THE ROLE OF THE NEYMAN-PEARSON LEMMA IN THE

THOERY OF DISCRETE SEARCH

Joseph B. Kadane

March 3, 1967
The Role of the Neyman-Pearson Lemma in the

Theory of Discrete Search*

by

Joseph B. Kadane

1. Introduction

Suppose an object is hidden in one of \( n \) boxes. It is in box \( k \) with probability \( p_k, k = 1, \ldots, n \). If it is in the \( k \)th box, a search of the \( k \)th box may overlook it with probability \( \alpha_k, 0 < \alpha_k < 1 \). The events \( E_{j,k} \) that the object is found in the \( j \)th search of the \( k \)th box are disjoint, and

\[
P_{j,k} = \Pr[E_{j,k}] = p_k \alpha_k^{j-1} (1-\alpha_k)
\]

for \( k = 1, \ldots, n \) and all positive integers \( j \). Suppose also that each search of box \( k \) costs \( c_k > 0 \). The main problem considered in this paper is how to search in order to maximize the probability of finding the object spending no more than a fixed amount \( C \).

For the moment we leave aside the consideration that if a procedure specifies that the tenth search of box 1 is to be conducted, the first nine searches must have been conducted. This is the requirement that a procedure be feasible. Later we will show that the best procedure for the more general problem is feasible, and hence is the best feasible procedure as well.

* Research undertaken by the Cowles Commission for Research in Economics under Contract Nonr-3055(00) with the Office of Naval Research.
Probability measure and cost measure are both measures on the space \( \{ E_{j,k} \mid 1 \leq k \leq n, 1 \leq j < \infty \} \) of searches. If both had total measure equal to 1, our problem would be roughly the problem solved by the Neyman-Pearson Lemma. This powerful result tells how to choose a set (critical region) so that its integral with respect to one measure (null hypothesis) is less than or equal to a fixed \( \alpha \) (size of the test), and its integral with respect to another measure (alternative hypothesis) is as large as possible (power of the test).

Therefore the first theorem of this paper extends the Neyman-Pearson Lemma to arbitrary \( \sigma \)-finite measures. This result is undoubtedly not new; however it does not appear to be conveniently available. It is stated in Section 2; the proof is a straightforward generalization of the proof in Lehmann [7, pp. 65, 66]. Using this theorem, Section 3 discusses the search problem stated above.

The results generalize a result of Chew [4], and give the solution to a slightly modified version of a problem stated by Mosteller in Bellman [1, pp. 49, 50]. A corollary generalized a theorem of Staroverov [9], and bears a close relation to work of Blackwell [3] and Black [2].

In Section 4, this theory is extended to searches with arbitrary probability \( p_{j,k} \) of success at the \( j \)th search of box \( k \), and arbitrary cost \( c_{j,k} \).

A Bibliography on search problems is given in [6].
2. Extension of the Neyman-Pearson Lemma

Theorem 1 Let $(\mathcal{X}, \mathcal{B})$ be a measurable space, and let $\mu_1$ and $\mu_2$ be any non-negative $\sigma$-finite measures on $(\mathcal{X}, \mathcal{B})$. Let $\mu$ be any measure with respect to which $\mu_1$ and $\mu_2$ are absolutely continuous ($\mu = \mu_1 + \mu_2$ will suffice). Let $f_1$ and $f_2$ be the Radon-Nikodym derivatives of $\mu_1$ and $\mu_2$, respectively, with respect to $\mu$.

