STOCHASTIC SEQUENTIAL ASSIGNMENT PROBLEMS

BY

SAMUEL CHRISTIAN ALBRIGHT, JR.

TECHNICAL REPORT NO. 147
APRIL 27, 1972

SUPPORTED BY THE ARMY AND NAVY
UNDER CONTRACT NO0014-67-A-0112-0052 (NR-042-002)
WITH THE OFFICE OF NAVAL RESEARCH

Reproduction in Whole or in Part is Permitted
for any purpose of the United States Government
Distribution of this Document is Unlimited

DEPARTMENT OF OPERATIONS RESEARCH
AND
DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
STOCHASTIC SEQUENTIAL ASSIGNMENT PROBLEMS

by

Samuel Christian Albright, Jr.

TECHNICAL REPORT NO. 147
April 27, 1972

Supported by the Army and Navy
with the Office of Naval Research

Gerald J. Lieberman, Project Director

Reproduction in whole or in part is permitted
for any purpose of the United States Government

Distribution of this document is unlimited

DEPARTMENT OF OPERATIONS RESEARCH

AND

DEPARTMENT OF STATISTICS

STANFORD UNIVERSITY

STANFORD, CALIFORNIA
# Table of Contents

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Preliminary Results</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>Assignment Problems with Exactly n jobs arriving</td>
<td>18</td>
</tr>
<tr>
<td>3.1</td>
<td>Successive Job Distributions Determined by a Markov Chain</td>
<td>19</td>
</tr>
<tr>
<td>3.2</td>
<td>Job Distribution Comes from a Random Mixture</td>
<td>41</td>
</tr>
<tr>
<td>3.3</td>
<td>Job Distribution Comes from a Random Mixture; r Categories of Men</td>
<td>42</td>
</tr>
<tr>
<td>3.4</td>
<td>Job Distributions Determined by a Markov Chain; r Categories of Men</td>
<td>57</td>
</tr>
<tr>
<td>3.5</td>
<td>Bayesian Model of the Assignment Problem</td>
<td>62</td>
</tr>
<tr>
<td>3.6</td>
<td>A Model Using Martingales</td>
<td>72</td>
</tr>
<tr>
<td>3.7</td>
<td>Choosing the Best p's.</td>
<td>73</td>
</tr>
<tr>
<td>4</td>
<td>Assignment Problems with a Variable Number of Jobs Arriving</td>
<td>85</td>
</tr>
<tr>
<td>4.1</td>
<td>All Jobs Must be Assigned; No Time Factor</td>
<td>86</td>
</tr>
<tr>
<td>4.2</td>
<td>All Jobs Must be Assigned; Time is a Factor</td>
<td>91</td>
</tr>
<tr>
<td>4.3</td>
<td>Jobs may be Bypassed; Time is a Factor</td>
<td>95</td>
</tr>
<tr>
<td>4.4</td>
<td>Jobs may be Bypassed; Time is no Factor</td>
<td>127</td>
</tr>
<tr>
<td>5</td>
<td>Several More Assignment Problems</td>
<td>133</td>
</tr>
<tr>
<td>5.1</td>
<td>Model 1: Men Reavailable After Finishing Jobs</td>
<td>133</td>
</tr>
<tr>
<td>5.2</td>
<td>Model 2: Random Replacements After Assignments</td>
<td>168</td>
</tr>
<tr>
<td>5.3</td>
<td>Model 3: Issuing Batteries</td>
<td>180</td>
</tr>
<tr>
<td></td>
<td>REFERENCES</td>
<td>190</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION

In a recent paper by Derman, Lieberman, and Ross [3], the following stochastic sequential assignment problem is treated. We have \( n \) men with "values" \( p_1 \leq \cdots \leq p_n \), and these men must be assigned to \( n \) jobs which arrive sequentially. The successive jobs have "values" \( X_1, \ldots, X_n \), which are assumed to independent and identically distributed (i.i.d.) random variables with a known distribution function \( F \). It is assumed that if a "p" man is assigned to an "x" job, a reward of \( px \) is received. An optimal policy is one which maximizes the expected reward from the \( n \) men.

For reference we record the main result from this paper:

**Theorem:** If \( n \) men and \( n \) jobs remain, then there exist numbers

\[ -\infty = a_0, n < a_1, n \leq \cdots \leq a_{n-1}, n < a_n, n = +\infty, \]

such that if the next job has value \( a_{i-1}, n < x \leq a_i, n \), it is optimal to assign the \( i \)th smallest job to this job. Furthermore, these \( a_i, n \)'s are independent of the values of the men, the p's, and they can be calculated by the recursion

\[
a_{i, n+1} = \int_{a_{i-1}, n}^{a_i, n} xF(dx) + a_{i-1}, n F(a_{i-1}, n) + a_{i, n} (1 - F(a_i, n)) .
\]

That this is true follows mainly from an earlier result of Hardy, Littlewood, and Pólya [5].
The present paper attempts to generalize the above model, which we refer to as the "original assignment model", as much as possible, thereby allowing more randomness and generality in the various aspects of the basic model. In all of these variations, however, there is one common characteristic, namely, we must assign one of several p's to an x, given that the future X's are random variables, and hence unknown, and we wish to do this so as to maximize the expected reward we receive. That is, we must make sequential decisions based on an uncertain future. Also, in almost all cases, we try to consider models which have fairly simple optimal policies, which are, in particular, independent of the p's. This has been possible in all but three of the models considered, and even in these, we show how the optimal policies can be found.

Throughout the paper, except in a few instances, we maintain the framework of men being assigned to jobs. This has been done for several reasons. First, it gives continuity to the subject matter, which we would not have if we discussed each model in a different setting. Secondly, most of the models we consider actually do make sense with the men and jobs interpretation, and thus it is not unnatural to keep this setting throughout. However, many of the models do have other interpretations which, from an applications point of view, are at least as useful as the men and jobs interpretation. Therefore, we will discuss these other interpretations both in this introduction and after the models where they arise.

The literature on the type of models discussed here is sparse. Only the paper by Derman, Lieberman, and Ross [3] discusses the assignment of a large number of items (the men) of various quality to a sequence of
incoming items (the jobs) of random quality, and this is basically the problem we are interested in. However, many authors have treated a single special case of this problem. This special case comes under the various headings of "selling a house", "disposing of an asset", "choosing a secretary", and others, and an example of at least one of these may be found in many texts on probability. The common characteristic of these problems is that there is one item (a house, a secretarial position, etc.) and it must be assigned, or disposed of, sometime in the future, in an optimal manner. Karlin [6] and Elfving [4], in particular, have written interesting papers on this problem, and their results are special cases of the results presented here.

We now summarize the contents of the present paper. In Chapter 2, we present several lemmas which are basic to the material which follows. In particular, we present the lemma of Hardy, Littlewood, and Pólya [5], and generalize it in several directions. Also, a lemma is presented which gives necessary and sufficient conditions for several of our models to have optimal policies which are independent of the p's.

In Chapter 3, we proceed with the most direct generalizations of the original assignment model. In all of these generalizations, we assume that exactly n X's are arriving sequentially, and that if a p is assigned to an x, a reward px is received. One interpretation of the problem which makes this reward function feasible is the following. Suppose the value p of a man is his productivity per unit time, and a job of value x lasts for a time x. Thus with the reward function px, we try to maximize the total expected productivity of the n men.
In the first few generalizations of the original model which we consider, we assume the successive job values come from one of $r$ distributions $F_1, \ldots, F_r$. We assume that if the previous job came from distribution $F_i$, then the next job comes from distribution $F_j$ with probability $q_{ij}$, where we allow the possibility that $q_{ij}$ is independent of $i$. Thus the successive job distributions are governed by a Markov Chain with transition probabilities $q_{ij}$, and when a job arrives, we observe not only the state of the chain $F_j$, but also a sample random variable $X$ from distribution $F_j$. This Markov chain formulation includes the possibilities that (1) $r = 1$, so that we are back to the original assignment model, (2) $F_j$ is degenerate at a point $x_j$, so that the actual values of the jobs are governed by a Markov chain, and (3) $q_{i,i+1(\text{mod } r)} = 1$, which shows cyclical, or seasonal, effects on the successive job distributions. One meaning of the different distributions is that they indicate different kinds of jobs, say plumbers as opposed to electricians.

In the first few models, we assume any man can do any of the incoming jobs. Under this assumption, optimal critical number policies are found which are independent of the $p$'s. We then let $n \to \infty$ and find the limiting behavior of these optimal policies. These limiting results not only provide elegant theory, but they also provide nearly optimal policies for large $n$ which are much easier to calculate than the optimal policies.

Next we assume there are $r$ categories of men, and that a man from category $i$ can do jobs only from distribution $F_i$. This is especially appealing when jobs from different distributions actually do
represent different kinds of jobs. Using the fact that men from different categories do not interact with each other, we are again able to find optimal critical number policies for these models which are independent of the p's in the various categories. Also, we again find limiting results for these optimal policies as \( n \to \infty \).

One model which fits nicely into the above framework is the following. Suppose all job values come from one distribution \( F \), but there exist numbers \( 0 < t_1 < \cdots < t_r < \infty \) and \( 0 < s_1 < \cdots < s_{r-1} < \infty \) such that if a job has value \( t_{i-1} < x \leq t_i \), then the man assigned to this job must have value \( s_{i-1} < p \leq s_i \). In the men and jobs interpretation, this might be appropriate if we never want to assign a man to a job for which he is either too good or not good enough. Similarly, it might also apply to the assigning of hospital rooms (men) to incoming patients (jobs).

We also consider a Bayesian model of the original assignment problem. Here all of the jobs come from one family of distributions \( F(x|\theta) \), but the parameter \( \theta \) is a random variable whose posterior distribution keeps changing as we see more and more \( x \)'s. It is shown that the optimal policy is again independent of the p's, and in the case where \( F \) is exponential, explicit results are obtained.

We conclude Chapter 3 with several models where the p's themselves are constrained to be in some set, and they must be chosen one at a time as the successive \( X \)'s are observed. We list two applications of these models. In the first, a government planner has a fixed amount of money \( N \) to spend on \( n \) incoming proposals of random value per dollar spent. As he observes the values of the successive proposals,
he must decide how much to spend on each of them. We allow the possibility that only $K_1$ may be spent on any one proposal, and the possibility that, for political reasons, at least $K_2$ must be spent on at least $M$ proposals, $M > 1$. The other application of this model is where a company must buy a quantity $N$ of some item in the next $n$ days, and the prices on the successive days are random variables. Here we allow the possibility that, for reasons of size or weight, only $K_3$ may be purchased on any one day. In these models we find optimal allocations which are closely connected to the results of the original assignment model.

In Chapter 4 we generalize in a different direction. We assume that all jobs come from the same distribution $F$, but that there is randomness in the number of jobs arriving and, possibly, in the times they arrive. We present two "timeless" models and two models where time is an explicit parameter.

In the "timeless" models we assume that at the beginning of the problem, the number of jobs which will ever arrive has a known distribution $\{q_k\}$. For $q_k$ arbitrary we consider only policies which assign to each job some man, if any remains. For $q_k = q^k(1-q)$, however, we allow the possibility of not assigning a job to anyone, if we so desire. In both cases optimal critical number policies are found which are independent of the $p$'s and which maximize the expected reward from $n$ men.

In the models where time is a parameter, however, we assume that successive offers arrive at random times according to a renewal process with given interarrival distribution $G$. Again we allow two
possibilities. First, we assume each job must be assigned to some man, if any remains, and we obtain the optimal policy rather easily. However, we then allow the possibility of not assigning a job, even if men remain. In this problem it becomes necessary to attach a discount factor $r(t)$ to a reward received at time $t$. (Otherwise we may never assign the men to any jobs.) We then treat this discounted problem in two stages. First we assume that all $n$ of the $p$'s are equal. This problem is then a generalized form of the house-selling model of Elfving [4] and of the disposal of an asset model of Karlin [6]. Once this identical men problem is solved, we use its solution and linearity of the reward function $p_x$ to find the optimal policy for the non-identical men problem. The optimal policies are again independent of the $p$'s.

Finally, in Chapter 5 we examine three more stochastic sequential assignment models which generalize the ideas in the original assignment model. In the first model, we assume for the first time that men can finish jobs and become reavailable. To be precise, we assume there are $n$ men with fixed $p$'s. At random times according to a renewal process, jobs arrive and their values are observed. A man is then assigned and he finishes the job in an exponentially distributed amount of time, the same for all men and all jobs. We show how to maximize the expected reward per unit time by means of Markov decision chain theory.

One possible application of this model is the following. Suppose we have $n$ repairmen, or repair teams, of various qualities. Machines, say airplanes, arrive at random times for repair, and once a repairman has been assigned, he works on the job for an exponential amount of time.
In Chapter 5, we list several interpretations of the $p$'s and the $x$'s in connection with this model, and we show their effects on the structure of the optimal policy.

The second model in Chapter 5 assumes that each time a man is assigned to a job, a new man replaces him in the set of available men. Thus the number of men available at any given time is $n$. However, the replacement is random, that is, the probability that a new man is a $p_i$ is $q_i$, independent of any decisions being made. Assuming that a job arrives at the beginning of each period of time, we show, by means of Markov decision chain theory, how to find the policy which maximizes the expected infinite-horizon discounted reward.

Besides the men and jobs interpretation, we have the following interpretation for this model. Suppose a machine arrives at the beginning of each period and requires a new part, say a motor, from one of the $n$ motors we have on hand. We issue one of these motors, and then we revamp the old motor which was in the machine some random amount, so that it can be added to our stockpile. In Chapter 5, we examine various interpretations of the $p$'s and the $x$'s in connection with this model, and we see what these interpretations imply about the structure of the optimal policy.

The final model we consider in Chapter 5 is somewhat different from the other models considered in this paper, but it is included since it falls under the heading of "stochastic sequential assignment" models. We assume that we have $n$ parts, say batteries, which deteriorate with age according to some function, or functions, $p_i(x)$. The interpretation of this $p_i$ function is that a car which uses a battery of
age \( x \) will survive at least \( t \) units of time on this battery with probability \( p_i(x)(1 - F(t)) \), where \( F \) is some distribution function, the same for all cars, and \( 0 \leq p_i(x) \leq 1 \). We assume that cars arrive at random times requesting these batteries, and we seek an issuing policy which maximizes the expected total lifetime of the cars we service. In particular, we seek conditions under which the optimal policy is FIFO (first-in-first-out, i.e., give worst away first) or LIFO (last-in-first-out, i.e., give best away first).

We examine three variations of this model: (1) the \( p \) function is the same for all batteries, but the initial ages are different; (2) the \( p \) function is the same for all batteries, and random replacements to inventory become available; and (3) no replacements are possible, the initial ages are the same, but the batteries deteriorate at different rates according to different \( p \) functions. In all three of these models we give simple conditions on the \( p \) functions which guarantee that LIFO or FIFO policies are optimal.

