^o > 0$. Then construct

$$\vec{x}^\epsilon \text{ such that: } \begin{cases} x_i^\epsilon = x_i^o - \epsilon & \text{where } 0 \leqq \epsilon \leqq x_i^o \\ x_j^\epsilon = x_j^o + \epsilon & \text{for same } j \\ x_k^\epsilon = x_k^o & \text{for } k \neq i,j \end{cases}$$

Obviously $\vec{x}^\epsilon$ is admissible. (Throughout this paper the term admissible will refer to a function that satisfies the constraints applied in the statement of the problem to functions which may be considered a solution.)

$$\text{Let } F(\varepsilon) = \sum_{k=1}^{n} f_k(x_k^\varepsilon) = f_i(x_i^o - \varepsilon) + f_j(x_j^o + \varepsilon) + \sum_{k \neq i,j} f_k(x_k^o)$$

Since $\vec{x}^o$ maximizes the above sum $F(\varepsilon)$ achieves its maximum when $\varepsilon = 0$ and so we must have $F'(\varepsilon)\Big|_{\varepsilon=0^+} \leqq 0.$

$$\text{but: } F'(\varepsilon)\Big|_{\varepsilon=0^+} = f_j'(x_j^o) - f_i'(x_i^o) \leqq 0$$

implies that (1) $f_i'(x_i^o) \geqq f_j'(x_j^o)$ for any $j \neq i$

We now observe that if for some $j \neq i$, $x_j^o > 0$, we can by the same method obtain the reverse inequality, thus giving us:

$$f_i'(x_i^o) \geqq f_j'(x_j^o) \quad \text{for} \quad x_i^o > 0 \ \& \ j \neq i$$

$$f_i'(x_i^o) = f_j'(x_j^o) \quad \text{for} \quad x_i^o > 0 \ \& \ x_j^o > 0.$$

Taking λ to be the common value of $f_i'(x_i^o)$ for those x_i^o which are positive the theorem is obtained.

Theorem II. Given the basic problem, if all the f_i are strictly concave ($f_i''(x) < 0$) and $\vec{x}^o = (x_1^o, x_2^o, \ldots x_n^o)$ yields the maximum, then there is a number λ such that $x_i^o > 0$ if and only if $f_i'(0) > \lambda$. If $f_i'(0) > \lambda$, then x_i^o is the unique solution of the equation $f_i'(x_i^o) = \lambda$.

Proof:

Apply Theorem I. If $x_i^o > 0$, we see that $f_i'(x_i^o) = \lambda$ but since $f_i''(x) < 0$, f_i' is strictly decreasing in x so that

$$f_i'(0) > f_i'(x_i^o) = \lambda.$$

On the other hand suppose $f_i'(0) > \lambda$. Theorem I states that if $x_i^o = 0$ then $f_i'(0) \leqq \lambda$ hence $x_i^o > 0$ and the if and only if statement is proven.

To finish the proof, suppose $f_i'(0) > \lambda$. Then $x_i^o > 0$ and from Theorem I $f_i'(x_i^o) = \lambda$. That there is a unique solution to this equation follows from the fact that $f_i'(x)$ is strictly decreasing in x.

Theorem III. Given the basic problem, if all the f_i are strictly convex ($f_i''(x) > 0$) and $\vec{x}^o = (x_1^o, x_2^o, \ldots x_n^o)$ yields the maximum, then there is an i such that $x_i^o = 1$ and $x_j^o = 0$ for $j \neq i$. (Geometrically, the solution lies on a corner of the simplex.)

<u>Proof</u>.

Consider any admissible $\vec{x} = (x_1, x_2, \ldots x_n)$ such that at least two, say x_i and x_j of its components are positive. We know (since $\vec{x}$ is admissible) that $x_i < 1$ and $x_j < 1$.

Take $\varepsilon > 0$ such that:

$$x_i - \varepsilon > 0$$

$$x_j - \varepsilon > 0$$

$$x_i + \varepsilon < 1$$

$$x_j + \varepsilon < 1$$

Define $\vec{x}^1$ and $\vec{x}^2$ as follows:

$$\begin{cases} x_i^1 = x_i - \varepsilon \\ x_j^1 = x_j + \varepsilon \\ x_k^1 = x_k \quad \text{for } k \neq i, j \end{cases}$$

$$\begin{cases} x_i^2 = x_i + \varepsilon \\ x_j^2 = x_j - \varepsilon \\ x_k^2 = x_k \quad \text{for } k \neq i, j \end{cases}$$

Then $\vec{x}^1$ and $\vec{x}^2$ are admissible and furthermore:

$$\frac{\vec{x}^1 + \vec{x}^2}{2} = \vec{x}$$

We may now write:

$$\sum_{i=1}^{n} f_i(x_i) = \sum_{i=1}^{n} f_i\left(\frac{x_i^1 + x_i^2}{2}\right) < \frac{1}{2}\left[\sum_{i=1}^{n} f_i(x_i^1) + \sum_{i=1}^{n} f_i(x_i^2)\right]$$

The inequality holds because each of the f_i is strictly convex and $\vec{x}^1$ and $\vec{x}^2$ differ. Hence $\sum_{i=1}^{n} f_i(x_i)$ must be less than $\sum_{i=1}^{n} f_i(x_i^1)$ or less than $\sum_{i=1}^{n} f_i(x_i^2)$ or both and thus $\vec{x}$ does not yield the maximum. Thus if $\vec{x}^o$ is not a corner of the simplex, $\vec{x}^o$ is not the solution.

Note: The solution for the above case is not necessarily unique. That is, more than one corner of the simplex could yield the same maximum value. However, we can see that all that is necessary to solve the problem for the convex case is to evaluate each f_i at $x_i = 1$ (or X if we have not normalized) and select that activity which yields the highest return.

Theorem IV. Given the basic problem. If $\vec{x}^o = (x_1^o, x_2^o, \ldots x_n^o)$ is the maximizing solution, then $f_i''(x_i^o) \leqq 0$ for all x_i^o which are positive, with at most one exception.

Proof. Suppose we have an admissible $\vec{x}^o$ with

$$x_i^o > 0,\ x_j^o > 0,\ f_i''(x_i^o) > 0 \text{ and } f_j''(x_j^o) > 0\,.$$

Let

$$\sigma = x_i^o + x_j^o$$

$$\varepsilon = \min(x_i^o, x_2^o)$$

Now consider the problem:

(1) Maximize $f_i(x_i) + f_j(x_j)$ subject to

$$x_i \geqq 0,\ x_j \geqq 0 \ \ \&\ \ x_i + x_j = \sigma$$

If $\vec{x}^o$ maximizes the original problem, we see that the pair (x_i^o, x_j^o) must provide the solution to (1).

Since both f_i'' and f_j'' are continuous, we can find a $\delta > 0$ such that $\delta < \varepsilon$ and for $|\eta| \leqq \delta$ we have $f_i''(x_i^o + \eta) > 0$ and $f_j''(x_j^o + \eta) > 0$. (This is the only place we utilize the full power of the constraints placed on the f_i's.)

Let $F(\eta) = f_i(x_i^o + \eta) + f_j(x_j^o - \eta)$. For $|\eta| < \delta$ we have $F''(\eta) > 0$ so that $F(\eta)$ is convex and achieves its maximum on the interval $[-\delta, \delta]$ at $\eta = -\delta$ or at $\eta = +\delta$ but since $(x_i^o + \eta) + (x_j^o - \eta) = \sigma$ this shows us that (x_i^o, x_j^o) can not be the solution to Problem (1). Thus any such $\vec{x}^o$ can not yield the maximum to our original problem.

3. Examples of the Finite Problem.

The theorems of the last section or variations of them have proved useful in a large number of mathematical applications. Included in these "allocation" problems are weapons assignment, solutions to games, and many other problems concerned with the optimum distribution of effort or resources.