Let $B = \int f_1(x) \, d\mu(x) = \mu_1(\mathcal{X})$, and let $\alpha$ be a number such that $0 \leq \alpha \leq B \leq \infty$. Then

(a) there exists a function $\phi(x), 0 \leq \phi(x) \leq 1$ and a number $r, 0 \leq r \leq \infty$, such that

\[(1) \quad \int_{\mathcal{X}} \phi(x) f_1(x) \, d\mu(x) = \alpha \]

\[(2) \quad \phi(x) = \begin{cases} 1 & f_2(x) > r f_1(x) \\ 0 & f_2(x) < r f_1(x) \end{cases} \]

(b) If $\phi$ satisfies (1) and (2), then it maximizes

\[\int_{\mathcal{X}} \phi(x) f_2(x) \, d\mu(x) \quad \text{subject to} \]

\[\int_{\mathcal{X}} \phi(x) f_1(x) \, d\mu(x) \leq \alpha \quad \text{and} \quad 0 \leq \phi(x) \leq 1 \]

(c) If $\phi$ maximizes $\int_{\mathcal{X}} \phi(x) f_2(x) \, d\mu(x)$ subject to

\[\int_{\mathcal{X}} \phi(x) f_1(x) \, d\mu(x) \leq \alpha \quad \text{and} \quad 0 \leq \phi(x) \leq 1 \]
then for some \( r \) it satisfies (2) a.e. \( \mu \) provided \( \int \phi(x) f_2(x) \, d\mu(x) < \infty \).

It also satisfies (1) unless there is a function \( \phi^* \) with
\[
\int \phi^*(x) f_2(x) \, d\mu(x) < \alpha \quad \text{and} \quad \int \phi^*(x) f_2(x) \, d\mu(x) = \mu_2 (\mathcal{X}) .
\]

Thus even in the infinite-measure case the ratio of densities (likelihood ratio in the testing problem) is the appropriate decision function.

The unfortunate device of the function \( \phi \) is forced on us by the necessity of randomization (or something similar) in the case of awkward \( \alpha \)'s. If (and only if) \( \alpha \) is a partial sum of costs of some optimal policy, \( \phi \) can be taken to have only the values zero and one, and the awkwardness does not occur. Thus the really general case in which the fixed cost is insisted upon is beyond the scope of this method. However this method does give very simple upper and lower bounds for the probability that the optimal policy will find the object.

This issue, and the problem of feasibility, will dominate the discussion in the next section.

3. Discrete Search

A policy is a set of pairs of integers \((j,k)\) specifying the searches to be conducted. A policy is feasible if the inclusion of \((j,k)\) implies the inclusion of \((j-1,k)\) for \(j=2, \ldots, \) and \(k=1, \ldots, n.\)

For the moment we will allow randomized policies, which specify for each pair of integers \((j,k)\) the probability \(q_{j,k}(j,k)\) that \((j,k)\) is included in the policy. The probability that such a (perhaps not feasible) policy will find the object is
\[
I. \quad E \phi(j,k) \rho_k \alpha_{j-1}^k (1-\alpha_k) .
\]

The expected cost (with the expectation taken only over the randomization of
the policy) is

\[ \sum_{j,k} \Phi(j,k) c_k. \]

We wish to find a policy which maximizes the probability of finding the object spending no more than the fixed amount \( C \).

In order to apply the extended Neyman-Pearson Lemma (Theorem 1), let \( \mathcal{X} = \{ (j,k) \mid j \geq 1, 1 \leq k \leq n, j \text{ and } k \text{ integers} \} \) and let \( \mathcal{B} \) be the class of all subsets of \( \mathcal{X} \). Now let \( \mu_1 \{ (j,k) \} = c_k, \mu_2 = \{ (j,k) \} = p_k \alpha_k^{j-1}(1-\alpha_k), \) and let \( \mu \) be counting measure.

We require, then, that \( 0 \leq \alpha = C \leq B = \infty \). Actually we can assume that \( 0 < C < B \) since if one is permitted to spend \( B \), the choice of all searches is trivially optimal, and if one is permitted no cost, only free searches are possible.