As the reader will probably suspect after reading this paper, or after reading only the introduction, the possible variations of the original assignment model are endless. Therefore, we do not mean to imply that this paper has covered all possibilities, or even that it has covered all possible models which allow "nice" results. We have simply tried to cover the variations which seem to be the most obvious and those which seem to be most amenable to application, as well as to solution.
CHAPTER 2
PRELIMINARY RESULTS

In this brief chapter, several preliminary results are set forth. We label these as lemmas since they will be useful in proving results of the later chapters. The first of these lemmas is a well-known result of Hardy, Littlewood, and Pólya [5], which will be referred to later on as Hardy's lemma.

Lemma 2.1: (Hardy's lemma). Let $p_1 \leq \cdots \leq p_n$ and $x_1 \leq \cdots \leq x_n$ be any numbers. Then for any permutation $\sigma(1), \ldots, \sigma(n)$ of $1, \ldots, n$, we have

$$\sum_{i=1}^{n} p_i x_i \geq \sum_{i=1}^{n} p_{\sigma(i)} x_{i}.$$

Proof: The proof follows by induction. For $n = 2$, the result is true since

$$(p_1 x_1 + p_2 x_2) - (p_2 x_1 + p_1 x_2) = (p_2 - p_1)(x_2 - x_1) \geq 0.$$

Assume the lemma is true for $n-1$, and let $\sigma(1), \ldots, \sigma(n)$ be any permutation of $1, \ldots, n$. If $\sigma(1) = 1$, then
\[ \sum_{i=1}^{n} p_{i}x_{i} - \sum_{i=1}^{n} p_{\sigma(i)}x_{i} = \sum_{i=2}^{n} p_{i}x_{i} - \sum_{i=2}^{n} p_{\sigma(i)}x_{i} \geq 0 \]

by the induction hypothesis. So assume \( \sigma(1) \neq 1 \), and let \( j \) be that integer such that \( \sigma(j) = 1 \). Now define a new permutation \( \tilde{\sigma} \) as follows. Let \( \tilde{\sigma}(1) = 1, \tilde{\sigma}(j) = \sigma(1), \) and \( \tilde{\sigma}(i) = \sigma(i), i \neq 1,j. \) Then

\[ \sum_{i=1}^{n} p_{\tilde{\sigma}(i)}x_{i} - \sum_{i=1}^{n} p_{\sigma(i)}x_{i} = (p_{\sigma(1)} - p_{1})(x_{j} - x_{1}) \geq 0, \]

and

\[ \sum_{i=1}^{n} p_{i}x_{i} - \sum_{i=1}^{n} p_{\tilde{\sigma}(i)}x_{i} = \sum_{i=2}^{n} p_{i}x_{i} - \sum_{i=2}^{n} p_{\tilde{\sigma}(i)}x_{i} \geq 0, \]

again by the induction hypothesis. This completes the induction,

Hardy's lemma may be strengthened as follows.

**Lemma 2.2**: Let \( r(p,x) \) be a real-valued function of 2 variables. Then

\[ (2.1) \quad \sum_{i=1}^{n} r(p_{i},x_{i}) - \sum_{i=1}^{n} r(p_{\sigma(i)},x_{i}) \geq 0, \]

for all \( p_{1} \leq \cdots \leq p_{n}, x_{1} \leq \cdots \leq x_{n} \), and permutations \( \sigma \), if and only if

\[ (2.2) \quad (r(p_{1},x_{1}) + r(p_{2},x_{2})) - (r(p_{2},x_{1}) + r(p_{1},x_{2})) \geq 0 \]

for all \( p_{1} \leq p_{2}, x_{1} \leq x_{2} \).
Furthermore, if \( r(p,x) \) is differentiable, then condition (2.2) is equivalent to

\[
\frac{\partial^2 r(p,x)}{\partial p \partial x} > 0 \quad \text{for all } p, x.
\]  

\[(2.3)\]  

**Proof:** (2.1) implies (2.2) trivially. (2.2) implies (2.1) by the same proof as was used to prove Hardy's lemma.

(2.3) implies (2.2) since

\[
0 \leq \int_{x_1}^{x_2} \int_{p_1}^{p_2} \frac{\partial^2 r(p,x)}{\partial p \partial x} \, dp \, dx = (r(p_1,x_1) + r(p_2,x_2)) - (r(p_2,x_1) + r(p_1,x_2))
\]

Finally, (2.2) implies (2.3) since

\[
0 \leq \lim_{p_1 \to p_2} \frac{1}{(p_2 - p_1)} \left\{ \lim_{x_1 \to x_2} \frac{1}{(x_2 - x_1)} ((r(p_2,x_2) - r(p_2,x_1)) - (r(p_1,x_2) - r(p_1,x_1))) \right\}
\]

\[
= \lim_{p_1 \to p_2} \frac{1}{(p_2 - p_1)} \left( \frac{\partial r(p_2,x_2)}{\partial x} - \frac{\partial r(p_1,x_2)}{\partial x} \right)
\]

Thus, if \( r(p,x) \) is differentiable, we need only show that (2.3) holds in order to show that (2.1) holds. Even if \( r(p,x) \) is not differentiable, we need only show that (2.1) holds for \( n = 2 \) in order to show that it holds for all \( n \).
Both Lemma 2.1 and Lemma 2.2 have obvious analogues when we are trying to minimize, not maximize

\[ \sum_{i=1}^{n} p_{\sigma(i)} x_i \quad \text{or} \quad \sum_{i=1}^{n} r(p_{\sigma(i)}, x_i). \]

In both cases the optimal permutation is \( \bar{\sigma}(i) = n-i+1 \). That is, if (2.2) holds, then

\[ \sum_{i=1}^{n} r(p_{\bar{\sigma}(i)}, x_i) \leq \sum_{i=1}^{n} r(p_{\sigma(i)}, x_i) \]

for all \( p_1 \leq \cdots \leq p_n, x_1 \leq \cdots \leq x_n \), and all permutations \( \sigma \).

On the other hand, if we have

(2.2') \[ r(p_1, x_1) + r(p_2, x_2) \leq r(p_2, x_1) + r(p_1, x_2) , \]

or equivalently

(2.3') \[ \frac{\partial^2 r(p, x)}{\partial p \partial x} \leq 0 , \]

when \( r(p, x) \) is differentiable, then \( \sum_{i=1}^{n} r(p_{\sigma(i)}, x_i) \) is maximized by taking \( \bar{\sigma}(i) = n-i+1 \), and is minimized by taking \( \bar{\sigma}(i) = i \).

The next lemma is an existence lemma which will be used several times throughout this paper to obtain the structure of optimal policies.
Lemma 2.3: Suppose $r(p, x)$ satisfies (2.2). Let $p_1 \leq \cdots \leq p_n$, $a_1, \ldots, a_n$ be any numbers. Then there exist numbers $b_1 \leq \cdots \leq b_{n-1}$, with $b_0 = -\infty$, $b_n = +\infty$, such that

$$\max_{1 \leq i \leq n} (r(p_i, x) + a_i) = r(p_k, x) + a_k$$

if and only if

$$b_{k-1} \leq x \leq b_k.$$

Proof:

$$\max_{1 \leq i \leq n} (r(p_i, x) + a_i) = r(p_k, x) + a_k$$

if and only if

$$(2.4) \quad r(p_k, x) - r(p_i, x) \geq a_i - a_k, \quad 1 \leq i \leq k-1,$$

and

$$(2.5) \quad r(p_i, x) - r(p_k, x) \leq a_k - a_i, \quad k+1 \leq i \leq n.$$

Now for $p_k \geq p_i$, $r(p_k, x) - r(p_i, x)$ is increasing in $x$ by (2.2), so (2.4) holds if and only if $x \geq \max_{1 \leq i \leq k-1} a_{ik}$, where $a_{ik}$ (possibly $+\infty$) satisfies

$$r(p_k, x) - r(p_i, x) \begin{cases} 
\geq a_i - a_k & \text{for } x \geq a_{ik}, \\
\leq a_i - a_k & \text{for } x \leq a_{ik}.
\end{cases}$$

Similarly, (2.5) holds if and only if
\[
x \leq \min_{k+1 \leq i \leq n} a_{ki} ,
\]

where \(a_{ki}\) (possibly \(+\infty\)) satisfies

\[
\begin{aligned}
    r(p_{i},x) - r(p_{k},x) &\leq a_{k} - a_{i} \quad \text{for } x \leq a_{ki} , \\
    &\geq a_{k} - a_{i} \quad \text{for } x \geq a_{ki} .
\end{aligned}
\]

Let \(b_{k} = \max_{1 \leq i \leq k-1} a_{ik}\), \(\bar{b}_{k} = \min_{k+1 \leq i \leq n} a_{ki}\), where \(b_{1} = -\infty\), \(\bar{b}_{n} = +\infty\). Then we have shown that

\[
\max_{1 \leq i \leq n} \left( r(p_{i},x) + a_{i} \right) = r(p_{k},x) + a_{k}
\]

if and only if

\[
b_{k} \leq x \quad \text{and} \quad x \leq \bar{b}_{k} .
\]

Also, we have for each \(i_{1}, i_{2}\) (\(1 \leq i_{1} < i_{2} \leq n\)),

\[
\bar{b}_{i} = \min_{i_{1}+1 \leq i \leq n} a_{i}, i_{1} \leq a_{i}, i_{2} \leq \max_{1 \leq i \leq i_{2}-1} a_{i}, i_{2} = \frac{b_{i}}{i_{2}} .
\]

This means it can never happen that there exists \(x_{1} < x_{2}\) and a pair \(i_{1} < i_{2}\), such that

\[
r(p_{i_{1}},x_{2}) + a_{i_{1}} = \max_{1 \leq i \leq n} (r(p_{i},x_{2}) + a_{i})
\]

and

15
\[ r(p_{i_2}, x_1) + a_{i_2} = \max_{1 \leq i \leq n} (r(p_i, x_1) + a_i). \]

This is sufficient to prove the lemma.

To actually find \( b_1, \ldots, b_{n-1} \) of the lemma, proceed as follows. Note that it is possible for \( r(p_k, x) + a_k \) to equal \( \max_{1 \leq i \leq n} (r(p_i, x) + a_i) \) for some \( x \) if and only if \( \frac{b_i}{x} \leq \frac{b_k}{x} \). When this is the case, take \( b_{k-1} = \frac{b_k}{x} \) and \( b_k = \frac{b_k}{x} \).

The same type of statements which follow Lemma 2.2 apply here also. For example, suppose \( r(p, x) \) satisfies \( (2.2') \) and we are interested in maximizing \( r(p_i, x) + a_i \). Then there exist numbers \( b_1' \leq \cdots \leq b_{n-1}' \), \( b_0' = -\infty, b_n' = +\infty \), such that

\[ \max_{1 \leq i \leq n} (r(p_i, x) + a_i) = r(p_k, x) + a_k \]

if and only if \( b_{n-k}' \leq x \leq b_{n-k+1}' \).

The following lemma will be useful when we try to show that the optimal policy is or is not independent of the values of the men, the p's.

**Lemma 2.4:** Let \( A \) be any subset of \( \mathbb{R}^n \) such that if \( x = (x_1, \ldots, x_n) \in A \), then \( x_1 + \cdots + x_n = c \), where \( c \) is a fixed constant. Also let \( B = \{ p \equiv (p_1, \ldots, p_n) \mid p_1 \leq \cdots \leq p_n \} \). Then a necessary and sufficient condition that \( \bar{x} \) maximize \( px = \sum_{i=1}^{n} p_i x_i \), for all \( x \in A, p \in B \), is that
\[(2.6) \quad \sum_{i=1}^{k} (\tilde{x}_i - x) \leq 0, \quad 1 \leq k \leq n-2, \quad \text{and} \quad \tilde{x}_n - x_n \geq 0, \]

for all \( x \in A \).

Note that since \( x \in A \) implies \( \sum_{i=1}^{n} x_i = c \), (2.6) is equivalent to

\[(2.7) \quad \sum_{i=k}^{n} (\tilde{x}_i - x_i) \geq 0, \quad 2 \leq k \leq n, \text{ for all } x \in A. \]

**Proof:** (Necessity) If \( \tilde{p} x \) is to be a maximum for all \( p \in B, x \in A \), it must be a maximum for \( p^1 = (0, 0, \ldots, 0, 1), p^2 = (0, 0, \ldots, 0, 1, 1), \) and \( p^{n-1} = (0, 1, \ldots, 1) \), which means condition (2.7) is necessary.

(Sufficiency) Suppose (2.6) holds for some \( \tilde{x} \in A \), and let \( x \in A, p \in B \). Then we have

\[
\tilde{p} x - px = \sum_{i=1}^{n} p_i (\tilde{x}_i - x_i) \\
= \sum_{i=1}^{n-2} p_i (\tilde{x}_i - x_i) + p_{n-1} (c - \sum_{i=1}^{n-2} \tilde{x}_i - \tilde{x}_n - c + \sum_{i=1}^{n-2} x_i + x_n) \\
+ p_n (\tilde{x}_n - x_n) \\
= \sum_{i=1}^{n-2} (p_i - p_{n-1}) (\tilde{x}_i - x_i) + (p_n - p_{n-1}) (\tilde{x}_n - x_n) \\
= \sum_{i=1}^{n-2} \left( \sum_{j=1}^{i} (p_j - p_{j+1}) (\tilde{x}_i - x_i) + (p_n - p_{n-1}) (\tilde{x}_n - x_n) \right) \\
= \sum_{j=1}^{n-2} (p_j - p_{j+1}) \left( \sum_{i=1}^{j} (\tilde{x}_i - x_i) \right) + (p_n - p_{n-1}) (\tilde{x}_n - x_n) \geq 0
\]

by (2.6) and the fact that \( p \in B \). This completes the proof.
CHAPTER 3

ASSIGNMENT PROBLEMS WITH EXACTLY n JOBS ARRIVING

This chapter deals with the most direct generalizations of the original assignment model of Derman, Lieberman, and Ross [3]. The common aspects of all models dealt with in this chapter are the following. A fixed number of jobs, say n, is going to arrive in the future. These arrive sequentially and take on random values. There are also men on hand, with known values, who are able to do all or some of the incoming jobs. When a job arrives, its value is observed, and a decision is made as to which of the available men to assign. If a "p" man is assigned to an "x" job, a reward px is obtained. We seek a policy which maximizes the expected reward from the n jobs.