It is the purpose of this section to illustrate through simple examples the application of the theorems of Section 2 to the solution of problems. References will be given to several papers and books which treat similar problems by this method and also to several which contain a different approach to the solution to these problems.

Example 1. (This example is based on a problem posed by Dr. John M. Danskin of the Institute of Naval Studies in his lectures on Search Theory during the spring of 1963.)

Suppose in the problem described prior to Theorem I we let the n activities be interpreted as n available targets while x_i represents the number (of a fixed number X) of weapons we allocate to the i'th target. Suppose the probability of destroying the i'th target with x weapons is $1 - e^{-b_i x}$ where $b_i > 0$ is a parameter representing weapon reliability, accuracy, target defenses, etc. Suppose further that the destruction of the i'th target results in a gain to us of $a_i > 0$ units. We now desire to allocate our weapons in such a manner as to maximize our expected payoff (gain) with respect to the probabilities of destruction. Our problem then is:

$$\text{maximize} \qquad \sum_{i=1}^{n} a_i (1-e^{-b_i x_i})$$

Subject to: $x_i \geqq 0$ for all i.

$$\sum_{i=1}^{n} x_i = X$$

by letting $y_i = \frac{x_i}{X}$ we can normalize this problem and rewrite it:

maximize $$\sum_{i=1}^{n} a_i(1-e^{-b_i X y_i})$$

subject to: $y_i \geqq 0$ for all i

$$\sum_{i=1}^{n} y_i = 1$$

where y_i represents the proportion of our total arsenal X that we allocate to the i'th target. In this example we shall determine the solution for both variations to show that they are indeed the same. (We shall adopt the notational scheme of numbering equations for the first problem and using the same numbers with a prime for the analogous equations for the second problem.)

We now have:

(1) $f_i(x_i) = a_i(1-e^{-b_i x_i})$

(2) $f_i'(x_i) = a_i b_i e^{-b_i x_i}$

(3) $f_i''(x_i) = -a_i b_i^2 \, e^{-b_i x_i} < 0$ for all i & $x_i \geqq 0$.

and: (1') $g_i(y_i) = a_i(1-e^{-b_i X y_i})$

(2') $g_i'(y_i) = a_i b_i X e^{-b_i X y_i}$

(3') $g_i''(y_i) = -a_i b_i^2 X^2 e^{-b_i X y_i} < 0$ for all i & $y_i \geqq 0$

We shall now let $c_i = b_i X$.

Since $f_i''(x) < 0$ for all i & x and $g_i''(y) < 0$ for all i & Y, we see that Theorem II applies in both cases. So we calculate:

(4) $f_i'(0) = a_i b_i$

(4') $g_i'(0) = a_i c_i$

and the application of Theorem II gives us:

$$(5)\quad \begin{cases} x_i^o > 0 & \text{iff } a_i b_i > \lambda_x \\ x_i^o = 0 & \text{iff } a_i b_i \leqq \lambda_x \end{cases}$$

$$(5')\quad \begin{cases} y_i^o > 0 & \text{iff } a_i c_i > \lambda_y \\ y_i^o = 0 & \text{iff } a_i c_i \leqq \lambda_y \end{cases}$$

We see from (5) & (5') that if the solutions are equivalent we must have $\lambda_y = X\lambda_x$.

Now applying the second half of Theorem II we obtain:

for all i such that $x_i^o > 0$:

$$f_i'(x_i^o) = a_i b_i \, e^{-b_i x_i^o} = \lambda_x$$

or after multiplying by $1/a_i b_i$ and taking logarithms of both sides:

$$(6)\quad x_i^o = \frac{1}{b_i} \ln \frac{a_i b_i}{\lambda_x}$$

Similarly for i such that $y_i^o > 0$:

$$(6')\quad y_i^o = \frac{1}{c_i} \ln \frac{a_i c_i}{\lambda_y}$$

We now re-label the targets so that $a_1 b_1 \leqq a_2 b_2 \leqq \dots \leqq a_n b_n$. For any λ_x we find its position in this string of inequalities. (Notice that λ_x must be less than $a_n b_n$ since otherwise we would have all $x_i = 0$.) Having chosen λ_x we set

$$(7)\quad \begin{cases} x_i^o = 0 & \text{for } a_i b_i \leq \lambda_x \\ x_i^o = \dfrac{1}{b_i} \ln \dfrac{a_i b_i}{\lambda_x} & \text{for } a_i b_i > \lambda_x \end{cases}$$

We now vary λ_x to obtain:

$$(8)\qquad \sum_{i=1}^{n} x_i^o = \sum_{i \ni a_i b_i > \lambda_x} \frac{1}{b_i} \ln \frac{a_i b_i}{\lambda_x} = X$$

Similarly using the same ordering ($a_i b_i > a_j b_j \Leftrightarrow a_i b_i X = a_i c_i > a_j c_j = a_j b_j X$), we obtain for the second problem:

$$(7')\quad \begin{cases} y_i^o = 0 & \text{for } a_i c_i \leq \lambda_y \\ y_i^o = \dfrac{1}{c_i} \ln \dfrac{a_i c_i}{\lambda_y}, & \text{for } a_i c_i > \lambda_y \end{cases}$$

and we vary λ_y to obtain:

$$(8')\qquad \sum_{i=1}^{n} y_i^o = \sum_{i \ni a_i c_i > \lambda_y} \frac{1}{c_i} \ln \frac{a_i c_i}{\lambda_y} = 1$$

Substituting for c_i into 8' we obtain

$$\sum_{i \ni a_i b_i X > \lambda_y} \frac{1}{b_i X} \ln \frac{a_i b_i X}{\lambda_y} = 1$$

or

$$\sum_{i \ni a_i b_i > \frac{\lambda_y}{X}} \frac{1}{b_i} \ln \frac{a_i b_i}{\lambda_y / X} = X$$

and we see that $\lambda_y = X \lambda_x$ and the two solutions are equivalent.

From this we see that the product $a_i b_i$ is the factor that determines whether or not a target is attacked. So, if a new target arises we can compute the new product ab and if the new ab is less than λ_x, the new target will not be attacked. If the new ab is greater than λ_x, then it will be attacked and a new λ_x must be calculated.

Example 2.

Consider the two-person zero-sum game with payoff Kernal

$K(\vec{x},\vec{y}) = \sum_{i=1}^{n} a_i x_i e^{-b_i y_i}$ where a_i & $b_i > 0$ and x_i & y_i are the coordinates of the n-dimensional vectors $\vec{x}$ and $\vec{y}$ constrained by $x_i, y_i \geqq 0$ for all i and $\sum_{i=1}^{n} x_i = c$; $\sum_{i=1}^{n} y_i = d$. We might consider the a_i's to represent target values; the b_i's target defensibility; x_i the allocation of weapons to target i by player I who desires to maximize $K(\vec{x}, \vec{y})$; and y_i the allocation of defenses to target i by player II who desires to minimize $K(\vec{x}, \vec{y})$. For convenience we shall assume c & d = 1. The existence of a solution to this game, that is a pair $(\vec{x}^o, \vec{y}^o)$ such that $K(\vec{x}^o, \vec{y}^o) \geqq K(\vec{x}, \vec{y}^o)$ for all $\vec{x}$ in the simplex and $K(\vec{x}^o, \vec{y}^o) \leqq K(\vec{x}^o, \vec{y})$ for all $\vec{y}$ in the simplex will not be proved here. Such a pair can be shown to exist.

Our problem then is to find the pair $(\vec{x}^o, \vec{y}^o)$ such that:

(1) $\sum_{i=1}^{n} x_i^o = 1$; $\sum_{i=1}^{n} y_i^o = 1$

(2) $x_i^o \geqq 0$; $y_i^o \geqq 0$ for all i

(3) $K(\vec{x}, \vec{y}^o) \leqq K(\vec{x}^o, \vec{y}^o) \leqq K(\vec{x}^o, \vec{y})$, for all $\vec{x}$ and $\vec{y}$ satisfying (1) and (2).