Then from Theorem 1 we know that

(a) there exists a function \( \Phi(j,k), 0 \leq \Phi(j,k) \leq 1 \) and a number \( r, 0 < r < \infty \) such that

\[ \sum_{j=1}^{\infty} \sum_{k=1}^{n} \Phi(j,k) c_k = C \]

\[ \Phi(j,k) = \begin{cases} 1 & \text{if } p_k \alpha_k^{j-1}(1-\alpha_k) > r c_k \\ 0 & \text{if } p_k \alpha_k^{j-1}(1-\alpha_k) < r c_k \end{cases} \]

(b) if \( \Phi \) satisfied (1) and (2), then it maximizes

\[ \sum_{j=1}^{\infty} \sum_{k=1}^{n} \Phi(j,k) p_k \alpha_k^{j-1}(1-\alpha_k) \]
subject to
\[ \sum_{j=1}^{n} \sum_{k=1}^{\infty} c_k \Phi(j,k) \leq C \quad \text{and} \quad 0 \leq \Phi(x) \leq 1 \]

(c) If \( \Phi \) maximizes \[ \sum_{j=1}^{n} \sum_{k=1}^{\infty} p_k c_k^{j-1} (1-\alpha_k) \Phi(j,k) \]
subject to \[ \sum_{j=1}^{n} \sum_{k=1}^{\infty} c_k \Phi(j,k) \leq C, \] then for some \( r \) it satisfies (2) (since total probability is less than or equal to 1). It also satisfies (1) (since \( C < B \) by assumption).

Then it is clear that we wish to include in our search all pairs \((j,k)\) for which
\[ \frac{p_k c_k^{j-1}(1-\alpha_k)}{c_k} > r \]
and exclude all those for which
\[ \frac{p_k c_k^{j-1}(1-\alpha_k)}{c_k} < r. \]

There are two possibilities for interpreting \( \Phi \) when it is neither zero nor one. The first, of course, is the idea of a randomized search, in which \( \Phi(j,k) \) is the probability that the \((j,k)\) search is included. It is clear that at most one search need be randomized.

The second is to permit a partial last search, such that in the last search one can expend some fraction \( s \) of the cost \( c_k \) of that search, and have probability \( s p_k c_k^{j-1}(1-\alpha_k) \) of finding the object in that search.
Under either of the above two interpretations of \( \phi \) when it is neither zero nor one, the sum I is exactly the probability of finding the object, and the sum II is exactly the cost (averaged in the case of randomization). Thus in both cases, part b) applies to give optimality. Let us say that a policy is \textit{locally optimal} if inclusion of \((j';k')\) and exclusion of \((j,k)\) implies

\[
\frac{p_{k} \alpha_{k}^{j'-1}(1-\alpha_{k'})}{c_{k'}} > \frac{p_{k} \alpha_{k}^{j-1}(1-\alpha_{k})}{c_{k}}
\]

By part (c) of Theorem 1, only locally optimal policies can be solutions to the problem.

Furthermore, all locally optimal policies are feasible since

\[
\frac{p_{k} \alpha_{k}^{j-1}(1-\alpha_{k})}{c_{k}}
\]

is monotone decreasing in \( j \) for all \( k \). Thus locally optimal policies, and only locally optimal policies, are solutions to the problem of finding a feasible policy maximizing the probability of finding the object spending no more than \( C \), \textit{provided} either a partial last search or randomization is permitted.

It is possible, of course, to consider the problem in which neither of these interpretations is acceptable. In general this is an integer programming problem similar to the knapsack problem [5, p. 517 ff].
Because the cost of each search of box \( k \) is the same, and \( p_{j,k} \) decreases in \( j \), any optimal solution of the unrestricted knapsack problem will be feasible. Algorithms for this problem are discussed in [10, chapter 4 and 5]. See also the references in [11].

However, from the above theory it is clear that both upper and lower bounds are obtainable. Thus the procedure with the largest probability of finding the object in a fixed cost \( C \) must have probability no smaller than that of the largest \( C' \), smaller than \( C \), which is a partial sum of costs of the policy defined by decreasing

\[
\frac{p_i \alpha_{i-1}(1-\alpha_k)}{c_k}
\]

Also, of course, the largest probability of finding the object in a fixed cost \( C \) when a partial last search is not permitted must be no larger than the probability of finding the object if a partial last search is permitted. These bounds can be expected to be very close if the \( c_k \)'s are much smaller than \( C \).