Although it is not necessary in all of the results of this chapter, we assume throughout that the values of the men, the p's, and the values of the jobs, the x's, are all non-negative. There are three main reasons for this assumption. First, negative values for either p's or x's have no apparent meaning in the present context of men and jobs, or in any other practical applications of the theory we will consider. Second, we will often encounter a case where there are more than n men. In this case it is not necessarily optimal to use only the n highest p's unless the x's are non-negative, and the theory is certainly simplified if we know we need only consider the n highest
p's. Third, if there are \( m < n \) men, it is convenient to add \( n-m \)
"fake" men, with p's equal to 0, to the \( m \) real men. In this case, we would intuitively want to use the real men for the best jobs, which is the case only if the p's are non-negative.

As mentioned before, however, the non-negativity restrictions on the p's and the x's are not always necessary, and it should be clear from the proofs which of the following results go through without these restrictions.

3.1. Successive Job Distributions Determined by a Markov Chain.

The result of Derman, Lieberman, and Ross has already been stated. We now present a straightforward generalization of their model. Suppose jobs come from \( r < \infty \) distributions \( F_1, \ldots, F_r \), with finite means. A job comes in with a value \( x \geq 0 \) from distribution \( F_i \), and then the next job comes from distribution \( F_j \) with probability \( q_{ij} \).

Thus the indices of the distributions of the successive jobs form a Markov chain. To be more rigorous, let \( T_1, T_2, \ldots \) be the successive states of a Markov chain, where \( 1, 2, \ldots, r \) are the possible states, and

\[
P(T_{k+1} = j | T_k = i) = q_{ij}.
\]

Associated with each state \( i \) is a sequence of i.i.d. non-negative random variables \( X_1^i, X_2^i, \ldots \) from distribution \( F_i \). We assume the \( X^i \)'s are independent of the \( X^j \)'s, and that both of these are independent.
of the $T$'s, for all states $i \neq j$. Now define $X_j$ by

$$X_j = \sum_{i=1}^{r} X_j^i 1(T_j=i) \quad {\dagger}$$

Then $X_j$ is the value of the $j$th job. (This all tacitly assumes an initial distribution for the Markov chain, which we regard as arbitrary, but known.)

We assume there are $n$ men with values $p_1 \leq \cdots \leq p_n$ and that each man is able to do each job. The following theorem gives the optimal assignment policy.

**Theorem 3.1:** Assume there are $n$ men and $n$ jobs left. Then there exist numbers $0 \leq a_{1,n}^i \leq \cdots \leq a_{n-1,n}^i$, along with $a_0,n^i = 0$ and $a_{n,n}^i = +\infty$, such that if the next job arrives from distribution $F_i$ with value $x$, it is best to assign man $p_j$ if and only if $a_{j-1,n}^i < x \leq a_{j,n}^i$. These $a_{j,n}^i$'s are independent of the $p$'s.

Furthermore, $a_{j,n}^i$ has the following interpretation. In an $n$ man problem, suppose the first job came from distribution $F_i$, was assigned, and that $n-1$ jobs and $n-1$ men $p_1 \leq \cdots \leq p_{n-1}$ are now left. Then $a_{j,n}^i$ is the expected value, under an optimal policy, of the job to which man $p_j$ is assigned.

**Proof:** The proof proceeds by induction. For $n=1$, the theorem is trivially true.

Assume the theorem is true for $n-1$. We introduce the following notation. Let

$${\dagger}$$ Throughout this paper $1_A$ is used to denote the indicator of the set $A$. 
\( f^i(p_1, \ldots, p_n) = \text{optimal expected reward from men } p_1 \leq \cdots \leq p_n \)
when there are \( n \) jobs left, given that the previous job was from distribution \( F_i \).

and

\( f(p_1, \ldots, p_n | x, i) = \text{optimal expected reward from men } p_1 \leq \cdots \leq p_n \)
when there are \( n \) jobs left, given that the first of these is from distribution \( F_i \) with value \( x \).

Then we have

\[
(3.1) \quad f(p_1, \ldots, p_n | x, i) = \max_{1 \leq k \leq n} \left( p_k x + f^i(p_1, \ldots, p_{k-1}, p_{k+1}, \ldots, p_n) \right).
\]

However, by the inductive hypothesis, it follows that the optimal policy for an \( n-1 \) job problem is independent of the \( n-1 \) \( p \)'s.

Hence, we may define \( a^i_{j,n} \) as the expected value, under the optimal policy, of the job to which the \( j \)th smallest \( p \) is assigned in the \( n-1 \) job problem, given that the previous job was from distribution \( F_i \).

Then the total expected reward from that policy is

\[
(3.2) \quad f^i(\tilde{p}_1, \ldots, \tilde{p}_{n-1}) = \sum_{j=1}^{n} \tilde{p}_j a^i_{j,n},
\]
for every \( \tilde{p}_1 \leq \cdots \leq \tilde{p}_{n-1} \). Furthermore, since \( a^i_{j,n} \) is independent of the \( p \)'s and other policies may be obtained by permuting the \( p \)'s, any
sum of the form \( \sum_{j=1}^{n-1} p_{\sigma(j)} a_{j,n}^i \), where \( \sigma(1), \ldots, \sigma(n-1) \) is a permutation of 1, \( \ldots, n-1 \), can be obtained for the total expected reward from the \( n-1 \) job problem. By Hardy's lemma this means that

\[
\frac{a_{1,n}^i}{a_{1,n}^i} \leq \cdots \leq \frac{a_{n-1,n}^i}{a_{n-1,n}^i},
\]

and, of course, \( 0 \leq a_{1,n}^i \), since all job values are non-negative.

Using (3.2) we may write (3.1) as

\[
f(p_1, \ldots, p_n|x, i) = \max_{1 \leq k \leq n} \left( p_k x + \sum_{j=1}^{k-1} p_j a_{j,n}^i + \sum_{j=k+1}^{n} p_j a_{j-1,n}^i \right).
\]

We may again use Hardy's lemma and (3.3) to obtain

\[
f(p_1, \ldots, p_n|x, i) = p_k x + \sum_{j=1}^{k-1} p_j a_{j,n}^i + \sum_{j=k+1}^{n} p_j a_{j-1,n}^i,
\]

where \( k^* \) is such that \( a_{k^*-1,n}^i < x \leq a_{k^*,n}^i \), with \( a_{0,n}^i = 0 \), \( a_{n,n}^i = +\infty \). Hence, the first choice in an \( n \) job problem, when the first job comes from distribution \( F_i \) and how value \( x \), is to use man \( p_j \) if and only if \( a_{j-1,n}^i < x \leq a_{j,n}^i \). This completes the induction.

**Corollary 3.1:** We may calculate the \( a_{j,n}^i \)'s by the following recursion:

\[
a_{j,n+1}^i = \sum_{i=1}^{r} \left\{ a_{j,n}^i \int x F_i(dx) + a_{j-1,n}^i F_i(a_{j-1,n}^i) + a_{j,n}^i (1-F_i(a_{j,n}^i)) \right\} q_{ki} \quad \text{for} \ 1 \leq k \leq r.
\]
Proof: We suppose that in an \( n+1 \) job problem, the first job comes from distribution \( F_k \). We then condition on the distribution and the value of the second job and use the interpretation of \( a^k_{j,n+1} \) given in Theorem 3.1. The result then follows immediately.

Special cases: (1) The result of Derman, Lieberman, and Ross is a special case of the above by taking \( r = 1 \).

(2) Suppose the actual values of the jobs form a Markov chain with \( r \) states. That is, a visit to state \( i \) at time \( k \) means that the \( k \)th job takes on value \( x_i \), where \( x_1, \ldots, x_r \) are the set of possible values. This fits into the above framework by taking \( F_i \) to be degenerate at the point \( x_i \). The recursion formula for the \( a^i_{j,n} \)'s becomes

\[
a^k_{j,n+1} = \sum_{i=1}^{r} \left\{ a^i_{j-1,n} \left( x_i < a^i_{j,n} \right) + a^i_{j-1,n} \left( x_i \leq a^i_{j-1,n} \right)
+ a^i_{j,n} \left( x_i > a^i_{j,n} \right) \right\} q_{ki}.
\]

In this case, if there are \( n \) jobs left and the next one takes on a value \( x_i \), it is best to assign the \( j \)th smallest man if and only if \( a^i_{j-1,n} < x_i \leq a^i_{j,n} \).

Limiting results:

This section deals with finding \( \lim_{n \to \infty} a^k_{[\pi n],n+1} \) for fixed \( k \) and \( 0 < \pi < 1 \), where \( [\cdot] \) denotes the greatest integer. In order to do
this, we will first find
\[
\lim_{n \to \infty} \frac{1}{n} \sum_{j=[nr]+1}^{n} a_{j,n+1}^k
\]
for fixed \( k \).

Suppose \( 0 < \pi < 1 \) and that in the \( n \) job problem, \( n-[nr] \)
of the \( p \)'s are \( 1 \) and the rest of the \( p \)'s are \( 0 \). Then if the previous
job was from distribution \( F_k \), the optimal reward from these \( p \)'s is
\[
\sum_{j=[nr]+1}^{n} a_{j,n+1}^k.
\]

We use this fact to find the limiting behavior of the average expected
reward per job,
\[
\frac{1}{n} \sum_{j=[nr]+1}^{n} a_{j,n+1}^k.
\]

We first make several definitions. Suppose the Markov chain
governing the successive distributions of the jobs is positive recurrent
and irreducible, and let \( \rho_1, \ldots, \rho_r \) be the steady-state probabilities.
(In case the chain is periodic, let
\[
\rho_i = \lim_{n \to \infty} \frac{1}{n} \sum_{j=1}^{n} q_{ki}^j.
\]

Suppose the chain starts in state \( k \), and let \( \ell_i(n) \) be the number of
visits to state \( i \) by time \( n \). Next, let \( \nu_1(i), \nu_2(i), \ldots \) denote
the times at which the Markov chain is in state \( i \), and for any \( a \)
and \( b, 1 \leq i \leq r \), let
\[ s_i(n) \equiv s_i(n; a, b) = \min \left( m : \sum_{j=1}^{m} l \left( Y_{j}^i(n) > a \right) \geq \left\lceil nb \right\rceil \right). \]

Define

\[ t_i(n) \equiv t_i(n; a, b) = \min(\ell_i(n), s_i(n; a, b)). \]

Also, for any fixed constant \( a \), define

\[ Y_j^i(a) = \mathbb{1}_{X_j^i} \left( X_j^i > a \right) \quad \text{and} \quad Z_j^i(a) = \mathbb{1}_{X_j^i} \left( X_j^i > a \right). \]

Finally, we use the following convention throughout this chapter. For any \( 0 \leq \sigma_i \leq 1, 1 \leq i \leq r \), let \( F_i^{-1}(\sigma_i) \) be any of the numbers \( \tau \) which satisfy \( F_i(\tau) = \sigma_i \). (The ambiguity here will not cause any difficulty.)

Using the above notation, we are now in a position to prove several lemmas.

**Lemma 3.1:** Let \( 0 \leq \sigma_i \leq 1 \) be any number for which the set \( \{ \tau | F_i(\tau) = \sigma_i \} \) is non-empty, and let \( s_i(n) = s_i(n; F_i^{-1}(\sigma_i), \rho_i(1-\sigma_i)) \). Then

\[ \lim_{n \to \infty} \frac{1}{n} E t_i(n) = \rho_i, \]

independent of the starting state \( k \).
Proof: Since $0 \leq t_i(n)/n \leq l_i(n)/n \leq 1$, it suffices by the bounded convergence theorem, to show that

$$\frac{t_i(n)}{n} \longrightarrow \rho_i \quad \text{in prob. as } n \to \infty.$$ 

Let $\epsilon > 0$ by arbitrary. Then

$$P\left(\frac{t_i(n)}{n} \geq \rho_i + \epsilon\right) \leq P\left(\frac{l_i(n)}{n} \geq \rho_i + \epsilon\right) \longrightarrow 0 \quad \text{as } n \to \infty,$$

by a well-known Markov chain theorem.

Next, we have

$$(3.4) \quad P\left(\frac{t_i(n)}{n} \leq \rho_i - \epsilon\right) \leq P\left(\frac{l_i(n)}{n} \leq \rho_i - \epsilon\right) + P\left(\frac{s_i(n)}{n} \leq \rho_i - \epsilon\right).$$

As above,

$$P\left(\frac{l_i(n)}{n} \leq \rho_i - \epsilon\right) \longrightarrow 0 \quad \text{as } n \to \infty.$$

For the other term in (3.4), we use the fact that $X^i_{\nu_1(i)}, X^i_{\nu_2(i)}, \ldots$ are distributed exactly like $X^i_1, X^i_2, \ldots$ to obtain
\[ p \left( \frac{s_i(n)}{n} \leq \rho_i - \epsilon \right) \]

\[ = P \left( s_i(n) \leq \lfloor n(\rho_i - \epsilon) \rfloor \right) \]

\[ = P \left( \sum_{j=1}^{\lfloor n(\rho_i - \epsilon) \rfloor} Z_{F_{-1}(\sigma_i)}^{i}(F^{-1}_{i}(\sigma_i)) \geq \lfloor n\rho_i(1 - \sigma_i) \rfloor \right) \]

\[ = P \left( \frac{\sum_{j=1}^{\lfloor n(\rho_i - \epsilon) \rfloor} Z_{F_{-1}(\sigma_i)}^{i}(F^{-1}_{i}(\sigma_i))}{\lfloor n(\rho_i - \epsilon) \rfloor} \geq \frac{\lfloor n\rho_i(1 - \sigma_i) \rfloor}{\lfloor n(\rho_i - \epsilon) \rfloor} \right) , \]

and this last term \( \to 0 \) as \( n \to \infty \). This follows directly from the weak law of large numbers (WLLN), since \( Z_{F_{-1}(\sigma_i)}^{i}(F^{-1}_{i}(\sigma_i)) \), \( Z_{F_{-1}(\sigma_i)}^{2}(F^{-1}_{i}(\sigma_i)) \), \ldots are i.i.d. with mean \( 1 - \sigma_i \), and

\[
\lim_{n \to \infty} \frac{\lfloor n\rho_i(1 - \sigma_i) \rfloor}{\lfloor n(\rho_i - \epsilon) \rfloor} = \frac{1 - \sigma_i}{1 - \epsilon/\rho_i} > 1 - \sigma_i .
\]

This completes the proof.

**Corollary 3.2**: (Special case when \( r = 1 \)) Let \( X_1, X_2, \ldots \) be i.i.d. random variables from a distribution \( F \), and let \( 0 \leq \sigma \leq 1 \) be any number for which \( \{ \tau | F(\tau) = \sigma \} \) is non-empty. Define \( s(n) \) by

\[ s(n) = \min(m | \sum_{j=1}^{m} 1(X_j > F^{-1}(\sigma)) \geq \lfloor n(1-\sigma) \rfloor) \]

and let \( t(n) = \min(n, s(n)) \). Then

\[ \lim_{n \to \infty} \frac{1}{n} Et(n) = 1 . \]
Proof: This is Lemma 3.1 with $r = 1$.