Looking at (3) in the light of the theorems of the last section we see we have a pair of maximization problems:

(4) Maximize $K(\vec{x}, \vec{y}^o)$ subject to $\sum_{i=1}^{n} x_i = 1$ and $x_i \geqq 0$ for all i.

(5) Maximize $-K(\vec{x}^o, \vec{y})$ subject to $\sum_{i=1}^{n} y_i = 1$ and $y_i \geqq 0$ for all i.

Upon differentiating $K(\vec{x}, \vec{y}^o) = \sum_{i=1}^{n} a_i x_i e^{-b_i y_i^o}$ by x_i and applying Theorem I, we see that there must be number λ such that:

$$(6) \quad a_i e^{-b_i y_i^o} \begin{cases} = \lambda & \text{if } x_i^o > 0 \\ \leqq \lambda & \text{if } x_i^o = 0 \end{cases}$$

Similarly from $-K(\vec{x}^o, \vec{y})$ we see there must be a number k such that:

$$(7) \quad a_i b_i x_i^o e^{-b_i y_i^o} \begin{cases} = k & \text{if } y_i^o > 0 \\ \leqq k & \text{if } y_i^o = 0 \end{cases}$$

From an examination of (6) and (7) we can determine the following:

(a) <u>$k > 0$</u> ; There must be at least one i for which $x_i > 0$. This with (7) gives $k > 0$.

(b) <u>$y_i^o > 0 \Rightarrow x_i^o > 0$</u> ; If $y_i^o > 0$, then $a_i b_i x_i^o e^{-b_i y_i^o} = k > 0 \therefore x_i > 0$.

(c) <u>$a_i < \lambda \Rightarrow x_i^o = 0 \Rightarrow y_i^o = 0$</u> ; The first implication follows directly from (6). The second is a consequence of (b).

(d) $\underline{a_i = \lambda \Rightarrow y_i = 0}$; If $a_i = \lambda$, then $y_i^o > 0 \Rightarrow x_i^o = 0$ by

(6); but, by (c) $x_i^o = 0 \Rightarrow y_i^o = 0$.

(e) $\underline{a_i = \lambda \Rightarrow x_i^o \leqq \frac{k}{a_i b_i}}$; Substitute $y_i^o = 0$ into (7)

(f) $\underline{a_i > \lambda \Rightarrow y_i^o > 0 \Rightarrow x_i^o > 0}$; This follows directly from

(6) and (b).

Now to determine a solution we proceed as follows:

We see from (c), (d), (e) and (f) that we need only consider those i for which $a_i \geqq \lambda$.

From (6): $a_i > \lambda \Rightarrow a_i e^{-b_i y_i^o} = \lambda$

Substituting this into (7) we obtain:

(8) $b_i \lambda x_i^o = k \Rightarrow x_i^o = \frac{k}{b_i \lambda}$, for $a_i > \lambda$.

From (e) we obtain:

(9) $x_i^o \leqq \frac{k}{b_i a_i} = \frac{k}{b_i \lambda}$, for $a_i = \lambda$.

If we let $Z = \sum_{i \ni a_i = \lambda} x_i^o \leqq \frac{k}{\lambda} \sum_{i \ni a_i = \lambda} \frac{1}{b_i}$,

then $\sum_{i=1}^{n} x_i^o = \sum_{i \ni a_i > \lambda} x_i^o + \sum_{i \ni a_i = \lambda} x_i^o = \frac{k}{\lambda} \sum_{i \ni a_i > \lambda} \frac{1}{b_i} + Z = 1.$

Now if we define $B_o = \sum_{i \ni a_i > \lambda} \frac{1}{b_i}$

$B_1 = \sum_{i \ni a_i \geqslant \lambda} \frac{1}{b_i}$

then for $Z = 0$ (that is $x_i^o = 0$ for $a_i = \lambda$) we have:

$$\sum_{i=1}^{n} x_i^o = 1 = \frac{k}{\lambda} B_o \Rightarrow \underline{k = \frac{\lambda}{B_o}}$$

However, if we set $x_i^o = \frac{k}{\lambda b_i}$ when $a_i = \lambda$, we have:

$$\sum_{i=1}^{n} x_i^o = \frac{k}{\lambda} \sum_{i \ni a_i \geq 0} \frac{1}{b_i} = \frac{k}{\lambda} B_1 = 1 \Rightarrow \underline{k = \frac{\lambda}{B_1}}$$

We can thus see that we have

$$(10) \quad \frac{k}{\lambda} B_o \leqq \sum_{i=1}^{n} x_i^o = 1 \leqq \frac{k}{\lambda} B_1$$

So we have:

$$(11) \quad \frac{\lambda}{B_1} \leqq k \leqq \frac{\lambda}{B_o}$$

Using (10) we see that for any value of k satisfying (11) we can find a number μ with $0 \leqq \mu \leqq 1$ such that:

$$(12) \quad \frac{k}{\lambda} B_o + \mu\left(\frac{k}{\lambda} B_1 - \frac{k}{\lambda} B_o\right) = 1.$$

Recalling the definition of B_o and B_1 (12) becomes:

$$\frac{k}{\lambda} \sum_{i \ni a_i > \lambda} \frac{1}{b_i} + \mu \frac{k}{\lambda} \sum_{i \ni a_i = \lambda} \frac{1}{b_i} = 1.$$

So by setting $x_i^o = \frac{\mu k}{\lambda b_i}$, for $a_i = \lambda$; $x_i^o = \frac{k}{\lambda b_i}$, for $a_i > \lambda$; and $x_i^o = 0$ for $a_i < \lambda$, we obtain a solution for player I.

For player II we have $y_i^o > 0$ only if $a_i > \lambda$. But for $a_i > \lambda$ we have

$$x_i^o = \frac{k}{\lambda b_i}$$

So we get from (7):

$$a_i b_i \frac{k}{\lambda b_i} e^{-b_i y_i^o} = k \quad \text{for } a_i > \lambda$$

or $\frac{a_i}{\lambda} e^{-b_i y_i^o} = 1.$ Multiplying by $\frac{\lambda}{a_i}$ and taking logarithms of both sides we obtain

$$-b_i y_i^o = \ln \frac{\lambda}{a_i} \quad \text{or} \quad y_i^o = \frac{1}{b_i} \ln \frac{a_i}{\lambda}$$

Thus we can obtain λ as the solution to the equation:

$$(13) \qquad \sum_{i=1}^{n} y_i = \sum_{i \ni a_i > \lambda} \frac{1}{b_i} \ln \frac{a_i}{\lambda} = 1$$

in the same manner as in Example 1.

Thus we can express the final solution to the game:

$$\text{for } a_i > \lambda \quad \begin{cases} x_i^o = k/\lambda b_i \\ y_i^o = \frac{1}{b_i} \ln \frac{a_i}{\lambda} \end{cases}$$

$$\text{for } a_i = \lambda \quad \begin{cases} x_i^o = \frac{\mu k}{\lambda b_i} \\ y_i^o = 0 \end{cases}$$

$$\text{for } a_i < \lambda \quad \begin{cases} x_i^o = 0 \\ y_i^o = 0 \end{cases}$$

where: (a) λ satisfies $\sum_{i \ni a_i > \lambda} \frac{1}{b_i} \ln \frac{a_i}{\lambda} = 1$

(b) k is not unique but any number satisfying:

$$\frac{k}{\lambda} \sum_{i \ni a_i > \lambda} \frac{1}{b_i} \leq 1 \leq \frac{k}{\lambda} \sum_{i \ni a_i \geq \lambda} \frac{1}{b_i}$$

(c) μ is the number for the chosen value of k such that:

$$\frac{k}{\lambda} \sum_{i \ni a_i > \lambda} \frac{1}{b_i} + \mu \frac{k}{\lambda} \sum_{i \ni a_i = \lambda} \frac{1}{b_i} = 1$$

Substituting the solution $(\vec{x}^o, \vec{y}^o)$ into the expression for K we see that the "value of the game" is:

$$\underline{\underline{v_o}} = K(\vec{x^o},\vec{y^o}) = \sum_{i=1}^{n} a_i x_i^o e^{-b_i y_i^o} = \sum_{i=1}^{n} \lambda x_i^o = \lambda .$$

The preceding examples are highly simplified and are meant to serve as an illustration of the method of applying theorems I through IV, not as examples of applications. For a much more sophisticated and realistic application of the theorems the reader is referred to the papers by J. M. Danskin on Convoy Routing [6] and Theory of Reconnaissance [5].