The special case in which \( c_k = 1 \) for all \( k \) has been studied by Chew [4]. He found that the policy of taking largest \( p_k \alpha_{k-1}(1-\alpha_k) \) maximizes the probability of finding the object in a fixed number of searches (cost when \( c_k = 1 \) for all \( k \)). Then for any integer \( C \) a partial or randomized last search is not required, and Theorem 1 applies to give Chew's result.

To summarize, we have the following result:
Theorem 2

Any policy which maximizes the probability of finding the object spending no more than a fixed cost $C$, $0 < C < \infty$, includes all searches for which

$$\frac{p_k \alpha_k^{j-1}(1-\alpha_k)}{c_k} > r$$

for some $r$, excludes all those for which

$$\frac{p_k \alpha_k^{j-1}(1-\alpha_k)}{c_k} < r$$

and includes enough of those with $\frac{p_k \alpha_k^{j-1}(1-\alpha_k)}{c_k} = r$ to spend exactly $C$.

Any such policy is feasible. A partial or randomized last search is unnecessary if and only if $C$ is a partial sum of some optimal policy.

We conclude this section by discussing a closely related problem, that of finding a sequence of searches which minimizes the expected cost of finding the object.

In the problem in which randomization or a partial last search is permitted a sequence ordered by decreasing $\frac{p_k \alpha_k^{j-1}(1-\alpha_k)}{c_k}$ and including all searches has the property that it maximizes the probability of finding the object with any fixed expenditure $C$. It must minimize the expected cost of finding the object. (Obviously any procedure hoping to finite expected cost must include all searches and we must assume $\Sigma p_k = 1$.)

However the expected cost any sequence when stopping is permitted the moment the object is found corresponds to the cost of all unsuccessful searches plus half the cost of the search in which the object is found in the discrete case. Hence we have
Corollary

Any sequence including all possible searches and ordered by decreasing

\[
\frac{p_k \alpha_j^{j-1} (1-\alpha_j)}{c_k}
\]

minimizes the expected cost of finding the object, with the searcher charged for only half the cost of the successful last search.

Staroverov [9] discusses the special case \( c_k = 1, p_{j,k} = p_k (1-p)^{j-1} \) and proves that the expected cost minimizing procedure chooses according to decreasing

\[ p_k (1-p)^{j-1} \]

Since the costs of all searches are equal, minimizing the expected cost of all unsuccessful searches plus half the cost of the successful one is equivalent to minimizing the expected cost of all searches, successful or not. Thus Staroverov's result is a special case of the corollary above.

Blackwell [3] and Black [2] show that the procedure in the corollary also minimizes the expected cost of all unsuccessful searches plus the cost of the successful last search. In a later paper, using entirely different methods, I will show that this same procedure minimizes the expected cost of all unsuccessful searches plus any fraction \( f \) of the cost of the last search, \( 0 \leq f \leq 1 \).

4. A More General Discrete Search

Suppose now that the object will be found in the \( j \)th search of the kth box with probability \( p_{j,k} \) and that the \( j \)th search of the kth box will cost \( C_{j,k} > 0 \) to conduct. The purpose of this section is to show to what extent the argument of section 3 carries over to this case.
Making the obvious extensions of definitions, we see that once again theorem 1 applies, and that the critical quantity is $P_{jk}/C_{jk}$. If $P_{jk}/C_{jk}$ is monotone decreasing in $jk$ then any optimal policy is feasible. Thus we obtain the following summary statement:

Assume that $P_{jk}/C_{jk}$ is monotone decreasing in $j$. Any policy which maximizes the probability of finding the object spending no more than a fixed cost $C$, $0 < C < \sum_{j,k} C_{jk}$, includes all searches for which

$$P_{jk}/C_{jk} > r$$

for some $r$, excludes all those for which

$$P_{jk}/C_{jk} < r$$

and includes enough of those with $P_{jk}/C_{jk} = r$ to spend exactly $C$. Any such policy maximizes the probability of finding the object spending no more than $C$. Any such policy is feasible. A partial or randomized last search is unnecessary if and only if $C$ is a partial sum of some optimal policy.