Lemma 3.2: Let $0 \leq \sigma_i \leq 1$, $1 \leq i \leq r$ be any numbers for which the sets $\{ \tau \mid F_i(\tau) = \sigma_i \}$ are all non-empty, and let $\ell_i(n)$, $s_i(n)$, and $t_i(n)$ be as in Lemma 3.1. Then for each $1 \leq i \leq r$,

$$\lim_{n \to \infty} \frac{1}{n} E\left( \sum_{j=1}^{t_i(n)} \frac{Y_i^{(i)}}{V_j(i)} \left( F^{-1}_i(\sigma_i) \right) \right) = \rho_i \int_{F^{-1}_i(\sigma_i)}^{\infty} xF_i(dx).$$

Proof: Define $\Lambda_\ell$ to be the set of possible realizations of $v_1(i), v_2(i), \ldots$ when $\ell_i(n) = \ell$. Also, for $\nu = (v_1, v_2, \ldots) \in \Lambda_\ell$, let

$$A(\ell, \nu) = \{ \ell_i(n) = \ell; v_j(i) = v_j, \ 1 \leq j < \infty \}.$$

Then we have

$$\frac{1}{n} E\left( \sum_{j=1}^{t_i(n)} \frac{Y_i^{(i)}}{V_j(i)} \left( F^{-1}_i(\sigma_i) \right) \right) = \frac{1}{n} \sum_{\ell} \sum_{\nu \in \Lambda_\ell} E\left( \sum_{j=1}^{t_i(n)} \frac{Y_i^{(i)}}{V_j(i)} \left( F^{-1}_i(\sigma_i) \right) | A(\ell, \nu) \right) P(A(\ell, \nu)).$$

Now on the set $A(\ell, \nu)$, $t_i(n)$ is a finite stopping rule for the i.i.d. $v_j^{i}$'s, so by Wald's equation and independence, the above expression is

$$\frac{1}{n} \sum_{\ell} \sum_{\nu \in \Lambda_\ell} E(t_i(n) | A(\ell, \nu)) \ E(Y_i^{(i)}(F^{-1}_i(\sigma_i))) \ P(A(\ell, \nu))$$

$$= \frac{1}{n} \ E(t_i(n)) \int_{F^{-1}_i(\sigma_i)}^{\infty} xF_i(dx) \longrightarrow \rho_i \int_{F^{-1}_i(\sigma_i)}^{\infty} xF_i(dx) \quad \text{as} \ n \to \infty,$$

by Lemma 3.1.
Lemma 3.3: Let $0 \leq \sigma_i \leq 1$, $1 \leq i \leq r$ by any numbers for which the sets \{\(\tau|F_i(\tau) = \sigma_i\)\} are all non-empty, and which satisfy

\[
\sum_{i=1}^{r} \rho_i (1 - \sigma_i) = 1 - \pi .
\]

(Since \(\sum_{i=1}^{r} \rho_i = 1\), this is equivalent to

\[
\sum_{i=1}^{r} \rho_i \sigma_i = \pi .
\]

Then we have

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{j=\lceil n\pi \rceil + 1}^{n} a_{j, n+1} \geq \sum_{i=1}^{r} \rho_i \int_{p_i^{-1}(\sigma_i)}^{\infty} x F_i(dx) .
\]

Proof: Consider the infeasible policy which assigns at most \(\lfloor n\rho_i (1 - \sigma_i) \rfloor\) of the p's equal to 1 to jobs from distribution \(F_i\), and, with this restriction, assigns a 1 to every job from distribution \(F_i\) with a value \(x > F_i^{-1}(\sigma_i)\). The reason this policy is infeasible is that the 1's may not all be assigned. However, since

\[
\sum_{i=1}^{r} \lfloor n\rho_i (1 - \sigma_i) \rfloor \leq \sum_{i=1}^{r} n\rho_i (1 - \sigma_i) = n(1 - \pi) \leq n - \lfloor n\pi \rfloor ,
\]

this policy will never try to assign more 1's than are available.

Letting \(R_n\) be the total reward from this policy and using the notation of Lemmas 3.1 and 3.2, we have
\[ \frac{1}{n} E(R_n) = \frac{1}{n} E \left( \sum_{i=1}^{t_i(n)} \sum_{j=1}^{v_i(i)} \chi_j(i)(F_i^{-1}(\tau_i)) \right) \xrightarrow{n \to \infty} \sum_{i=1}^{t_i(n)} \rho_i \int_{F_i^{-1}(\tau_i)}^\infty x F_i(dx) \]

as \( n \to \infty \), by Lemma 3.2.

Now we compare this infeasible policy with a similar, but feasible, non-optimal policy. Namely, assign a 1 to every job from distribution \( F_i \) with value \( x > F_i^{-1}(\tau_i) \) until there are as many jobs left as there are 1's remaining, if this ever happens. After this, assign a 1 to every job. Let \( \tilde{R}_n \) be the total reward from this policy and let \( D_n = \tilde{R}_n - R_n \). Then \( D_n \geq 0 \), since more 1's are assigned under the feasible policy and the \( x \)'s are non-negative. Therefore, since \( E(\tilde{R}_n) \) is non-optimal, we have

\[ \sum_{i=1}^{t_i(n)} \rho_i \int_{F_i^{-1}(\tau_i)}^\infty x F_i(dx) = \lim_{n \to \infty} \frac{1}{n} E(R_n) \]

\[ \leq \lim_{n \to \infty} \frac{1}{n} E(\tilde{R}_n) \]

\[ \leq \lim_{n \to \infty} \frac{1}{n} \sum_{j=[n\tau]+1}^{n} e_{j,n+1} \cdot \]

(\text{It is easy to show that } \frac{1}{n} \left| E D_n \right| \to 0 \text{ as } n \to \infty, \text{ even if the } x \text{'s are not necessarily non-negative, so that this is no restriction.})

We now compare the optimal policy with an infeasible, but certainly better-than-optimal, policy in order to obtain an upper bound. First we prove the following existence lemma.
Lemma 3.4: Let $F_1, \ldots, F_r$ be continuous distribution functions. Then for any $0 < \pi < 1$, there exist fractions $0 \leq \sigma_i \leq 1$, $1 \leq i \leq r$, satisfying

$$\sum_{i=1}^{r} \rho_i \sigma_i = \pi \quad \text{and} \quad F_1^{-1}(\sigma_1) = \cdots = F_r^{-1}(\sigma_r).$$

Proof: For $b$ sufficiently large, the solutions $\sigma_1, \ldots, \sigma_r$ to $F_1^{-1}(\sigma_1) = \cdots = F_r^{-1}(\sigma_r) = b$ satisfy

$$\sum_{i=1}^{r} \rho_i \sigma_i > \pi.$$  

Similarly, for $b$ sufficiently small, the solutions $\sigma_1, \ldots, \sigma_r$ to $F_1^{-1}(\sigma_1) = \cdots = F_r^{-1}(\sigma_r) = b$ satisfy

$$\sum_{i=1}^{r} \rho_i \sigma_i < \pi.$$  

Since each $\sigma_i$ is a continuous function of $b = F_i^{-1}(\sigma_i)$, this means there exists $b^*$ satisfying

$$\sum_{i=1}^{r} \rho_i \sigma_i = \pi \quad \text{and} \quad F_1^{-1}(\sigma_1) = \cdots = F_r^{-1}(\sigma_r) = b^*,$$  

so that $\sigma_i = F_i(b^*)$. This prove the lemma.

Lemma 3.5: Suppose $F_1, \ldots, F_r$ are continuous and $0 \leq \sigma_i \leq 1$, $1 \leq i \leq r$ are the fractions whose existence Lemma 3.4 guarantees. Then we have
Proof: Recall that $X_j$ is the value of the $j$th job, and define $X_{(1)}, \ldots, X_{(n)}$ to be the order statistics associated with $X_1, \ldots, X_n$. Then we have

$$\lim_{n \to \infty} \frac{1}{n} \sum_{j=[n\nu]+1}^{n} a_{j,n+1} = \frac{1}{n} \sum_{i=1}^{r} \rho_i \int_{F_i^{-1}(\sigma_i)}^{\infty} x F_i(dx) .$$

since no sequential assignment can do better than assigning the 1's to the highest order statistics.

Now we show that

$$\lim_{n \to \infty} \frac{1}{n} \frac{1}{n} \sum_{j=[n\nu]+1}^{n} X_{(j)} - \frac{1}{n} \sum_{i=1}^{r} \sum_{j=1}^{n \nu} Y_{j,i}(F_i^{-1}(\sigma_i)) \leq 0 .$$

This will be sufficient to prove the lemma since

$$\lim_{n \to \infty} \frac{1}{n} \frac{1}{n} \sum_{i=1}^{r} \sum_{j=1}^{n \nu} Y_{j,i}(F_i^{-1}(\sigma_i)) = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{r} \int_{F_i^{-1}(\sigma_i)}^{\infty} x F_i(dx) .$$

Define $N_i(n)$ by

$$N_i(n) = \sum_{j=1}^{n \nu} Y_{j,i}(F_i^{-1}(\sigma_i)) .$$
Then the expression in (3.5) may be written as

\[
\frac{1}{n} \left\{ \sum_{i=1}^{r} N_i(n) \leq n - \lfloor \frac{n}{n+1} \rfloor \right\} \text{ terms, } \forall i \leq F_1^{-1}(\sigma_i)
\]

\[
+ \int \left( \sum_{i=1}^{r} N_i(n) - n + \lfloor \frac{n}{n+1} \rfloor \right) \text{ terms, } \forall i \leq F_1^{-1}(\sigma_i)
\]

\[
\leq \frac{1}{n} F_1^{-1}(\sigma_i) \cdot (n - \lfloor \frac{n}{n+1} \rfloor - \sum_{i=1}^{r} N_i(n))
\]

\[
= F_1^{-1}(\sigma_i) \left\{ \frac{1}{n} (n - \lfloor \frac{n}{n+1} \rfloor - \sum_{i=1}^{r} \lfloor n \rho_i \rfloor (1 - \sigma_i)) \right\} \rightarrow 0 \quad \text{as } n \rightarrow \infty,
\]

since \( \sum_{i=1}^{r} \rho_i (1 - \sigma_i) = 1 - \pi. \)

Thus the inequality in (3.5) holds, and the proof is complete.

\[\text{Theorem 3.2: Let } F_1, \ldots, F_r \text{ and } \sigma_1, \ldots, \sigma_r \text{ be as in Lemma 3.5. Then}
\]

\[
\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=\lfloor \frac{n}{n+1} \rfloor + 1}^{n} \sum_{i=1}^{k} a_{i,j,n+1} = \sum_{i=1}^{r} \rho_i \int_{F_1^{-1}(\sigma_i)}^{\infty} x F_i(dx).
\]

\[\text{Proof: The proof follows directly from Lemmas 3.3 and 3.5.}
\]

\[\text{Corollary 3.3: The solutions } 0 \leq \sigma_i \leq 1, 1 \leq i \leq r \text{ of}
\]

\[
\sum_{i=1}^{r} \rho_i \sigma_i = \pi \quad \text{and} \quad F_1^{-1}(\sigma_1) = \cdots = F_r^{-1}(\sigma_r)
\]
also maximize

\[ \sum_{i=1}^{r} \rho_i \int_{F_i^{-1}(\sigma_i)}^{\infty} xF_i(dx) , \]

subject to

\[ \sum_{i=1}^{r} \rho_i \sigma_i = \pi, \quad 0 \leq \sigma_i \leq 1 . \]

**Proof:** Suppose \( 0 \leq \sigma_i \leq 1 \) satisfy \( \sum_{i=1}^{r} \rho_i \sigma_i = \pi \) and that

\[ F_i^{-1}(\sigma_i) = F_i^{-1}(\sigma_i) \]

for some \( i_1 \neq i_2 \). Let \( \epsilon_1, \epsilon_2 > 0 \) be such that \( 0 \leq \sigma_{i_1} + \epsilon_1 \leq 1, 0 \leq \sigma_{i_2} - \epsilon_2 \leq 1 \), and

\[ \sum_{i_1, i_2} \rho_{i_1} \sigma_{i_1} + \rho_{i_1} (\sigma_{i_1} + \epsilon_1) + \rho_{i_2} (\sigma_{i_2} - \epsilon_2) = \pi . \]

Then we have \( F_i^{-1}(\sigma_{i_1} + \epsilon_1) > F_i^{-1}(\sigma_{i_2} - \epsilon_2) \). Furthermore,

\[ \sum_{i=1}^{r} \rho_i \int_{F_i^{-1}(\sigma_i)}^{\infty} xF_i(dx) \leq \left\{ \sum_{i_1, i_2} \rho_{i_1} \int_{F_i^{-1}(\sigma_i)}^{\infty} xF_i(dx) \right\} 
+ \rho_{i_1} \int_{F_i^{-1}(\sigma_{i_1} + \epsilon_1)}^{\infty} xF_i(dx) + \rho_{i_2} \int_{F_i^{-1}(\sigma_{i_2} - \epsilon_2)}^{\infty} xF_i(dx) 
= \rho_{i_1} \int_{F_i^{-1}(\sigma_{i_1} + \epsilon_1)} xF_i(dx) + \rho_{i_2} \int_{F_i^{-1}(\sigma_{i_2} - \epsilon_2)} xF_i(dx) 
\geq \rho_{i_1} \epsilon_1 F_i^{-1}(\sigma_{i_1}) - \rho_{i_2} \epsilon_2 F_i^{-1}(\sigma_{i_2}) = 0 , \]

since \( \rho_{i_1} \epsilon_1 - \rho_{i_2} \epsilon_2 = 0 \) and \( F_i^{-1}(\sigma_{i_1}) = F_i^{-1}(\sigma_{i_2}) \).
This is sufficient to prove the corollary.

This corollary shows in perhaps a more intuitive way why the numbers $\sigma_1, \ldots, \sigma_r$ of Lemma 3.4 are associated with the optimal assignment policy.