For an appraisal of the earlier work done in this field with much more severe restrictions on the nature of the problem the reader is referred to the papers by Charnes and Cooper [2] and deGuenin [7].

The reader has probably observed that the solution obtained by application of Theorems I through IV will not, in general, provide integral assignments. Thus we may be faced with the problem of interpreting the meaning of assigning 0.7 bombs to a target. Further, in some cases the solution to an allocation problem is very unstable (this is the case when the payoff functions $f_i(x)$ are S-shaped, first convex and then concave as x increases). In cases of this type our method may not give us any useful information. The paper by Karush [11] addresses this problem and contains some very interesting examples pointing up the danger of attempting to deduce an integer solution from the analytical solution (these examples are readily apparent upon reading the article although not specifically referred to therein). Bellman's book [1] provides a fine computer oriented approach to the solution of allocation problems by the generation of sequences of functional equations. Other approaches to special types of allocation problems can be found in the papers by den Broeder, Ellison & Emerling [8], Lemus & David [13], Manne [14] and Piccariello [16].

4. The Case of the Continuum.

In this section we shall consider the analog of the problem of Section 2 in which we are allocating our resource over a continuum of activities. Examples might be the distribution of searching effort (effort spent on a point represented by the velocity at which the point is traversed) over an area or the positioning of barrier forces in order to protect an installation. We will develop four theorems analogous to those of Section 2.

The following lemma and four theorems are based on Lebesgue measure and integrals. In what follows we shall let:

m be Lebesgue measure

W be an integrable function

g be a measurable, non-negative function

X be the class of all integrable, non-negative functions bounded above by g and such that $\int x = c$ whenever $x \in X$.

All of the above are defined on E_n.

Lemma I. $x^o \in X$ maximizes $\int Wx$ over all $x \in X$ if and only if there exists a number λ such that for almost all t:

$$(1) \quad W(t) \begin{cases} \geqq \lambda & \text{if } x^o(t) = g(t) \\ = \lambda & \text{if } 0 < x^o(t) < g(t) \\ \leqq \lambda & \text{if } x^o(t) = 0 \end{cases}$$

Proof:

a. Sufficiency.

Suppose $x^o(t)$ satisfies (1). We must show that:

$$\int W(t) \left[x^o(t) - x(t) \right] \geqq 0 \quad \text{for all } x \in X.$$

We can write:

$$(2) \quad \int W(t)\left[x^o(t) - x(t)\right] = \int\left[W(t) - \lambda\right]\left[x^o(t) - x(t)\right]$$

since $\int \lambda\left[x^o(t) - x(t)\right] = \lambda C - \lambda C = 0$

We now decompose E_n into three disjoint measurable sets:

$$S_1 = \{t : W(t) < \lambda\}$$

$$S_2 = \{t : W(t) = \lambda\}$$

$$S_3 = \{t : W(t) > \lambda\}$$

Then (2) becomes:

$$\int_{S_1} \left[W(t) - \lambda\right]\left[x^o(t) - x(t)\right] + \int_{S_2} \left[W(t) - \lambda\right]\left[x^o(t) - x(t)\right]$$

$$+ \int_{S_3} \left[W(t) - \lambda\right]\left[x^o(t) - x(t)\right]$$

We observe that:

On S_1 : $W(t) - \lambda < 0$ and $x^o(t) = 0$ Thus $x^o(t) - x(t) \leq 0$. So, the integrand and consequently the integral is non-negative.

On S_2 : $W(t) - \lambda = 0$ so the integral is zero.

On S_3 : $W(t) - \lambda > 0$ and $x^o(t) = g(t)$ thus $x^o(t) - x(t) \geq 0$ and so, as on S_1 the integral is non-negative.

This proves sufficiency.

b. Necessity.

Suppose $x^o(t)$ maximizes the integral subject to the constraints of X.

We note that Lebesgue measurability is equivalent to approximate continuity almost everywhere. (See Munroe [15], p. 292 or Saks [17], p. 132.)

Let t_o and t_1 be two distinct points of approximate continuity of x^o, W, and g such that $x^o(t_o) < g(t_o)$ and $x^o(t_1) > 0$ (If no two such points exist, the problem is degenerate.)

The following property of a function in a neighborhood of a point of approximate continuity is vital to the remainder of the proof.

If x is approximately continuous at t_o, then for any $\varepsilon > 0$ and $\eta > 0$ there exists $\delta > 0$ such that for any interval I containing t_o such that $m(I) < \delta$ we have:

$$m(E_o^x \cap I) > (1 - \eta) m(I)$$

where $E_o^x = \{t : |x(t) - x(t_o)| < \varepsilon\}$

Now let us choose $\eta = 1/6$, an arbitrary $\sigma > 0$, and $\varepsilon > 0$ such that: $x^o(t_o) + 4\varepsilon < g(t_o)$ and $x^o(t_1) - 3\varepsilon > 0$.

Define:

$$E_o^x = \{t : x^o(t) < g(t_o) - 3\varepsilon\}$$

$$E_o^W = \{t : W(t) > W(t_o) - \sigma\}$$

$$E_o^g = \{t : g(t) > g(t_o) - \varepsilon\}$$

$$E_1^x = \{t : x^o(t) > 2\varepsilon\}$$

$$E_1^W = \{t : W(t) < W(t_1) + \sigma\}$$

$$E_1^g = \{t : g(t) > g(t_1) - \varepsilon\}$$

Then we can find $\delta_o^x, \delta_o^W, \delta_o^g, \delta_1^x, \delta_1^W, \delta_1^g$ all > 0 such that for:

$$t_o \in I_o^x \ \& \ m(I_o^x) < \delta_o^x, \ m(E_o^x \cap I_o^x) > 5/6 \ m(I_o^x)$$

$$t_o \in I_o^W \ \& \ m(I_o^W) < \delta_o^W, \ m(E_o^W \cap I_o^W) > 5/6 \ m(I_o^W)$$

$$t_o \in I_o^g \ \& \ m(I_o^g) < \delta_o^g, \ m(E_o^g \cap I_o^g) > 5/6 \ m(I_o^g)$$

$$t_1 \in I_1^x \ \& \ m(I_1^x) < \delta_1^x, \ m(E_1^x \cap I_1^x) > 5/6 \ m(I_1^x)$$

$$t_1 \in I_1^W \ \& \ m(I_1^W) < \delta_1^W, \ m(E_1^W \cap I_1^W) > 5/6 \ m(I_1^W)$$

$$t_1 \in I_1^g \ \& \ m(I_1^g) < \delta_1^g, \ m(E_1^g \cap I_1^g) > 5/6 \ m(I_1^g)$$

We now take δ such that:

$$0 < \delta < \min\,(\delta_o^x, \delta_o^W, \delta_o^g, \delta_1^x, \delta_1^W, \delta_1^g).$$

and I_o and I_1 such that $t_o \in I_o$ and $t_1 \in I_1$, $I_o \cap I_1 = \phi$ and:

(3) $\quad 0 < m(I_o) = m(I_1) \leq \delta$

Let:

(4) $\quad E_o = E_o^x \cap E_o^W \cap E_o^g \cap I_o$

(5) $\quad E_1 = E_1^x \cap E_1^W \cap E_1^g \cap I_1$

then:

$$\begin{aligned} m(E_o) &= m(I_o) - m(I_o - E_o) = m(I_o) - m(I_o - [E_o^x \cap E_o^W \cap E_o^g]) \\ &= m(I_o) - m((I_o - E_o^x) \cup (I_o - E_o^W) \cup (I_o - E_o^g)) \\ &\geqslant m(I_o) - m(I_o - E_o^x) - m(I_o - E_o^W) - m(I_o - E_o^g) \\ &> m(I_o) - \tfrac{1}{6} m(I_o) - \tfrac{1}{6} m(I_o) - \tfrac{1}{6} m(I_o) \\ &= 1/2 \ m(I_o). \end{aligned}$$

Thus we have

(6) $\quad m(E_o) > 1/2 \ m(I_o)$

and by the same reasoning we have

(7) $\quad m(E_1) > 1/2\, m(I_1)$

Using (3), (4), (5), (6) and (7) we obtain:

$$1/2\, m(E_o) \leq 1/2\, m(I_o) = 1/2\, m(I_1) < m(E_1) \leq m(I_1) = m(I_o) < 2\, m(E_o)$$

or $\quad 1/2\, m(E_o) < m(E_1) < 2\, m(E_o)$

and since $m(E_o) > 0$ we can write

$$1/2 < \frac{m(E_1)}{m(E_o)} < 2 .$$

Next define

$$x^1(t) = \begin{cases} x^o(t) + \varepsilon \dfrac{m(E_1)}{m(E_o)} & \text{on } E_o \\ x^o(t) - \varepsilon & \text{on } E_1 \\ x^o(t) & \text{elsewhere.} \end{cases}$$

We observe

$$\text{on } E_o: \quad x^1(t) = x^o(t) + \varepsilon \frac{m(E_1)}{m\, E_o} < x^o(t) + 2\varepsilon$$

$$< g(t_o) - \varepsilon < g(t)$$

$$\text{on } E_1: \quad x^1(t) = x^o(t) - \varepsilon > \varepsilon > 0$$

$$x^1(t) = x^o(t) - \varepsilon < x^o(t) < g(t)$$

$$\text{and} \quad \int x^1 = \int x^o + \int_{E_o} \frac{m(E_1)}{m(E_o)} \varepsilon - \int_{E_1} \varepsilon = \int x^o + \varepsilon\, m(E_1) - \varepsilon\, m(E_1) = C.$$

so $\quad x^1 \in X$.

Now since $x^o(t)$ maximizes our integral we can write:

$$\int W(t)\, x^o(t) \geqslant \int W(t)\, x^1(t)$$

and using the definition of $x^1(t)$ this becomes

$$(8) \qquad \varepsilon\int_{E_1} W(t) \geqslant \frac{m(E_1)}{m(E_o)}\,\varepsilon\int_{E_o} W(t)$$

but E_o and E_1 were chosen such that

$$W(t) > W(t_o) - \sigma \text{ on } E_o$$

$$W(t) < W(t_1) + \sigma \text{ on } E_1$$

so we have upon dividing both sides of (8) by ε :

$$(W(t_1)+\sigma)\, m(E_1) \geqslant \int_{E_1} W(t) \geqslant \frac{m(E_1)}{m(E_o)}\int_{E_o} W(t) \geqslant (W(t_o) - \sigma)\, m(E_1)$$

Upon dividing through by $m(E_1)$ which was chosen to be non zero we obtain

$W(t_1)+\sigma \geqslant W(t_o) - \sigma$ where σ is arbitrary.

Thus $W(t_1) \geqslant W(t_o)$.

Since all we required was that $x^o(t_1) > 0$ and $x^o(t_o) < g(t_o)$ we can see that if also $x^o(t_o) > 0$ and $x^o(t_1) < g(t_1)$ we can obtain the reverse inequality in the same manner, thus obtaining $W(t_o) = W(t_1)$. Taking λ to be this common value we obtain the conditions:

$$W(t) \geqslant \lambda \text{ if } x^o(t) = g(t)$$

$$W(t) = \lambda \text{ if } 0 < x^o(t) < g(t)$$

$$W(t) \leq \lambda \text{ if } x^o(t) = 0$$

These conditions hold for all points at which x, W, and g are approximately continuous and hence a.e. and the necessity is proven.

With this lemma we are now ready to prove our first theorem.

Theorem V. Given g and X as before and a function f defined on E_{n+1} (or at least that portion having non-negative first coordinate) such that $f_x(x(t),t)$ (the first partial derivative by the first coordinate) is an integrable function of t for all $x \in X$. If $x^{o}(t)$ maximizes $\int f(x(t),t)$ over $x \in X$, then there is a number λ such that for almost all t:

$$f_x(x^{o}(t),t) \begin{cases} \geq \lambda & \text{if } x^{o}(t) = g(t) \\ = \lambda & \text{if } 0 < x^{o}(t) < g(t) \\ \leq \lambda & \text{if } x^{o}(t) = 0 \end{cases}$$

(The problem and constraints set forth in the above hypothesis shall be referred to as "the Basic Problem" throughout the remainder of this section.)

Proof:

Suppose $x^{o}(t)$ maximizes the integral subject to the constraints on X. Then let $x^{1}(t)$ be any other function from X.

Define

$$x^{\theta}(t) = \theta x^{1}(t) + (1-\theta)x^{o}(t)$$

for $0 \leq \theta \leq 1$

Obviously $x^{\theta}(t) \in X$ for all $\theta \in [0,1]$.

Let $F(\theta) = \int f(x^{\theta}(t),t) = \int f(\theta x^{1} + (1-\theta)\, x^{o}, t)$

Now since x^{o} maximizes the integral, $F(\theta)$ has its maximum at $\theta = 0$ and since f_x exists $dF/d\theta$ exists for $\theta \in [0,1]$ (we take the derivative from the right at 0 and from the left at 1). So $\left.\frac{dF}{d\theta}\right|_{\theta=0}$ must be negative. But:

$$\frac{dF}{d\theta} = \int f_x(\theta x^{1} + (1-\theta)x^{o},t)\,[x^{1} - x^{o}]$$

So we have:

$$\left.\frac{dF}{d\theta}\right|_{\theta=0} = \int f_x(x^{o}(t),t)\,[x^{1} - x^{o}] \leq 0$$

for all $x^{1} \in X$

But this is the problem stated in the lemma with $W(t) = f_x(x^o(t), t)$. So by application of the lemma the theorem is obtained.

When the function $f(x,t)$ has convexity or concavity properties we can determine even more about the nature of the maximizing solution as the following theorems will show.

For simplicity we shall consider two functions that differ on a set of measure zero to be equal in the following three theorems:

<u>Theorem VI</u>. Given the basic problem, if $f_x(x,t)$ is strictly decreasing in x for all t, then there exists a number λ such that:

$x^o(t) > 0$ if and only if $f_x(0,t) > \lambda$

where $x^o(t)$ is the maximizing solution.

<u>Proof</u>.

a. $x^o(t) > 0 \Rightarrow f_x(0,t) > \lambda$:

Since $f_x(x,t)$ is strictly decreasing we have for $x^o(t) > 0$

(1) $f_x(0,t) > f_x(x^o(t),t)$

From Theorem I we have

(2) $x^o(t) > 0 \Rightarrow f_x(x^o(t),t) \geqslant \lambda$

Combining (1) and (2) we obtain

$$f_x(0,t) > \lambda$$

b. $f_x(0,t) > \lambda \Rightarrow x^o(t) > 0$:

or equivalently $x^o(t) = 0 \Rightarrow f_x(0,t) \leq \lambda$

but this is the conclusion of Theorem V.