**Corollary**

Any sequence including all possible searches and ordered by decreasing $P_{jk}/C_{jk}$ minimizes the expected cost of the unsuccessful searches plus half the cost of the successful search.

However a special difficulty can occur in this more general case for the problem of minimizing the expected cost. Consider, for example, a two-box situation where
\[ P_{j1} = \frac{1}{2} e^{-\lambda} \frac{\lambda^{j-1}}{(j-1)!} \quad \Sigma P_{j1} = 1/2 \]
\[ C_{j1} = \frac{1}{2} e^{-\lambda} \frac{\lambda^{j-1}}{(j-1)!} \cdot \frac{j}{j+1} \quad \Sigma C_{j1} = \frac{\lambda^2}{2} [\lambda^2 + 1 - e^{-\lambda}] \]
\[ P_{j2} = (1/2)^{j+1} \quad \Sigma P_{j2} = 1/2 \]
\[ C_{j2} = 1 \]

Then
\[ \frac{P_{j1}}{C_{j1}} = \frac{j+1}{j} = 1 + \frac{1}{j} \quad \text{and} \quad \frac{P_{j2}}{C_{j2}} = (1/2)^{j+1}. \]

Thus to maximize the probability of finding the object spending \( C < \frac{\lambda^2}{2} [\lambda^2 + 1 - e^{-\lambda}] \), as many searches as possible of box 1 should be conducted. For \( C = \frac{\lambda^2}{2} [\lambda^2 + 1 - e^{-\lambda}] \) all searches of box 1 should be conducted. For \( C = \frac{\lambda^2}{2} [\lambda^2 + 1 - e^{-\lambda}] + 1 \), all searches of box 1 and the first search of box 2 should be conducted.

For the problem of minimizing the expected cost of finding the object, however, order matters, and it would be optimal, if it were possible, to have all searches of box 1 followed by all searches of box 2. This difficulty arises because the solution to the problem of maximizing probability is a set, while the solution to the problem of minimizing expected cost is a sequence. For the above problem, only \( \varepsilon \)-optimal solutions exist to the problem of finding the expected-cost minimizing sequence.

Thus the corollary above is true, but in this case no sequence including all possible searches and ordered by decreasing \( P_{j,k}/C_{j,k} \) exists.
We have assumed that \( \frac{P_{jk}}{C_{jk}} \) is strictly monotone decreasing in \( j \). Since these ratios are bounded from below by zero, for each \( k \):

\[
a_k = \lim_{j \to \infty} \frac{P_{jk}}{C_{jk}} \text{ exists.}
\]

Because of the decreasing character of the ratios, \( a_k \) is never attained. In the example above, \( a_1 = 1 \) and \( a_2 = 0 \). A necessary and sufficient condition that a sequence including all possible searches and ordered according to decreasing \( \frac{P_{jk}}{C_{jk}} \) exist is that there exists an \( a \) such that

\[
a = a_k \quad k = 1, \ldots, n.
\]

Thus in particular if \( C_{jk} \)'s are bounded away from zero, (3) will be satisfied with \( a = 0 \).

If instead we assume that \( \frac{P_{jk}}{C_{jk}} \) is merely non-increasing in \( j \), several changes must be made in the above argument. In the probability maximizing problem, no longer is every optimal policy feasible. However, an optimal feasible policy does exist, and the theorem holds for all optimal feasible policies.

In the problem of minimizing expected cost only searches with \( P_{jk} \neq 0 \) need be included in the desired sequence. Then a necessary and sufficient condition for the existence of a feasible sequence ordered by \( \frac{P_{jk}}{C_{jk}} \) is that

(1) \( a = a_k = \lim_{j \to \infty} \frac{P_{jk}}{C_{jk}} \) for all \( k \) such that \( P_{jk} > 0 \) \( \forall j \), and

(2) if the limit \( a \) is attained by any \( \frac{P_{jk}}{C_{jk}} \) such that \( P_{jk} > 0 \), then it must be attained, for some \( j \), in each other sequence \( \frac{P_{jk}}{C_{jk}} \) with \( P_{jk} > 0 \) for all \( j \).
(3) \( \frac{P_{jk}}{C_{jk}} \geq a \) for all \((j,k)\) such that \( P_{jk} > 0 \)

In conclusion, we have the following

Theorem 3 [Final Version].