**Corollary 3.4:** Under the same assumptions as in Theorem 3.2, we have

$$
\lim_{n \to \infty} \sum_{j=1}^{[n\xi]} a_j^{k, n+1} = \sum_{i=1}^{r} \rho_i \frac{F_i^{-1}(\sigma_i)}{ \int_0^{\infty} x F_i(dx) }
$$

**Proof:** By taking all the $p$'s equal to 1, we obtain

$$
\frac{1}{n} \sum_{j=1}^{n} a_j^{k, n+1} = \frac{1}{n} E\left( \sum_{j=1}^{n} X_j \right)
$$

$$
= \frac{1}{n} E\left( \sum_{i=1}^{r} \sum_{j=1}^{n} 1(T_j = i) X_j^i \right)
$$

$$
= \frac{1}{n} \sum_{i=1}^{r} \sum_{j=1}^{n} q_{ki} \int_0^{\infty} x F_i(dx)
$$

$$
\rightarrow \sum_{i=1}^{r} \rho_i \int_0^{\infty} x F_i(dx) \quad \text{as } n \to \infty.
$$

Combining this with Theorem 3.2 gives the result of the corollary.

**Corollary 3.4:** Under the same assumptions as in Theorem 3.2, we have

$$
\lim_{n \to \infty} \frac{1}{n} E\left( \sum_{j=\lceil n\pi \rceil + 1}^{n} X(j) \right) = \sum_{i=1}^{r} \rho_i \int_0^{\infty} x F_i^{-1}(\sigma_i) F_i^{-1}(\sigma_i)
$$

and

$$
(3.6) \quad \lim_{n \to \infty} \frac{1}{n} E\left( \sum_{j=\lceil n\pi \rceil + 1}^{n} X(j) \right) = \sum_{i=1}^{r} \rho_i \int_0^{\infty} x F_i(dx)
$$

35
(3.7) \[ \lim_{n \to \infty} \frac{1}{n} E\left( \sum_{j=1}^{[n \pi]} X_{(j)} \right) = \sum_{i=1}^{r} \rho_i \int_{0}^{\pi} x F_i^{-1}(\sigma_i) \]

Proof: From the proof of Lemma 3.5,

\[ \lim_{n \to \infty} \frac{1}{n} E\left( \sum_{j=[n \pi]+1}^{n} X_{(j)} \right) \leq \sum_{i=1}^{r} \rho_i \int_{0}^{\pi} x F_i^{-1}(\sigma_i) \]

Using Theorem 3.2 and the fact that

\[ \frac{1}{n} \sum_{j=[n \pi]+1}^{n} a^k_{j,n+1} \leq \frac{1}{n} E\left( \sum_{j=[n \pi]+1}^{n} X_{(j)} \right) \]

we have

\[ \lim_{n \to \infty} \frac{1}{n} E\left( \sum_{j=[n \pi]+1}^{n} X_{(j)} \right) \geq \sum_{i=1}^{r} \rho_i \int_{0}^{\pi} x F_i^{-1}(\sigma_i) \]

This proves (3.6). Equation (3.7) then follows since

\[ \frac{1}{n} E\left( \sum_{j=1}^{n} X_{(j)} \right) = \frac{1}{n} E\left( \sum_{j=1}^{n} X_{j} \right) \to \sum_{i=1}^{r} \rho_i \int_{0}^{\infty} x F_i^{-1}(\sigma_i) \] as \( n \to \infty \).

Now we are in a position to find \( \lim_{n \to \infty} a^k_{[n \pi],n+1} \). First we consider the solutions \( \sigma_1, \ldots, \sigma_r \) of

\[ \sum_{i=1}^{r} \rho_i \sigma_i = \pi \quad \text{and} \quad F_1^{-1}(\sigma_1) = \cdots = F_r^{-1}(\sigma_r) \]

Define \( x(\pi) \) to be the common number \( F_1^{-1}(\sigma_1) = \cdots = F_r^{-1}(\sigma_r) \). Also define
\[ g(x) = \sum_{i=1}^{r} \rho_i F_i(x), \]

and assume that \( g \) is strictly increasing in a neighborhood around \( x(\pi) \). In this case, \( g^{-1} \) is defined in a neighborhood around \( \pi \), since \( g(x(\pi)) = \pi \), so that \( x(\pi) = g^{-1}(\pi) \). We use this to obtain the following theorem.

**Theorem 3.3:** Suppose each \( F_i \) is absolutely continuous with density \( f_i \). Also let \( \sigma_1, \ldots, \sigma_r \) be as in Theorem 3.2, and suppose \( g = \sum_{i=1}^{r} \rho_i F_i \) is strictly increasing in a neighborhood around \( x(\pi) = F_i^{-1}(\sigma_i) \). Then for \( 0 < \pi < 1 \),

\[ \lim_{n \to \infty} a_{[nr],n+1}^k = x(\pi). \]

**Proof:** Suppose \( 0 < \pi' < \pi < 1 \). Then from Theorem 3.2,

\[ \lim_{n \to \infty} \frac{1}{n} \sum_{j=[nr']+1}^{[nr]} a_{j,n+1}^k = \sum_{i=1}^{r} \rho_i \int_{x(\pi)}^{x(\pi)} xF_i(dx). \]

Since we have \( a_{1,n+1}^k \leq \cdots \leq a_{n,n+1}^k \), it follows that

\[ \frac{1}{n} ([nr] - [nr']) a_{[nr],n+1}^k \geq \frac{1}{n} \sum_{j=[nr']+1}^{[nr]} a_{j,n+1}^k, \]

for large enough \( n \) such that \( [nr] \geq [nr'] + 1 \).

Now let \( n \to \infty \) to obtain
\[
(\pi - \pi') \lim_{n \to \infty} a_{[\pi', \pi], n+1}^k \geq \sum_{i=1}^{r} \rho_i \int_{\pi}^{\pi'} x^{(\pi)} x F_i (dx),
\]

or, dividing by \( \pi - \pi' \) and letting \( \pi' \uparrow \pi \),

\[
\lim_{\pi' \uparrow \pi} \frac{1}{(\pi - \pi')} \sum_{i=1}^{r} \rho_i \int_{\pi}^{\pi'} \frac{x^{(\pi)} x F_i (dx)}{x^{(\pi')}} \leq \lim_{n \to \infty} a_{[\pi', \pi], n+1}^k.
\]

But we have

\[
\lim_{\pi' \uparrow \pi} \frac{1}{(\pi - \pi')} \sum_{i=1}^{r} \rho_i \int_{\pi}^{\pi'} \frac{x F_i (dx)}{x^{(\pi')}} = \lim_{\pi' \uparrow \pi} \frac{1}{(\pi - \pi')} \sum_{i=1}^{r} \rho_i \int_{0}^{\frac{g^{-1}(\pi)}{g^{-1}(\pi')}} x F_i (dx)
\]

\[
= \frac{d}{d\pi} \left( \sum_{i=1}^{r} \rho_i \int_{0}^{g^{-1}(\pi')} x F_i (dx) \right)
\]

\[
= \sum_{i=1}^{r} \rho_i \left( \frac{d}{d\pi} \right) g^{-1}(\pi) \cdot g^{-1}(\pi) f_i (g^{-1}(\pi))
\]

\[
= \sum_{i=1}^{r} \rho_i \frac{\rho_i}{\sum_{i=1}^{r} \rho_i f_i (g^{-1}(\pi'))} \cdot g^{-1}(\pi) f_i (g^{-1}(\pi))
\]

\[
= g^{-1}(\pi) = x(\pi).
\]

Similarly, letting \( 0 < \pi < \pi' < 1 \) and going through the same argument, we obtain

\[
\lim_{n \to \infty} a_{[\pi], n+1}^k \leq x(\pi).
\]
Therefore, \( \lim_{n \to \infty} a_{[n \pi],n+1}^k = x(\pi) \), and the proof is complete.

The paper by Albright and Derman [1] treats the case where \( r = 1 \). Then Theorem 3.3 reduces to

\[
\lim_{n \to \infty} a_{[n \pi],n+1}^l = F^{-1}(\pi).
\]

The results of the above section may be used in two ways. First, we may want to calculate the \( a_{j,n}^i \)'s only for \( n \leq N \), at which point it appears that the \( a_{j,N}^i \)'s are close enough to the limits of Theorem 3.3 that further calculations seem unnecessary.

A second application is when we have a large number of men, all with one of two values, \( p < \bar{p} \). Then Theorem 3.2 shows that the feasible non-optimal policy described in Lemma 3.2 is nearly optimal in the sense of expected average reward per job. That is, if \([n\pi] = \) number of \( p \)'s, then

\[
\frac{1}{n} \left( \sum_{j=1}^{[n\pi]} p a_{j,n+1}^k + \sum_{j=[n\pi]+1}^{n} \bar{p} a_{j,n+1}^k \right) \approx \frac{1}{n} E(p R_n + \bar{p} \bar{R}_n)
\]

where \( R_n \) is the sum of the values of the jobs given to the \( p \) men and \( \bar{R}_n \) is the sum of the values of the jobs given to the \( \bar{p} \) men, under the non-optimal policy. Therefore, since the non-optimal policy is much easier to administer, we may choose to use it rather than the optimal policy.
Example: An example of Theorem 3.3 is the following. Suppose \( r = 2 \),

\[
F_1(x) = \begin{cases} 
0 & x \leq 0 \\
\frac{1}{2} x & 0 \leq x \leq 2 \\
1 & x \geq 2 
\end{cases},
\]

and

\[
F_2(x) = \begin{cases} 
0 & x \leq 1 \\
\frac{1}{2} x - \frac{1}{2} & 1 \leq x \leq 3 \\
1 & x \geq 3 
\end{cases}.
\]

Suppose \( \frac{1}{2} \rho_1 \leq \pi \leq 1 - \frac{1}{2} \rho_2 \). Then \( x(\pi) = 2\pi + \rho_2 \), since this means \( 1 \leq x(\pi) \leq 2 \), and therefore,

\[
\rho_1 F_1(x(\pi)) + \rho_2 F_2(x(\pi)) = \rho_1 (\pi + \frac{\rho_2}{2}) + \rho_2 (\pi + \frac{\rho_2}{2} - \frac{1}{2}) = \pi.
\]

Thus,

\[
\lim_{n \to \infty} a_{[n\pi], n+1}^1 = \lim_{n \to \infty} a_{[n\pi], n+1}^2 = 2\pi + \rho_2.
\]

Next, suppose \( 1 - \frac{1}{2} \rho_2 \leq \pi < 1 \). Then \( x(\pi) = 1 + (2\pi - 2\rho_1)/\nu_2 \), since this means \( 2 \leq x(\pi) < 3 \), and therefore,

\[
\rho_1 F_1(x(\pi)) + \rho_2 F_2(x(\pi)) = \rho_1 + \rho_2 \left( \frac{1}{2} + \frac{(\pi - \rho_1)}{\rho_2} - \frac{1}{2} \right) = \pi.
\]

Thus,

\[
\lim_{n \to \infty} a_{[n\pi], n+1}^1 = \lim_{n \to \infty} a_{[n\pi], n+1}^2 = 1 + \frac{(2\pi - 2\rho_1)}{\rho_2}.
\]
Finally, suppose $0 < \pi \leq \frac{1}{2} \rho_1$. Then $x(\pi) = \frac{2\pi}{\rho_1}$, since this means $0 < x(\pi) \leq 1$, and therefore,

$$\rho_1 F_1(x(\pi)) + \rho_2 F_2(x(\pi)) = \rho_1 \frac{\pi}{\rho_1} + \rho_2 \cdot 0 = \pi.$$ 

Thus,

$$\lim_{n \to \infty} a_{[n\pi],n+1}^1 = \lim_{n \to \infty} a_{[n\pi],n+1}^2 = \frac{2\pi}{\rho_1}.$$ 

3.2. Job Distribution Comes from a Random Mixture.

We now turn our attention to a slightly different version of the above model. We assume that the probabilities $q_{k1}$ do not depend on $k$. That is, each job comes from distribution $F_i$ with probability $q_i$, independent of the previous jobs. This model, however, can easily be put in the framework of the previous model, with $r = 1$, by considering the mixture

$$F(x) = \sum_{i=1}^{r} q_i F_i(x).$$ 

Since the successive job values are really a random sample from $F$, we may directly use the results of the previous model, in particular, Theorems 3.1, 3.2, and 3.3, for the present model.
3.3. Job Distribution Comes from a Random Mixture; r Categories of Men.

Next we look at several generalizations of the previous two models. The main difference in these will be that not all of the men can do all of the jobs. For the first, assume there are \( r \) job distributions \( F_1, \ldots, F_r \). Corresponding to these are \( r \) types of men, with \( m_i \) of type \( i \). We label their values \( p_1^i \leq \cdots \leq p_{m_i}^i \), and we assume a job from distribution \( F_i \) can be done only by a man of type \( i \). The jobs arrive sequentially, and we assume each job comes from distribution \( F_i \) with probability \( q_i \), where \( \sum_{i=1}^{r} q_i = 1 \). The value of the job is then a non-negative random variable from distribution \( F_i \). The values of the successive jobs are independent of each other and of the process which determines the distributions of the successive jobs. We then have the following theorem.

**Theorem 3.4:** Assume there are \( n \) jobs remaining. Then there exist numbers \( a_{1,n}^i \leq \cdots \leq a_{n-1,n}^i \), along with \( a_{0,n}^i = 0 \), \( a_{n,n}^i = +\infty \), such that if the next job comes from distribution \( F_i \) with value \( x \), then it is best to assign the \( j \)th smallest of the \( n \) men of type \( i \) to this job if and only if \( a_{j-1,n}^i < x \leq a_{j,n}^i \). (We may assume there are \( n \) \( p^i \)'s by using only the \( n \) highest \( p^i \)'s if \( m_i > n \) and by adding \( n-m_i \) "fake" men, with value 0, if \( m_i < n \).) The \( a_{j,n}^i \)'s are independent of the \( p^i \)'s.

The interpretation of the \( a_{j,n}^i \)'s is that if there are \( n-1 \) jobs remaining, then \( a_{j,n}^i \) is the expected value, under an optimal policy, of the job to which the \( j \)th smallest man of type \( i \) is assigned. The \( a_{j,n}^i \)'s may be calculated from the recursion:

\[ a_{j,n}^i = \max \left\{ \frac{1}{n} \sum_{k=j}^{n} a_{k,n}^i, \frac{1}{n} \sum_{k=j}^{n} a_{k,n-1}^i \right\} \]
\[ a_{l,n+1} = q_i \left\{ \int_0^{a_{l,n}} x F_i(dx) + a_{l,n} (1 - F_i(a_{l,n})) \right\} , \]

and

\[ a_{j,n+1} = (1-q_i) a_{j-1,n} + q_i \left\{ \int_{a_{j-1,n}}^{a_{j,n}} x F_i(dx) + a_{j-1,n} F_i(a_{j-1,n}) + a_{j,n} (1 - F_i(a_{j,n})) \right\} , \]

for \( 2 \leq j \leq n. \)

**Proof:** Consider there are \( r \) decision-makers where decision-maker \( i \) has control of men \( p_1^i \leq \cdots \leq p_n^i \). He tries to maximize the expected reward from these men, and since there is no interaction between the \( p^i \)'s and the \( p^j \)'s, \( i \neq j \), he may do this independently of decision-maker \( j \)'s decisions. Therefore, we focus our attention only on decision-maker \( i \), for some \( 1 \leq i \leq r. \)

With this in mind, the proof is very similar to the proof of Theorem 3.1 and proceeds by induction on \( n \), the number of jobs remaining. For \( n = 1 \), the proof is trivial, with \( a_{0,1}^i = 0, a_{1,1}^i = +\infty \).