<u>Theorem VII</u>. Given the basic problem, if $f_x(x,t)$ is strictly increasing in x for all t and $x^o(t)$ is a maximizing solution, then

$m\left(\{t : 0 < x^o(t) < g(t)\}\right) = 0.$

Proof: Suppose we have an $x^o(t)$ such that $m(S) = m(\{t : 0 < x^o(t) < y(t)\}) > 0$. Let t_1 & t_2 be distinct points of S such that x^o and g are approximately continuous at t_1 & t_2. Recalling the variational technique employed in Lemma I we take $\eta = 1/4$ and $\varepsilon > 0$ such that:

$$x^o(t_1) + 4\varepsilon < g(t_1)$$

$$x^o(t_2) + 4\varepsilon < g(t_2)$$

$$x^o(t_1) > 3\varepsilon$$

$$x^o(t_2) > 3\varepsilon$$

Define:

$$E_1^x = \{t : |x^o(t) - x^o(t_1)| < \varepsilon\}$$

$$E_1^g = \{t : g(t) > g(t_1) - \varepsilon\}$$

$$E_2^x = \{t : |x^o(t) - x^o(t_2)| < \varepsilon\}$$

$$E_2^g = \{t : g(t) > g(t_2) - \varepsilon\}$$

We then find $\delta_1^x, \delta_1^g, \delta_2^x, \delta_2^g > 0$ such that:

for $t_1 \in I_1^x$ & $m(I_1^x) < \delta_1^x$, $m(E_1^x \cap I_1^x) > 3/4\ m(I_1^x)$

$t_1 \in I_1^g$ & $m(I_1^g) < \delta_1^y$, $m(E_1^g \cap I_1^g) > 3/4\ m(I_1^g)$

$t_2 \in I_2^x$ & $m(I_2^x) < \delta_2^x$, $m(E_2^x \cap I_2^x) > 3/4\ m(I_2^x)$

$t_2 \in I_2^g$ & $m(I_2^g) < \delta_2^g$, $m(E_2^g \cap I_2^g) > 3/4\ m(I_2^g)$

Now select $\delta > 0$ such that $\delta < \min(\delta_1^x, \delta_1^g, \delta_2^x, \delta_2^y)$ and I_1 & I_2 such that $t_1 \in I_1$, $t_2 \in I_2$,

$I_1 \cap I_2 = \phi$ and $m(I_1) = m(I_2) = \delta$.

Let: $E_1 = E_1^x \cap E_1^g \cap I_1$

$E_2 = E_2^x \cap E_2^g \cap I_2$

Now: $m(E_1) = m(I_1) - m(I_1 - E_1) = m(I_1) - m(I_1 - (E_1^x \cap E_1^g))$

$$= m(I_1) - m([I_1 - E_1^x] \cup [I_1 - E_1^g])$$

$$\geq m(I_1) - m(I_1 - E_1^x) - m(I_1 - E_1^g)$$

$$> m(I_1) - 1/4\, m(I_1) - 1/4\, m(I_1) = 1/2\, m(I_1)$$

Similarly we obtain $m(E_2) > 1/2\, m(I_2)$ and as before we obtain the inequality:

$$1/2 < \frac{m(E_1)}{m(E_2)} < 2 .$$

Now define:

$$x^1(t) = \begin{cases} x^o(t) + \varepsilon \ , & \text{on } E_1 \\ x^o(t) - \dfrac{m(E_1)}{m(E_2)}\varepsilon \ , & \text{on } E_2 \\ x^o(t) \ , & \text{on } -(E_1 \cup E_2) \end{cases}$$

$$x^2(t) = \begin{cases} x^o(t) - \varepsilon \ , & \text{on } E_1 \\ x^o(t) + \dfrac{m(E_1)}{m(E_2)}\varepsilon \ , & \text{on } E_2 \\ x^o(t) \ , & \text{on } -(E_1 \cup E_2) \end{cases}$$

We observe:

on E_1: $\quad x^1(t) = x^o(t) + \varepsilon > x^o(t) \geq 0$

$$x^1(t) = x^o(t) + \varepsilon < x^o(t_1) + 2\varepsilon < g(t_1) - \varepsilon < g(t)$$

$$x^2(t) = x^o(t) - \varepsilon > x^o(t_2) - 2\varepsilon > 0$$

$$x^2(t) = x^o(t) - \varepsilon < x^o(t) < g(t)$$

on E_2: $x^1(t) = x^o(t) - \frac{m(E_1)}{m(E_2)}\varepsilon > x^o(t) - 2\varepsilon > x^o(t_1) - 3\varepsilon > 0$

$$x^1(t) = x^o(t) - \frac{m(E_1)}{m(E_2)}\varepsilon < x^o(t) \leq g(t)$$

$$x^2(t) = x^o(t) + \frac{m(E_1)}{m(E_2)}\varepsilon > x^o(t) \geqslant 0$$

$$x^2(t) = x^o(t) + \frac{m(E_1)}{m(E_2)}\varepsilon < x^o(t) + 2\varepsilon < x^o(t_1) + 3\varepsilon < g(t_1) - \varepsilon < g(t).$$

$$\int x^1 = \int x^o + \int_{E_1}\varepsilon - \frac{m(E_1)}{m(E_2)}\int_{E_2}\varepsilon = \int x^o$$

$$\int x^2 = \int x^o + \frac{m(E_1)}{m(E_2)}\int_{E_2}\varepsilon - \int_{E_1}\varepsilon = \int x^o$$

Thus x^1 & $x^2 \in X$.

We now observe that $x^o(t) = \dfrac{x^1(t) + x^2(t)}{2}$

but, since f is convex ($f_x(x,t)$ is strictly increasing in x) we have:

$f(x^o(t),t) < 1/2\ f(x^1(t),t) + 1/2\ f(x^2(t),t)$ for all t and thus

$$\int f(x^o(t),t) < 1/2 \int f(x^1(t),t) + 1/2 \int f(x^2(t),t)$$

so that either x^1 or x^2 (or both) yields an integral greater than x^o. Thus any x^o such that $m(\{x : 0 < x^o(t) < g(t)\}) > 0$ can not be the maximizing solution and the theorem is proven.

<u>Theorem VIII</u>. Given the basic problem with $f(x,t)$ twice continuously differentiable by x. If $x^o(t)$ is the maximizing solution and we let

$S_1(x^o) = \{t : f_{xx}(x^o,t) > 0\}$

$S_2(x^o) = \{t : 0 < x^o(t) < g(t)\}$

and $S(x^o) = S_1(x^o) \cap S_2(x^o)$, then $m(S(x^o)) = 0$.

Proof: Suppose we have an $x^o(t)$ such that $m(S(x^o)) > 0$.

On $S(x^o)$ choose $\mathcal{E}_1(t) > 0$ such that:

$$x^o(t) - \mathcal{E}_1(t) > 0$$

$$x^o(t) + \mathcal{E}_1(t) < g(t)$$

and $\mathcal{E}_2(t) > 0$ such that:

for $x^o(t) - \mathcal{E}_2(t) \leq x(t) \leq x^o(t) + \mathcal{E}_2(t)$, $f_{xx}(x,t) > 0$.

This can be done since f_{xx} is continuous.

Let $\mathcal{E}(t) = \min(\mathcal{E}_1(t), \mathcal{E}_2(t))$.

Now consider the problem:

(1) Maximize $\int_{S(x^o)} f(x,t)$ subject to:

(2) $\int_{S(x^o)} x(t) = \int_{S(x^o)} x^o(t)$

(3) $x^o(t) - \mathcal{E}(t) \leq x(t) \leq x^o(t) + \mathcal{E}(t)$

If $x^o(t)$ maximizes the original integral, then if we let $x*(t) = x^o(t)$ restricted to $S(x^o)$, $x^*(t)$ must provide a maximizing solution to this new problem.