Assume that \( \frac{P_{jk}}{C_{jk}} \) is non-increasing in \( j \) for each \( k \). Any policy which maximizes the probability of finding the object spending no more than a fixed cost \( C \), \( 0 < C < \sum_j C_{jk} \) \( \text{Sgn} (P_{jk}) \) includes all searches for which

\[ \frac{P_{jk}}{C_{jk}} > r \]

for some \( r \), excludes all those for which

\[ \frac{P_{jk}}{C_{jk}} < r \]

and includes enough of those with \( \frac{P_{jk}}{C_{jk}} \) to spend exactly \( C \). Any such policy maximizes the probability of finding the object spending no more than \( C \).

Not any such policy is feasible, but a feasible such policy does exists. A partial or randomized last search is unnecessary if and only if \( C \) is a partial sum of some optimal feasible policy.

Corollary

Any feasible sequence including all possible searches such that \( P_{jk} \neq 0 \) and ordered by non-increasing \( \frac{P_{jk}}{C_{jk}} \) minimizes the expected cost of the unsuccessful searches plus half the cost of the successful search.

Such a sequence exists if and only if

(1) \( a = a_k = \lim_{j \to \infty} \frac{P_{jk}}{C_{jk}} \) for all \( k \) such that \( P_{jk} > 0 \) \( \forall j \)

and
(2) If the limit $a$ is attained by any $P_{jk}/C_{jk}$ such that $P_{jk} > 0$, then it must be attained, for some $j$, in each other sequence $P_{jk}/C_{jk}$ with $P_{jk} > 0$ for all $j$.

(3) $P_{jk}/C_{jk} \geq a$ for all $(j,k)$ such that $P_{jk} > 0$.

Acknowledgements

I wish to thank Morris DeGroot for bringing this problem to my attention, and L. J. Savage for helpful discussions on it.
Appendix

Proof of Theorem 1

For $\alpha = 0$ and $\alpha = B$, the theorem is true when $K = \infty$ and $k = 0$ are permitted. Therefore assume $0 < \alpha < B$.

(2) Let $\alpha(S) = \int f_1(x) \, d\mu(x)$

\[
(x \mid f_2(x) > S f_1(x), f_1(x) > 0)
\]

$\alpha(0) = B$, $\alpha(\infty) = 0$, and $\alpha$ is right continuous and non-increasing. Then for any $0 \leq \alpha \leq B$, there is an $S_0$, $0 \leq S_0 \leq \infty$ such that $\alpha(S_0) \leq \alpha \leq \alpha(S_0 - 0)$.

If $\alpha(S_0) = \alpha(S_0 - 0)$, then let

$\phi(x) = \begin{cases} 1 & f_2(x) > S_0 f_1(x) \\ 0 & f_2(x) \leq S_0 f_1(x) \end{cases}$

and

\[
\int \phi(x) f_1(x) \, d\mu(x) = \int f_1(x) \, d\mu(x) = \alpha(S_0) = \alpha
\]

\[
(x \mid f_2(x) > S_0 f_1(x), f_1(x) > 0)
\]

If $\alpha(S_0) \neq \alpha(S_0 - 0)$, let

\[
\frac{1}{\alpha(S_0 - 0) - \alpha(S_0)} \begin{pmatrix} 1 & f_2(x) > S_0 f_1(x) \\ \alpha - \alpha(S_0) & f_2(x) = S_0 f_1(x) \\ 0 & f_2(x) < S_0 f_1(x) \end{pmatrix}
\]
Then \(\int \phi(x) f_2(x) \, d\mu(x) = \int f_1(x) \, d\mu(x)\)

\[ \{x | f_2(x) > S_0 f_1(x), f_1(x) > 0\} \]

\[+ \frac{c - c(S_0)}{\alpha(S_0) - \alpha(0)} \int f_1(x) \, d\mu(x)\]

\[= \alpha(S_0) + \frac{\alpha(S_0)}{\alpha(S_0) - \alpha(0)} \cdot c(S_0) \cdot c(0) - \alpha(S_0)\]

\[= \alpha\]

Thus in both cases \(S_0\) will suffice as \(r\) in the theorem, and part (a) is complete.