Now assume the theorem is true for \( n-1 \). Introduce the following notation. Let

\[ f(p_1^1, \ldots, p_n^i) = \text{the optimal expected reward from men} \]

\[ p_1^1 \leq \cdots \leq p_n^i \text{ when } n \text{ jobs remain,} \]

and
\[ f(p_1^i, \ldots, p_n^i|x,i) = \text{the optimal expected reward from men} \]
\[ p_1^i \leq \cdots \leq p_n^i \text{ when } n \text{ jobs remain,} \]
given that the first of these is from distribution \( F_i \) with value \( x \).

Then we have

\[
(3.8) \quad f(p_1^i, \ldots, p_n^i) = (1-q_i) f(p_2^i, \ldots, p_n^i) \\
+ q_i \int_0^\infty f(p_1^i, \ldots, p_n^i|x,i) \ F_i(dx)
\]

and

\[
(3.9) \quad f(p_1^i, \ldots, p_n^i|x,i) = \max_{1 \leq k \leq n} \left( p_k^i x + f(p_1^i, \ldots, p_{k-1}^i, p_{k+1}^i, \ldots, p_n^i) \right).
\]

By the inductive hypothesis, it follows that the optimal policy for the \( n-1 \) job problem is independent of the \( p \)'s. Hence we may define \( a_{j,n}^i \) as the expected value, under an optimal policy, of the job to which the \( j \)th smallest \( p_i^i \) is assigned in the \( n-1 \) job problem. Then for any \( p_1^i \leq \cdots \leq p_{n-1}^i \),

\[
(3.10) \quad f(p_1^i, \ldots, p_{n-1}^i) = \sum_{j=1}^{n-1} p_j^i a_{j,n}^i.
\]

Also, since \( a_{j,n}^i \) is independent of the \( p_i^i \)'s and other policies may be obtained by permuting the \( p_i^i \)'s, it follows that any sum of the form \( \sum_{j=1}^{n-1} p_{\sigma(j)}^i a_{j,n}^i \), where \( \sigma(1), \ldots, \sigma(n-1) \) is a permutation of 1, \( \ldots, n-1 \), can be obtained as the expected reward from the \( \tilde{p}_i^i \)'s for the \( n-1 \) job problem. Hence, by Hardy's lemma
\[(3.11) \quad a_{1,n}^i \leq \cdots \leq a_{n-1,n}^i, \]

and \(0 \leq a_{1,n}^i\), since all job values are non-negative.

Using (3.10) we may write (3.9) as

\[
f(p_1^i, \ldots, p_n^i|x,i) = \max_{1 \leq k \leq n} \left( p_k^x + \sum_{j=1}^{k-1} p_j^i a_{j,n}^i + \sum_{j=k+1}^{n} p_j^i a_{j-1,n}^i \right).
\]

Again, we use Hardy's lemma and (3.11) to obtain

\[
f(p_1^i, \ldots, p_n^i|x,i) = p_k^x + \sum_{j=1}^{k-1} p_j^i a_{j,n}^i + \sum_{j=k+1}^{n} p_j^i a_{j-1,n}^i,
\]

where \(k^*\) is such that \(a_{k^*-1,n}^i < x \leq a_{k^*,n}^i\), with \(a_{0,n}^i = 0\), \(a_{n,n}^i = +\infty\).

Hence, the first choice in an \(n\) job problem, when the first job comes from distribution \(F_i\) with value \(x\), is to use the jth smallest \(p^i\) if and only if \(a_{j-1,n}^i < x \leq a_{j,n}^i\). This completes the induction.

We may now use (3.8) to prove the recursion formulas for the \(a_{j,n}^i\)'s. If \(j = 1\), then in an \(n\) job problem, man \(p_1^i\) will be used only if the first job comes from distribution \(F_i\). Using this and the meaning of \(a_{1,n+1}^i\) gives

\[
a_{1,n+1}^i = q_i \left[ \int_0^{a_{1,n}^i} xF_i(dx) + a_{1,n}^i(1 - F_i(a_{1,n}^i)) \right].
\]
For $2 \leq j \leq n$, if the first job is not from distribution $F_i$, man $p^i_j$ is dropped, so that man $p^i_j$ becomes the $(j-1)$st smallest $p^i$ for the $n-1$ job problem. Hence

$$a^i_{j,n+1} = (1-q^i_n) a^i_{j-1,n} + q^i_n \int_{a^i_{j-1,n}} F^i(x) \, dx + a^i_{j-1,n} F^i(a^i_{j-1,n}) + a^i_{j,n} (1 - F^i(a^i_{j,n}))$$

We note parenthetically that the expression

$$\int_{a^i_{j-1,n}} x F^i(x) \, dx + a^i_{j-1,n} F^i(a^i_{j-1,n}) + a^i_{j,n} (1 - F^i(a^i_{j,n}))$$

is the expected value of the job to which the $j$th smallest $p^i$ is assigned in the $n$ job problem, given that the first of the $n$ jobs comes from distribution $F_i$. This fact follows directly from the meaning of the $a^i_{j,n}$'s.

One interpretation of the preceding model is the following.

Suppose there are $m$ men present and $n$ jobs remain. All job values come from one distribution $F$, but not all of the men can do all of the jobs. In particular, there exist numbers $t_1 < \cdots < t_{r-1}$ and $s_1 < \cdots < s_{r-1}$, with $t_0 = s_0 = 0$ and $t_r = s_r = +\infty$, such that if a job has a value $t_{i-1} < x \leq t_i$, $1 \leq i \leq r$, then the man assigned to this job, if any, must have $s_{i-1} < p \leq s_i$.

Such might be the case if we never wanted to assign a man to a job for which he was either too good or not good enough. It could also
apply to the assignment of hospital rooms to incoming patients, where hospital rooms take the place of "men" and incoming patients take the place of "jobs".

This model fits the exact framework of the model in Theorem 3.4 by defining, for \( 1 \leq i \leq r \),

\[
q_i = P(t_{i-1} < \text{value of job, } X \leq t_i) = F(t_i) - F(t_{i-1})
\]

and

\[
F_i(x) = \begin{cases} 
0 & x \leq t_{i-1} \\
1 & x > t_i
\end{cases}
\]

\[
\frac{F_i(x) - F(t_{i-1})}{F(t_i) - F(t_{i-1})}
\]

Now the ith decision-maker uses the numbers \( a^i_1, n \leq \cdots \leq a^i_{n-1}, n \) to make the optimal assignments of the men with p's in \( (s_{i-1}, s_i) \). Note that if there are \( m_i < n \) p's in \( (s_{i-1}, s_i) \), we add \( n-m_i \) "fake" men with value \( 0 \notin (s_{i-1}, s_i) \) to this category of men. This, however, we allow, since assigning a "fake" man to a job is the same as not assigning the job at all.

Limiting Results: Now we examine the limiting behavior of the \( a^i_{j,n} \)'s of Theorem 3.4 for large \( n \). As before we proceed in two parts. First we show that for \( 1-q_i < \pi < 1 \), we have

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{j=[mr]+1}^{n+1} a^i_{j,n} = q_i \int_{F_i^{-1}(\sigma_i)}^{\infty} xF_i(dx)
\]
where $\sigma_i = 1 - (1 - \eta)/q_i$. Then we show that

$$\lim_{n \to \infty} a_i^{[nF_i]}_{n+1} = F_i^{-1}(\sigma_i).$$

We first introduce the following random variables. Let $T_1, T_2, \ldots$ be i.i.d. random variables, with

$$P(T_j = i) = 1 - P(T_j = 1, \ldots, i-1, i+1, \ldots, r) = q_i, \quad 1 \leq i \leq r.$$ 

Next let $X_1^i, X_2^i, \ldots$ be i.i.d. random variables from distribution $F_i$, and assume these are independent of the $T$'s. Then $X_j^i \{T_j = i\}$ is the value of the $j$th job if the $j$th job is from distribution $F_i$. Furthermore, for any fixed constant $a$, let

$$Y_j^i(a) = 1_{X_j^i > a, T_j = i} \quad \text{and} \quad Z_j^i(a) = 1_{X_j^i > a, T_j = i}.$$ 

Finally, call the $j$th job a "success from distribution $F_i"$ if and only if $Z_j^i(a) = 1$. Then we have the following lemma.

**Lemma 3.6:** Consider independent Bernoulli trials which result in success with probability

$$\rho_i = P(Z_j^i(F_i^{-1}(\sigma_i)) = 1) = q_i(1 - \sigma_i), \quad \text{where } 0 < \sigma_i < 1,$$

and the set $\{\tau|F_i(\tau) = \sigma_i\}$ is non-empty. Let $s_i(n; \rho_i) \equiv s_i(n)$ be the number of trials required to obtain $[np_i] + 1$ successes from
distribution $F_i$, and let $t_i(n) = \min(n, s_i(n))$. Then we have

$$\lim_{n \to \infty} \frac{1}{n} E t_i(n) = 1.$$ 

**Proof:** Since $0 \leq \frac{1}{n} t_i(n) \leq 1$, it suffices, by the bounded convergence theorem, to show that

$$\frac{t_i(n)}{n} \to 1 \quad \text{in prob, as } n \to \infty.$$

Let $\varepsilon > 0$. Then

$$P(\frac{t_i(n)}{n} \geq 1 + \varepsilon) = 0.$$

and

$$P(\frac{t_i(n)}{n} \leq 1 - \varepsilon) = P(\frac{s_i(n)}{n} \leq 1 - \varepsilon)$$

$$= P(s_i(n) \leq \left\lfloor n(1-\varepsilon) \right\rfloor)$$

$$= P\left( \sum_{j=1}^{\left\lfloor n(1-\varepsilon) \right\rfloor} Z_j^{i}\left( F_i^{-1}(\sigma_i) \right) \geq \left\lceil np_i \right\rceil + 1 \right)$$

$$= P\left( \frac{\sum_{j=1}^{\left\lfloor n(1-\varepsilon) \right\rfloor} Z_j^{i}\left( F_i^{-1}(\sigma_i) \right)}{\left\lfloor n(1-\varepsilon) \right\rfloor} \geq \frac{\left\lceil np_i \right\rceil + 1}{\left\lfloor n(1-\varepsilon) \right\rfloor} \right) \to 0$$

as $n \to \infty$, by the WLLN, since

$$\lim_{n \to \infty} \frac{\left\lceil np_i \right\rceil + 1}{\left\lfloor n(1-\varepsilon) \right\rfloor} = \frac{\rho_i}{1-\varepsilon} > \rho_i.$$
and \( E Z^i_j(F^{-1}_i(\sigma_i)) = \rho_i \) by definition.

This completes the proof of the lemma.

Now in the assignment problem, suppose there are \( n \) jobs left for some \( 1 - q_i < \pi < 1 \),

\[
p^i_1 = \cdots = p^i_{[n\tau]} = 0, \quad p^i_{[n\tau]+1} = \cdots = p^i_n = 1.
\]

Using the optimal policy, the expected reward from the \( p^i \)'s is

\[
\frac{1}{n} \sum_{j=[n\tau]+1}^{n} a^i_{j,n+1}.
\]

As before we now proceed to obtain bounds on the lower and upper limits of \( \frac{1}{n} \sum_{j=[n\tau]+1}^{n} a^i_{j,n+1} \), as \( n \to \infty \).

**Lemma 3.7:** For \( 1 - q_i < \pi < 1 \), define \( \sigma_i = 1 - (1-\pi)/q_i \), and assume the set \( \{ \tau | F_i(\tau) = \sigma_i \} \) is non-empty. Then

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{j=[n\tau]+1}^{n} a^i_{j,n+1} \geq q_i \int_{F^{-1}_i(\sigma_i)}^{\infty} x F'_i(dx).
\]

**Proof:** Consider the infeasible policy which assigns a \( p^i = 1 \) to a job from distribution \( F_i \) if and only if the value of the job is \( x > F^{-1}_i(\sigma_i) \), until all the \( p^i \)'s equal to 1 are used up. Let \( R_i^n \) be the reward obtained from the \( p^i \)'s under this policy. Then since there are \( n-[n\tau] = [n(1-\pi)] + 1 \) 1's to assign,
where, using the notation of Lemma 3.6,

$$t_i(n) = \min(n, s_i(n; \rho_i)),$$

and $\rho_i = q_i(1 - \sigma_i) = 1 - n$.

Now $t_i(n)$ is a finite stopping rule for the $Y_j^i$s, so by Wald's equation,

$$E(R_n^i) = E(t_i(n)) q_i \int_{F_i^{-1}(\sigma_i)}^{\infty} xF_i(dx).$$

By Lemma 3.6, we have

$$\lim_{n \to \infty} \frac{1}{n} E(R_n^i) = q_i \int_{F_i^{-1}(\sigma_i)}^{\infty} xF_i(dx).$$

Now define a feasible policy which does the same as the above policy until there are as many $p_i$'s equal to 1 left as there are jobs remaining. At this point the new policy assigns a $p_i - 1$ to every remaining job from distribution $F_i$. Let $\tilde{R}_n^i$ be the reward from this policy and let $D_n^i = \tilde{R}_n^i - R_n^i$. Then $D_n^i > 0$ since the feasible policy assigns more 1's than the infeasible policy and since the $x$'s are all non-negative. Hence
\[ q_i \int_{F_i^{-1}(\sigma_i)}^{\infty} x F_i(dx) = \lim_{n \to \infty} \frac{1}{n} E(R_i^n) \]

\[ \leq \lim_{n \to \infty} \frac{1}{n} E(\bar{R}_1^n) \]

\[ \leq \lim_{n \to \infty} \frac{1}{n} \sum_{j=\lceil n\ell \rceil + 1}^{n} a_{j,n+1} \]

since \( E(\bar{R}_1^n) \) is non-optimal. This completes the proof.

Next we obtain an upper bound.