If we now let $y(t) = x(t) - x^o(t) + \mathcal{E}(t)$, we obtain:

(1') Maximize $\int_{S(x^o)} f(y(t) + x^o(t) - \mathcal{E}(t), t)$

subject to:

(2') $\int_{S(x^o)} y(t) = \int_{S(x^o)} \mathcal{E}(t)$

(3') $0 \leq y(t) \leq 2\mathcal{E}(t)$

We note that $f(y(t) + x^o(t) - \mathcal{E}(t), t)$ can be written $h(y(t), t)$ since x^o and $\mathcal{E}$ are fixed functions of t. We further see that $h_y(y(t),t) = f_x(y(t),t)$ and $h_{yy}(y(t),t) = f_{xx}(y(t),t)$ so we have

$$h_{yy}(y(t),t) > 0 \quad \text{on} \quad S(x^o)$$

and by Theorem VII for y^o to be a maximizing solution we must have $y^o(t) = 0$ or $2\mathcal{E}$ for almost all t in $S(x^o)$. But;

$y^o(t) = 0$ implies $x^*(t) = x^o(t) + \mathcal{E}(t)$

$y^o(t) = 2\mathcal{E}$ implies $x^*(t) = x^o(t) - \mathcal{E}(t)$

Thus any $x^o(t)$ such that $m(S(x^o)) > 0$ can not be the maximizing solution to the original problem and the proof is complete.

5. Examples of the Continuous Case.

In this section we shall illustrate some applications of the theorems of the last section by means of examples and shall cite references for further reading for the interested reader. We shall use as our primary example the problem of search as formulated by B. O. Koopman [12]. A study of Koopman's article and his indicated generalizations will make readily apparent how much less restrictive the theorems of Section 4 are than similar theorems from the calculus of variations. We shall also show the Neyman-Pearson lemma of statistics to be a special case of Lemma I when the selection of a most powerful test of a hypothesis is looked upon as an allocation problem.

Example 1. The Koopman Search Problem.

Suppose we are searching for an object which may be represented as a point on the real line. We desire to apply a fixed amount of search effort along the line in such a manner as to maximize the probability of detecting the object. Detection can be represented as a Bernoulli trial with a parameter dependent on location, X, and effort ϕ. More specifically the conditional probability of detection with effort $\phi(x)$ given the location of the object is in fact x is assumed to be $(1 - e^{-\phi(x)})$. (This assumes that all portions of the line are equally susceptible to search, which assumption could be deleted.) Assuming X has the marginal density function p(x), we have the unconditional probability of detection $P(\phi)$ given by:

$$P(\phi) = \int p(x)(1 - e^{-\phi(x)})\, dx$$

So our problem is to find $\phi^o(x)$ that will maximize $P(\phi)$ subject to:

$$(2) \quad \begin{cases} \phi(x) \geq 0 \\ \int \phi(x)\, dx = \Phi \end{cases}$$

We now observe that in the terminology of Section 4 we have:

$$(3) \quad f(\phi, x) = p(x)(1 - e^{-\phi(x)})$$

$$(4) \quad f_{\phi}(\phi, x) = p(x)e^{-\phi(x)}$$

$$(5) \quad f_{\phi\phi}(\phi, x) = -p(x)\, e^{-\phi(x)} \leq 0 \quad \text{for } \forall\, x$$

From (5) we see that Theorem VI applies wherever $p(x) > 0$ and where $p(x) = 0, \phi(x)$ obviously will be zero, so we can assume Theorem VI to hold throughout. So we know that there is a number λ such that:

$$(6) \quad \phi^{o}(x) > 0 \quad \text{iff} \quad p(x) > \lambda$$

Since we have no upper bound on $\phi(x)$, $(g(x) \equiv \infty)$ we also have from Theorem V and (6):

$$(7) \quad \begin{cases} p(x)\, e^{-\phi^{o}(x)} = \lambda & \text{for } p(x) > \lambda \\ \phi^{o}(x) = 0 & \text{for } P(x) \leq \lambda \end{cases}$$

Taking logarithms of both sides of the upper equation in (7) we obtain:

$$\phi^{o}(x) = \begin{cases} \ln \dfrac{p(x)}{\lambda} & \text{for } p(x) > \lambda \\ 0 & \text{for } p(x) \leq \lambda \end{cases}$$

from (2):

$$\Phi = \int \phi^{o}(x)\, dx = \int_{x \ni p(x) > \lambda} \ln \frac{p(x)}{\lambda}\, dx = \int_{x \ni p(x) > \lambda} (\ln p(x) - \ln \lambda)\, dx$$

Thus we can use the following graphical method to find the solution $\phi^{o}(x)$.

Step (1). Graph ln(p(x)) for all x at which $p(x) > 0$ this graph may look like Fig. 1.

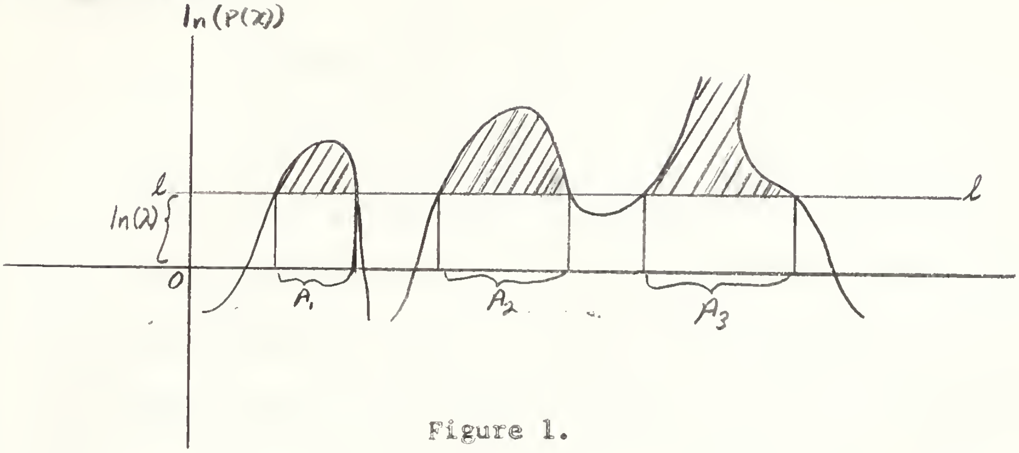

Figure 1.

Step (2). Draw a horizontal line ℓ parallel to the x axis and move it up and down until the area (shaded in Fig. 1) above the line and under the curve ln(p(x)) is equal to Φ. Mark off $A = (A_1, A_2, \ldots)$ the projections on the x axis of the segments cut off from ℓ by the graph of ln(p(x)).

Step (3). $\phi(x) = 0$ outside of $A_1, A_2, \ldots$ Inside these intervals $\phi(x) = \ln p(x) - \ln(\lambda)$ is equal to the length of the vertical segment from ℓ to ln(p(x)) for the x in question.

This completes the solution to the problem. However, we can obtain further information which might prove useful.

Step (4). Graph on the same graph p(x) and the line $y = \lambda$ (Fig. 2)

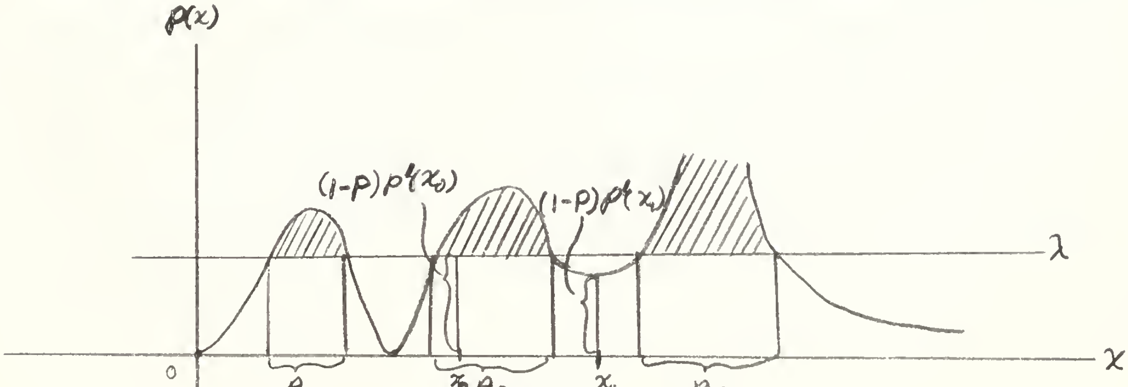

Figure 2.