(b) Suppose that \(\phi\) satisfies (1) and (2). If

\[\int \phi(x) f_2(x) \, d\mu(x) = \infty, \phi\] clearly maximizes

\[\int \phi(x) f_2(x) \, d\mu(x)\] and we are done. Therefore

suppose \[\int \phi(x) f_2(x) \, d\mu(x) < \infty\].

Let \(\phi^*\) be any other function, \(0 \leq \phi^*(x) \leq 1\),

and \[\int \phi^*(x) f_1(x) \, d\mu(x) \leq \alpha\].

Let \(S^+ = \{x | \phi(x) - \phi^*(x) > 0\}\) and \(S^- = \{x | \phi(x) - \phi^*(x) < 0\}\).

Then \(x \in S^+ \Rightarrow \phi(x) \neq 0 \Rightarrow f_2(x) > r f_1(x)\)

\[x \in S^- \Rightarrow \phi(x) \neq 1 \Rightarrow f_2(x) \leq r f_1(x),\] so

\[\int (\phi(x) - \phi^*(x))(f_2(x) - rf_1(x)) \, d\mu(x) = \int (\phi(x) - \phi^*(x))(f_2(x) - rf_1(x)) \, d\mu(x)\]

\[S^+ \cup S^- \geq 0\]
Then \( \int (\phi(x) - \phi^*(x)) f_2(x) \, d\mu(x) \geq k \int (\phi(x) - \phi^*(x)) f_1(x) \, d\mu(x) \geq 0 \)
as was to be shown.

(c) Suppose \( \phi^*(x) \) maximizes

\[ \int \phi(x) f_2(x) \, d\mu(x) \text{ subject to } \int \phi(x) f_1(x) \, d\mu(x) = \alpha \]

(with \( \int \phi^*(x) f_2(x) \, d\mu(x) < \infty \)) and suppose \( \phi(x) \)
satisfies (1) and (2). Let \( S \) be the intersection of the set \( S^+ \cup S^- \), on which \( \phi \) and \( \phi^* \) differ, with the set \( \{ x \mid f_2(x) \neq r f_1(x) \} \), and suppose \( \mu(S) > 0 \).

Then since \( (\phi(x) - \phi^*(x))(f_2(x) - r f_1(x)) \) is positive on \( S \), it follows that

\[ \int_{S^+ \cup S^-} (\phi(x) - \phi^*(x))(f_2(x) - r f_1(x)) \, d\mu(x) = \int_{S} (\phi(x) - \phi^*(x))(f_2(x) -rf_1(x)) > 0 \]

and therefore

\[ \int (\phi(x) - \phi^*(x)) f_1(x) \, d\mu(x) > r \int (\phi(x) - \phi^*(x)) f_1(x) \geq 0 \]

which contradicts the assumption on \( \phi^* \). Then \( \mu(S) = 0 \) as claimed.

If \( \phi^* \) were such that \( \int \phi^*(x) f_1(x) \, d\mu(x) < \alpha \) and \( \int \phi^*(x) f_2(x) \, d\mu(x) < \mu_2(\mathcal{X}) \)
it would be possible to increase \( \phi^* \) slightly, increasing both integrals until either \( \int \phi^*(x) f_1(x) \, d\mu(x) = \alpha \) or \( \int \phi^*(x) f_2(x) \, d\mu(x) = \mu_2(\mathcal{X}) \).

Q.E.D.
REFERENCES