Lemma 3.8: Under the same assumptions as in Lemma 3.7, we have

\[ \lim_{n \to \infty} \frac{1}{n} \sum_{j=\lceil n\ell \rceil + 1}^{n} a_{j,n+1} = q_i \int_{F_i^{-1}(\sigma_i)}^{\infty} x F_i(dx) \]

Proof: Let \( \ell = \ell_1(n) \) be the number of jobs in the first \( n \) which are from distribution \( F_i \), and let \( \nu_1(i), \ldots, \nu_\ell(i) \), be the indices of these jobs. Finally, let \( X_{\nu_1(i)}^i, \ldots, X_{\nu_\ell(i)}^i \) be the order statistics associated with \( X_{\nu_1(i)}^i, \ldots, X_{\nu_\ell(i)}^i \). Then since no sequential assignment can do better than assigning the \( n-[n\ell] \) 1's to the highest order statistics, we have

\[ \frac{1}{n} \sum_{j=\lceil n\ell \rceil + 1}^{n} a_{j,n+1} \leq \frac{1}{n} E\left( \sum_{j=\ell}^{\ell} X_{(j)}^i \right) \]

where \( m = \max(0, \ell - (n-[n\ell])) + 1 \).
Now we show that

\[ \lim_{n \to \infty} \frac{1}{n} \mathbb{E} \left\{ \sum_{j=m}^{\ell} X_{(j)}^i - \sum_{j=1}^{n} Y_{j}^i(F_i^{-1}(\sigma_i)) \right\} \leq 0, \]

which, combined with (3.13), will suffice to prove (3.12) since

\[ \frac{1}{n} \mathbb{E} \left( \sum_{j=1}^{n} Y_{j}^i(F_i^{-1}(\sigma_i)) \right) = q_i \int_{F_i^{-1}(\sigma_i)}^{\infty} xF_i(dx). \]

First, we have

\[ P(\ell < n - [\lceil n \rceil]) = P\left( \sum_{j=1}^{n} 1(T_j = i) < n - [\lceil n \rceil] \right) \]

\[ \leq P\left( \frac{1}{n} \sum_{j=1}^{n} 1(T_j = i) < q_i - \epsilon \right), \]

for some \( \epsilon > 0 \), for large enough \( n \), since \( 1 - \pi < q_i \). So by the WLLN,

\[ P(\ell < n - [\lceil n \rceil]) \to 0 \quad \text{as} \quad n \to \infty. \]

Since \( Y_{j}^i(F_i^{-1}(\sigma_i)) = X_{j}^i 1(X_{j}^i > F_i^{-1}(\sigma_i)) 1(T_j = i) \), we may write

\[ \sum_{j=1}^{n} Y_{j}^i(F_i^{-1}(\sigma_i)) = \sum_{j=1}^{\ell} X_{j}^i 1(X_{j}^i > F_i^{-1}(\sigma_i)) \]

Now, by defining

\[ N_i(n) = \sum_{j=1}^{\ell} 1(X_{j}^i > F_i^{-1}(\sigma_i)), \]

53
we have

\[
\frac{1}{n} \mathbb{E}\left\{ \sum_{j=m}^{\ell} X_{\mathbb{V}}^i(j) - \sum_{j=1}^{\ell} X_{\mathbb{V}}^i(j) 1\left(X_{\mathbb{V}}^i(j) > F_i^{-1}(\sigma_i)\right) \right\} \\
= \frac{1}{n} \int_{\{\ell < n-[n_{\mathbb{V}}]\}} \sum_{j=1}^{\ell} X_{\mathbb{V}}^i(j) - \sum_{j=1}^{\ell} X_{\mathbb{V}}^i(j) 1\left(X_{\mathbb{V}}^i(j) > F_i^{-1}(\sigma_i)\right) \\
+ \int_{\{\ell \geq n-[n_{\mathbb{V}}], N_i(n) \leq n-[n_{\mathbb{V}}]\}} (n-[n_{\mathbb{V}}]-N_i(n)) \text{ terms, all } \leq F_i^{-1}(\sigma_i) \\
+ \int_{\{\ell \geq n-[n_{\mathbb{V}}], N_i(n) \geq n-[n_{\mathbb{V}}]\}} (N_i(n)-n+[n_{\mathbb{V}}]) \text{ terms, all } \leq -F_i^{-1}(\sigma_i) \\
\leq \frac{1}{n} \int_{\{\ell < n-[n_{\mathbb{V}}]\}} \sum_{j=1}^{\ell} X_{\mathbb{V}}^i(j) 1\left(X_{\mathbb{V}}^i(j) \leq F_i^{-1}(\sigma_i)\right) \\
+ F_i^{-1}(\sigma_i) \int_{\{\ell \geq n-[n_{\mathbb{V}}]\}} \left(n-[n_{\mathbb{V}}]-N_i(n)\right) \\
\leq \frac{1}{n} F_i^{-1}(\sigma_i) \left\{ (n-[n_{\mathbb{V}}])P(\ell < n-[n_{\mathbb{V}}]) + (n-[n_{\mathbb{V}}])P(\ell \geq n-[n_{\mathbb{V}}]) \right\} \\
- \int_{\{\ell \geq n-[n_{\mathbb{V}}]\}} N_i(n) \\
= \frac{1}{n} F_i^{-1}(\sigma_i) \left\{ n-[n_{\mathbb{V}}] - \sum_{k=n-[n_{\mathbb{V}}]}^{n} kP(\ell = k) (1 - \sigma_i) \right\} \\
= \frac{1}{n} F_i^{-1}(\sigma_i) \left\{ n-[n_{\mathbb{V}}] - nq_i \left(1 - \frac{1}{q_i}\right) + \frac{1}{n} F_i^{-1}(\sigma_i) \sum_{k=0}^{n-[n_{\mathbb{V}}]-1} kP(\ell = k)(1-\sigma_i) \rightarrow 0 \right\} \\
as n \rightarrow \infty.
\]
Note that the second term goes to 0 since

\[
\frac{1}{n} \sum_{k=0}^{n-\lfloor \frac{n}{\pi} \rfloor -1} k P(\ell = k) \leq \frac{(n-\lfloor \frac{n}{\pi} \rfloor -1)}{n} P(\ell < n-\lfloor \frac{n}{\pi} \rfloor) \longrightarrow 0
\]

as \( n \to \infty \).

Therefore, we have

\[
\lim_{n \to \infty} \frac{1}{n} E\left( \sum_{j=m}^{\ell} X_i^j(j) \right) \leq q_i \int_{F_i^{-1}(\pi)}^{\infty} x F_i(dx)
\]

and the proof is complete.

**Theorem 3.5:** Under the same assumptions as in Lemma 3.8, we have

\[
\lim_{n \to \infty} \sum_{j=\lfloor \frac{n}{\pi} \rfloor, n+1}^{n} a_{i,n+1,j} = q_i \int_{F_i^{-1}(\pi)}^{\infty} x F_i(dx)
\]

**Proof:** The proof follows directly from Lemmas 3.7 and 3.8.

**Theorem 3.6:** Suppose \( F_i \) is absolutely continuous with density \( f_i \), and suppose \( 1 - q_i < \pi < 1 \). Then we have

\[
(3.14) \quad \lim_{n \to \infty} a_{\lfloor \frac{n}{\pi} \rfloor, n+1} = F_i^{-1}(\pi)
\]

where \( \pi = 1 - (1-\pi)/q_i \). Furthermore, let

\[
R_{pi} = \sup\{x \mid F_i(x) = 0\}
\]

55
Then for $0 < \pi \leq 1 - q_i$, we have

$$(3.15) \quad \lim_{n \to \infty} a_{\lfloor n\pi \rfloor}, n+1 \leq \frac{R_{q_i}}{F_i}.$$ 

**Proof:** The proof of (3.14) is omitted since it follows from Theorem 3.5 in exactly the same manner as the proof of Theorem 3.3 followed from Theorem 3.2. The only thing to note is that the factor $q_i$ disappears because

$$\frac{d}{dt} q_i \int_0^{F_i^{-1}(\sigma_i)} x F_i(dx)$$

$$= q_i \frac{1}{q_i} \frac{1}{f_i(F_i^{-1}(\sigma_i))} F_i^{-1}(\sigma_i) f_i(F_i^{-1}(\sigma_i)) = F_i^{-1}(\sigma_i).$$

To prove (3.15), we note that for $\pi' < 1 - q_i < \pi$,

$$
\lim_{n \to \infty} a_{\lfloor n\pi' \rfloor}, n+1 \leq \lim_{n \to \infty} a_{\lfloor n(1-q_i) \rfloor}, n+1 \leq \lim_{n \to \infty} a_{\lfloor n\pi \rfloor}, n+1 = F_i^{-1}(\sigma_i).
$$

But $F_i^{-1}(\sigma_i)$ as $\pi \to 1 - q_i$, which proves (3.15).

Consider again the model described after Theorem 3.4. In that case

$$q_i = F(t_i) - F(t_{i-1}),$$

and

56
\[
F_i(x) = \begin{cases} 
\frac{(F(x) - F(t_{i-1}))}{(F(t_i) - F(t_{i-1}))} & t_{i-1} < x \leq t_i \\
0 & x \leq t_{i-1} \\
1 & x > t_i 
\end{cases}
\]

This implies the relation

\[F_i^{-1}(\sigma_i) = F_i^{-1}(\sigma_i(F(t_i) - F(t_{i-1})) + F(t_{i-1})), \quad 0 \leq \sigma_i \leq 1.\]

Therefore, if \(1 - q_i < \pi < 1\) and \(\sigma_i = 1 - (1-\pi)/q_i\), we have

\[
\lim_{n \to \infty} a_{[nT],n+1}^i = F_i^{-1}(\sigma_i) = F_i^{-1}(q_i - 1 + \pi + F(t_{i-1})) ;
\]

if \(0 < \pi \leq 1 - q_i\),

\[
\lim_{n \to \infty} a_{[nT],n+1}^i \leq t_{i-1}.
\]

3.4. **Job Distributions Determined by a Markov Chain; r Categories of Men.**

A slightly different version of the above model is to assume that the successive job distributions form a Markov chain. That is, if one job comes from distribution \(F_i\), the next job comes from distribution \(F_j\) with probability \(q_{ij}\). Again, only men \(p_1^i \leq \cdots \leq p_n^i\) can do jobs from distribution \(F_i\). We have the following theorem.
Theorem 3.7: Suppose there are \( n \) jobs remaining. Then there exist numbers \( a_{1,n}^i \leq \cdots \leq a_{n-1,n}^i \), along with \( a_{0,n}^i = 0 \) and \( a_{n,n}^i = +\infty \), such that if the next job comes from distribution \( F_i \) with value \( x \), it is best to assign the \( j \)th smallest \( p^i \) to this job if and only if \( a_{j-1,n}^i < x \leq a_{j,n}^i \). (Again, if there are more than \( n \) \( p^i \)'s, use only the highest \( n \); if there are \( m_i < n \) \( p^i \)'s, add \( n-m_i \) "fake" men, with value \( 0 \), to the \( p^i \)'s.) The \( a_{j,n}^i \)'s are independent of the \( p^i \)'s.

The interpretation of the \( a_{j,n}^i \)'s is as follows. In an \( n \) job problem, suppose the first job was from distribution \( F_i \), was assigned, and that \( n-1 \) \( p^i \)'s now remain (plus, of course, any other \( p^k \)'s, \( k \neq i \)). Then \( a_{j,n}^i \) is the expected value, under an optimal policy, of the job to which the \( j \)th smallest \( p^i \) is assigned during the remainder of the problem. This gives the recursion formulas:

\[
a_{j,n+1}^i = q_{ii} \left\{ \int_0^{a_{1,n}^i} xF_i(dx) + a_{1,n}^i (1 - F_i(a_{1,n}^i)) \right\}
\]

and

\[
a_{j,n+1}^i = (1 - q_{ii}) a_{j-1,n}^i
\]

\[
+ q_{ii} \left\{ \int_{a_{j,n}^i}^{a_{1,n}^i} xF_i(dx) + a_{j-1,n}^i F_i(a_{j-1,n}^i) - a_{j,n}^i (1 - F_i(a_{j,n}^i)) \right\}
\]

for \( 2 \leq j \leq n \).

Proof: The proof is omitted since it is entirely similar to the proof of Theorem 3.4.
Limiting Results: Suppose the Markov chain which determines the successive job distributions is irreducible and positive recurrent with steady state probabilities \( \rho_1, \ldots, \rho_r \). Suppose the chain starts in state \( k \) and let \( T_1, T_2, \ldots \) denote the successive states. Define \( \ell_i(n) \) by

\[
\ell_i(n) = \sum_{j=1}^{n} 1(T_j = i),
\]

the number of visits to state \( i \) by time \( n \). Then we have

\[
\frac{\ell_i(n)}{n} \rightarrow \rho_i \quad \text{in prob. as } n \rightarrow \infty,
\]

by Markov chain theory.

Using (3.16) we may prove the following two theorems in the same way as Theorems 3.5 and 3.6 were proved. Therefore, the proofs are again omitted.

**Theorem 3.8:** Suppose \( 1 - \rho_i \leq \pi < 1 \), \( \sigma_i = 1 - (1-\pi)/\rho_i \), and

\[\{\tau \mid F_i(\tau) = \sigma_i\} \neq \emptyset.\]

Then

\[
\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=[n\pi]+1}^{n} a_{i,j,n+1} = \rho_i \int_{F_i(\sigma_i)}^{\infty} x F_i(dx).
\]

**Theorem 3.9:** Suppose \( F_i \) is absolutely continuous with density \( f_i \), and \( 1 - \rho_i < \pi < 1 \). Then

\[
\lim_{n \rightarrow \infty} a_{[n\pi],n+1} = F_i^{-1}(\sigma_i),
\]
where \( \sigma_i = 1 - (1-\pi)/\rho_i \). If \( 0 < \pi \leq 1 - \rho_i \), then

\[
\lim_{n \to \infty} a_i^{[n\pi],n+1} \leq R_{F_i}.
\]

Note that in the last couple of models, it was possible to have \( a_{j,n}^i < R_{F_i} \) for some \( i,j,n \). For example, in the model discussed in Theorem 3.4, suppose \( q_i = \frac{1}{2} \) and \( F_i \) is uniform on \([1,2]\). Then

\[
a_{1,2}^i = q_i \int_1^2 x F_i (dx) = \frac{3}{4} < R_{F_i} = 1.
\]

This means that in the 2 job problem, if the first job comes from distribution \( i \), it will always be assigned to \( p_2^i \), never to \( p_1^i \).

This fact may be to our advantage computationally. Suppose for some \( i,j,n \), we have \( a_{j,n}^i \leq R_{F_i} < a_{j+1,n}^i \). Then we do not need to calculate \( a_{j,n}^i \) for \( 1 < j < j \), since their exact values will never be needed in the decision-making. This suggests that a good way to proceed is to calculate in the order \( a_{n-1,n}^i \), then \( a_{n-2,n}^i \), and so on.