Then the area (shaded in Fig. 2) above λ and below p(x) is the probability of successful search, P. The a-postiori probability density of the location of the object p'(x) is:

$$p'(x) = \begin{cases} p(x)/_{1-P} & \text{for } x \notin A \\ \lambda/_{1-P} & \text{for } x \in A \end{cases}$$

So, if after an unsuccessful expenditure of effort Φ a new quantity Φ' of effort becomes available, its optimum distribution can be computed using p'(x) instead of p(x). In this case (exponential) it turns out that our optimum probability of detection for allocation of $\Phi + \Phi'$ units of effort is the same whether we apply it in successive stages or all in one stage.

Example 2. The Neyman-Pearson Lemma.

We shall now show that the Neyman-Pearson Lemma of Statistics is an application of Lemma I under suitable identification.

Suppose we have observations taken on n random variables (or as is usually the case n observations of the same process). We desire to test the hypothesis that the joint density function $f(x_1, x_2, \ldots x_n)$ is a given function f_o against the alternative that $f(x_1, x_2, \ldots x_n)$ is another function f_1 at significance level α.

By significance level α we mean we want the probability of rejecting the hypothesis when it is true to be less than or equal to α. If we define the power of the test to be the probability of rejecting the hypothesis when the alternative is true we might then desire to find a criterion based upon the observations that will maximize the power for the given α. If such a test exists, it is called the most powerful level α test.

If we let $t = (x_1, x_2, \ldots x_n)$, we have the probability of rejecting the hypothesis when it is true is $\int_A f_o(t)$ where A is the set of points t whose observation would cause us to reject the hypothesis. Thus a test (or procedure) is a selection of a set A (or a partition of E_n) such that we reject the hypothesis if the observation x lies in the set A. The set A, so determined, is called the critical region. Similarly the power of the test may be written: $\int_A f_1(t)$.

Our problem then is to find the set A^o that yields the maximum power for our level α test. Letting C_A be the characteristic function of the set A our problem is to:

maximize $\int f_1(t)\ C_A(t)$ — The power of the test using critical region A

subject to: (1) $0 \leq C_A \leq 1$

(2) $\int f_o(t)\ C_A(t) \leq \alpha$ — The test must be of level α.

We now let:

$$x(t) = f_o(t)\ C_A\ (t)$$

$$W(t) = \frac{f_1(t)}{f_o(t)}$$

If $f_o(t_o)$ were zero we would obviously reject the hypothesis upon observation of t_o so we can restrict our problem to the region wherein $f_o(t) > 0$. Thus we shall assume $f_o(t) > 0$ throughout the problem.

Our problem now becomes:

(3) maximize $\int W(t)\, x(t)$

subject to: (4) $0 \leq x(t) \leq f_o(t)$

(5) $\int x(t) \leq \alpha$

Since W(t) is non-negative we can without loss of generality write (5) as:

(6) $\int x(t) = \alpha$

Application of Lemma I to the problem of (3), (4) and (6) tells us that for $x^o(t)$ to be a maximizing solution there must be a number λ such that if

$$(7) \qquad W(t) \begin{cases} > \lambda, & x^o(t) = f_o(t) \\ < \lambda, & x^o(t) = 0 \\ = \lambda, & x^o(t) \text{ is undetermined.} \end{cases}$$

Substituting for $x^o(t)$ and $W(t)$ in (7) we obtain:

$$(8) \quad \text{if} \quad \frac{f_1(t)}{f_o(t)} \begin{cases} > \lambda, & C_{A^o}(t)\, f_o(t) = f_o(t) \\ < \lambda, & C_{A^o}(t)\, f_o(t) = 0 \\ = \lambda, & \text{undetermined} \end{cases}$$

Recalling the meaning of $A^o(t)$ we have the conclusion that the most powerful test of hypothesis at level α consists of determining a number λ (depending on f_1, f_o and α) and to:

reject the hypothesis if $\frac{f_1(t)}{f_o(t)} > \lambda$

accept the hypothesis if $\frac{f_1(t)}{f_o(t)} < \lambda$

we have no test if $\frac{f_1(t)}{f_o(t)} = \lambda$.

where t is the observed vector.

These two examples, while more meaningful than those of Section 3, certainly fall short of illustrating the full power of the theorems of Section 4. The interested reader can find much more sophisticated and meaningful examples in the papers by Danskin and Gillman [3] and Danskin [4]. A particularly interesting problem is investigated by

Karlin, Madow, and Pruitt [10] who attack the problem of determining optimum firing rates as a function of range for a weapon of given characteristics against a particular type of target and then proceed using these maximum payoff functions to consider optimum mixes of several types of weapons against supposed mixes of targets subject to an overall cost or weight restriction on the weapons. A recently published attack on this same class of problems but more restricted than the one of this paper can be found in the article by Zahl [18].

For the reader who is interested in a particularly thorough discussion of the approach to problems of this type the two-volume set of books by Karlin [9] is recommended, especially Chapter 8 of Volume II.

BIBLIOGRAPHY

1. Bellman, R. E. and S. E. Dreyfus. Applied Dynamic Programming. Princeton University Press, 1962.

2. Charnes, A. and W. W. Cooper. The Theory of Search: Optimum Distribution of Search Effort. Management Science, v. 5, 1958: 44-49.

3. Danskin, J. M. and L. Gillman. A Game Over Function Space. Rivista di Matematica della Universita di Parma 4, 1953: 83-94.

4. Danskin, J. M. Mathematical Treatment of a Stockpiling Problem. Naval Research Logistics Quarterly, v. 2, 1955: 99-109.

5. Danskin, J. M. A Theory of Reconnaissance: I & II. Operations Research, v. 10, May-June 1962: 285-309.

6. Danskin, J. M. A Game Theory Model of Convoy Routing. Operations Research, v. 10, Nov.-Dec. 1962: 774-785.

7. de Guenin, J. Optimum Distribution of Effort: An Extension of the Koopman Basic Theory. Operations Research, v. 9, Jan.-Feb. 1961: 1-7.

8. den Broeder, G. G., R. E. Ellison, and L. Emerling. On Optimum Target Assignments. Operations Research, v. 7, 1959: 322-326.

9. Karlin, S. Mathematical Methods and Theory in Games, Programming, and Economics, Vol. I & II. Addison Wesley, 1959.

10. Karlin, S., W. E. Pruitt, and W. G. Madow. On Choosing Combinations of Weapons. Naval Research Logistics Quarterly, v. 10, June 1963: 95-119.

11. Karush, W. A General Algorithm for the Optimal Distribution of Effort. Management Science, v. 9, Oct. 1962: 50-72.

12. Koopman, B. O. The Theory of Search III: The Optimum Distribution of Searching Effort. Operations Research, v. 5, Oct. 1957: 613-626.

13. Lemus, F. and H. K. David. An Optimum Allocation of Different Weapons to a Target Complex. Operations Research, v. 11, Sept.-Oct. 1963: 787-794.

14. Manne, A. S. A Target Assignment Problem. Operations Research, v. 6, 1958: 346-351.

15. Munroe, M. E. Introduction to Measure and Integration. Addison Wesley, 1953.

16. Piccariello, H. J. A Missile Allocation Problem. Operations Research, v. 10, Nov.-Dec. 1962: 795-798.

17. Saks, S. Theory of the Integral. Hafner Publishers, 1937.

18. Zahl, S. An Allocation Problem With Applications to Operations Research and Statistics. Operations Research, v. 11, May-June 1963: 426-441.