Another Model with Categories of Men:

One more interesting generalization of the original assignment problem, which, however, does not have such a nice solution, is the following. Suppose again that all jobs come from one distribution \( F \). There exist numbers \( t_1 < \cdots < t_{r-1} \) and \( s_1 < \cdots < s_{r-1} \), along with \( t_0 = s_0 = 0, t_r = s_r = +\infty \), such that if a job has value \( x \in (t_{i-1}, t_i) \),
and there are any p's in \((s_{i-1}, s_i]\), then this job must either be done by one of these men or not be assigned at all. If there are no men in this category, then any man in the next highest non-empty category may be assigned to this job, if desired. That is, if there are no p's in \((s_{i-1}, s_i]\) or \((s_i, s_{i+1}],\) say, but there are some in \((s_{i+1}, s_{i+2}],\) then any of these men are eligible to do the job.

This model is probably at least as useful as the similar model already mentioned, for the following reason. We may never want to use the good men for poor jobs when there are poor men still available. But if there are no more poor men available, we would probably want to be able to use the good men on poor jobs rather than let these men sit idle.

Given that there are \(n\) jobs remaining and there are \(m_i\) p's in \((s_{i-1}, s_i]\), we may use dynamic programming to find the optimal policy. However, it can be shown (for example, with \(r = 2, n = 3, m_1 = m_2 = 1\)), that the optimal policy is not independent of the p's. Nevertheless, the optimal policy for any given p's is still a "critical numbers" policy as before. That is, there exist numbers \(a_{i1} \leq \cdots \leq a_{im_i}\), with \(a_{i0} = 0, a_{im_i+1} = +\infty\), depending upon the p's, the \(m_i\)'s, and \(n\), such that if \(m_i > 0\) and the next job has a value \(x \in (t_{i-1}, t_i]\), with \(a_{ij} < x \leq a_{i,j+1}, 0 \leq j \leq m_i\), it is best to assign the jth smallest p in \((s_{i-1}, s_i]\) to this job. Here the 0th smallest p, \(p_0 = 0\), corresponds to not assigning the job at all. These statements follow from Lemma 2.3, since we must find

\[
\max_{0 \leq j \leq m_i} (p_j x + c_j),
\]

61
where \( p_0 = 0, p_j \in (s_{i-1}, s_i], 1 \leq j \leq m_i, \) and \( c_j \) is the optimal expected reward possible after assigning \( p_j \).

Similarly, if \( m_i = 0 \), let \( i' \) be the smallest integer greater than \( i \) with \( m_{i'} > 0 \). Then there exist numbers \( a_1^{i'} \leq \cdots \leq a_{m_{i'}}^{i'} \) along with \( a_0^{i'} = 0, a_{m_{i'}+1}^{i'} = +\infty \), depending on the \( p \)'s, the \( m_i \)'s, and \( n \), such that if the next job has value \( x \in (t_{i-1}, t_i] \), with \( a_j^{i'} < x \leq a_{j+1}^{i'}, 0 \leq j \leq m_{i'} \), then it is best to assign the \( j \)th smallest \( p \) in \((s_{i'-1}, s_{i'}]\) to this job. Again this follows since we must find

\[
\max_{0 \leq j \leq m_{i'}} (p_j x + c_j^{i'})
\]

where \( p_0 = 0, p_j \in (s_{i'-1}, s_{i'}], 1 \leq j \leq m_{i'}, \) and \( c_j^{i'} \) is the optimal expected reward possible after assigning \( p_j \).

3.5. **Bayesian Model of the Assignment Problem.**

We now briefly explore a Bayesian model of the original assignment problem. Suppose that \( n \) men with known values \( p_1 \leq \cdots \leq p_n \) are on hand and \( n \) jobs are arriving sequentially. The values of these \( n \) jobs are random variables, all from the distribution \( F(x \mid \theta) \) with density \( f(x \mid \theta) \). However, the parameter \( \theta \) is a random variable with a prior density \( q(\theta) \). After each job value is observed, a new posterior distribution for \( \theta \) must be computed from Baye's formula. For instance, given that the first job has a value \( x \), we must change the density of \( \theta \) from \( q \) to \( q(x) \), where
\[ q_x(\theta) = \frac{f(x|\theta) q(\theta)}{\int f(x|\sigma) q(\sigma) d\sigma} \]

As before, if a man is assigned to an \( x \) job, a reward \( px \) is obtained. We seek a policy which maximizes the expected man reward.

This model is better than the original assignment model when we are not sure of the value of the parameter \( \theta \). Rather than guessing at a value for \( \theta \) and sticking with it, we may now learn more about the "real" \( \theta \) as each job arrives. Hence we should make better decisions. However, this model has the drawback that the optimal policy is in general difficult to calculate explicitly, as indicated by the following theorem and example.

**Theorem 3.10:** Suppose \( n \) men remain, and after using all the information we have received so far, the updated density of \( \theta \) is \( q(\theta) \). Then there exist numbers

\[ 0 = a_{0,n} \leq a_{1,n}(q_x) \leq \cdots \leq a_{n-1,n}(q_x) \leq a_{n,n} = +\infty , \]

such that if the next job has value \( x \), it is best to assign the \( i \)th smallest \( p \) if and only if \( a_{i-1,n}(q_x) \leq x \leq a_{i,n}(q_x) \), where \( q_x \) is the updated density of \( \theta \), given \( x \). These \( a_{i,n} \)'s are independent of the \( p \)'s.

The interpretation of \( a_{i,n}(q) \) is as follows. Suppose that \( n-1 \) men remain and the present density of \( \theta \) is \( q \). Then \( a_{i,n}(q) \)
is the expected value, under an optimal policy, to which the ith smallest
p is assigned. This gives the recursion

\[(3.17) \quad a_{i,n+1}(q) = \int_{\{a_{i-1,n}(q_x) < x \leq a_{i,n}(q_x)\}} xh(x) \, dx + \int_{\{x \leq a_{i-1,n}(q_x)\}} a_{i-1,n}(q_x) h(x) \, dx + \int_{\{a_{i,n}(q_x) < x\}} a_{i,n}(q_x) h(x) \, dx,\]

where

\[h(x) = \int f(x|\theta) \, q(\theta) \, d\theta\]

is the marginal distribution of the value of the first (nth from last)
job.

**Proof:** The proof proceeds in the same way as the proof of the original
assignment problem. However, we include it for completeness.

The proof is trivial for \( n = 1 \). Suppose that it holds for
n-1. We wish to show it is true for \( n \). First, define the following.

Let

\[f(p_1, \ldots, p_n; q) = \text{optimal expected reward from men}\]

\[p_1 \leq \cdots \leq p_n \text{ when there are } n \text{ jobs left, and the present density}\]

of \( \theta \) is \( q \),

and
\[ f(p_1, \ldots, p_n; q|x) \] is optimal expected reward from men

\[ p_1 \leq \cdots \leq p_n \] when the first of

the remaining \( n \) jobs has value

\( x \), and given that the density of

\( \theta \) was \( q \) before the \( x \) was observed.

Then we have

\[ f(p_1, \ldots, p_n; q|x) = \max_{1 \leq k \leq n} (p_k x + f(p_1, \ldots, p_{k-1}, p_{k+1}, \ldots, p_n; q|x)) \] however, by the inductive hypothesis, it follows that the

optimal policy for an \( n-1 \) man problem is independent of the \( n-1 \) p's.

Hence we may define \( a_{i,n}(q') \) to be the expected value, under an

optimal policy, of the job to which the \( i \)th smallest \( p \) is assigned

in the \( n-1 \) man problem, given that at the beginning of that problem,

the density of \( \theta \) is \( q' \). Then the optimal expected reward from any

\( n-1 \) men \( \bar{p}_1 \leq \cdots \leq \bar{p}_{n-1} \) is

\[ a_{i,n}(q') \]

Furthermore, since \( a_{i,n}(q') \) is independent of the \( p \)'s and other

policies may be obtained by permuting the \( p \)'s, any sum of the form

\[ \sum_{i=1}^{n} \bar{p}_o(i) a_{i,n}(q') \]
where $\sigma(1), \ldots, \sigma(n-1)$ is any permutation of $1, \ldots, n-1$, can be obtained for the expected reward from the $n-1$ man problem. By Hardy's lemma this means that

\begin{equation}
(3.20) \quad a_{1,n}(q') \leq \cdots \leq a_{n-1,n}(q')
\end{equation}

for any $q'$, since (3.19) is optimal.

Using (3.19), we may write (3.18) as

\[
f(p_1, \ldots, p_n; q|x) \]

\[
= \max_{1 \leq k \leq n} \left( p_k x + \sum_{i=1}^{k-1} p_i a_{i,n}(q_x) + \sum_{i=k+1}^{n} p_i a_{i-1,n}(q_x) \right).
\]

We may again use Hardy's lemma and (3.20) to obtain

\[
f(p_1, \ldots, p_n; q|x) \]

\[
= p^{k^*}_k x + \sum_{i=1}^{k^*-1} p_i a_{i,n}(q_x) + \sum_{i=k^*+1}^{n} p_i a_{i-1,n}(q_x),
\]

where $k^*$ is such that $a_{k^*-1,n}(q_x) < x \leq a_{k^*,n}(q_x)$, with $a_{0,n} = 0$, $a_{n,n} = +\infty$.

Hence, the first choice in an $n$ man problem, when the density of $\theta$ is $q$ and the first job has value $x$, is to use the $i$th smallest $p$ if and only if $a_{i-1,n}(q_x) < x \leq a_{i,n}(q_x)$. This completes the induction.
The recursion then follows immediately by conditioning on the value of the first (nth from last) job, which has density \( h(x) \), and by using the interpretation of \( a_{i,n+1}(q) \) given in the theorem.

**Remark:** The critical numbers, the \( a's \), now depend on the density of \( \theta \), and therefore, on the information which has been received so far. This makes the calculations quite tedious, as will be seen in the following example.

**Example:** In this example we examine the case where \( f \) is exponential with mean \( \frac{1}{\theta} \) and \( q \) is gamma. That is

\[
f(x|\theta) = \theta e^{-x\theta}, \quad x > 0,
\]

and

\[
q(\theta; y, m) = \frac{\theta^m y^{m+1}}{m! e^{-y\theta}} \quad \theta > 0.
\]

We choose \( q \) to be gamma because the gamma is the conjugate of the exponential, and hence all updated densities of \( q \) will be gamma. For instance

\[
(3.21) \quad q_x(\theta; y, m) = q(\theta; x+y, m+1),
\]

by a simple calculation. Since \( q \) is completely determined by the parameters \( y \) and \( m \), we will indicate only these parameters to show which gamma we are talking about.
We now show that for any \( n \) and sufficiently large \( m \), we have the representation

\[
(3.22) \quad a_{i,n}(y,m) = y g_{i,n}(m), \quad 0 \leq g_{1,n}(m) \leq \cdots \leq g_{n-1,n}(m) < 1.
\]

This implies

\[
a_{i,n}(q_x) = a_{i,n}(x+y, m+1) = (x+y) g_{i,n}(m+1).
\]

Hence there exist numbers \( a_{i,n}^\ast, 1 \leq i \leq n-1 \), such that

\[
a_{i-1,n}(q_x) \leq x \leq a_{i,n}(q_x) \quad \text{if and only if} \quad a_{i-1,n}^\ast \leq x \leq a_{i,n}^\ast, \quad \text{with} \quad a_{0,n}^\ast = 0, \ a_{n,n}^\ast = +\infty.
\]

This may be seen easily from the following diagram:
We assume (3.22) is true for \( n \) and proceed by induction to prove if for \( n+1 \). By (3.17) we have

\[
\begin{align*}
a_{1,n+1}(y,m) &= \int_0^\infty \int_0^\infty a_{1,n}^*(x+y, m+1) h(x) \, dx \\
&= \int_0^\infty \left\{ \int_0^\infty x \theta e^{-\theta x} \, dx + \int_0^\infty (x+y) g_{1,n}^*(m+1) \theta e^{-\theta x} \, dx \right\} \\
&\quad \cdot \frac{\theta^m}{m!} e^{-y \theta} \, d\theta.
\end{align*}
\]

This may be calculated by using the fact that \( a_{1,n}^* = (a_{1,n}^* + y) g_{1,n}^*(m+1) \), by definition of \( a_{1,n}^* \). The result is that

\[
a_{1,n+1}(y,m) = \frac{\theta^m}{m!} \left( 1 - (1 - g_{1,n}^*(m+1))^{m+1} \right),
\]

so that we may define

\[
(3.23) \quad g_{1,n+1}(m) = \frac{1}{m} \left( 1 - (1 - g_{1,n}^*(m+1))^{m+1} \right).
\]

Similarly we obtain

\[
\begin{align*}
a_{1,n+1}(y,m) &= \int_0^\infty \left\{ \int_0^\infty a_{1-1,n}^*(x+y) g_{1-1,n}^*(m+1) \theta e^{-\theta x} \, dx \\
&\quad + \int_0^\infty x \theta e^{-\theta x} \, dx + \int_0^\infty (x+y) g_{1,n}^*(m+1) \theta e^{-\theta x} \, dx \right\} \\
&\quad \cdot \frac{\theta^m}{m!} e^{-y \theta} \, d\theta \\
&= yg_{1-1,n}^*(m+1) + \frac{\theta^m}{m!} \left( (g_{1-1,n}^*(m+1) + (1 - g_{1-1,n}^*(m+1))^{m+1} \\
&\quad - (1 - g_{1,n}^*(m+1))^{m+1} \right),
\end{align*}
\]
so that

\[(3.24) \quad g_{i,n+1}(m) = g_{i-1,n}(m+1)\]

\[+ \frac{1}{m} \left( g_{i-1,n}(m+1) + (1-g_{i-1,n}(m+1))^{m+1} - (1-g_{i,n}(m+1))^{m+1} \right),\]

for \(1 < i < n\), and

\[a_{n,n+1}(y,m) = \int_0^\infty \int_0^{a_{n-1,n}^*} (x+y)g_{n-1,n}(m+1)e^{-x\theta} dx\]

\[+ \int_0^\infty x\theta e^{-x\theta} dx \cdot \frac{e^{m\frac{\theta}{m}} - e^{-\theta}}{m^\frac{\theta}{m}} e^{-y\theta} d\theta\]

\[= yg_{n-1,n}(m+1)\]

\[+ \frac{Y}{m} \left( g_{n-1,n}(m+1) + (1-g_{n-1,n}(m+1))^{m+1} \right),\]

so that

\[(3.25) \quad g_{n,n+1}(m) = g_{n-1,n}(m+1)\]

\[+ \frac{1}{m} \left( g_{n-1,n}(m+1) + (1-g_{n-1,n}(m+1))^{m+1} \right).\]

Now, we know that

\[a_{1,n+1}(y,m) \leq \cdots \leq a_{n,n+1}(y,m)\]

for any \(y,m\), by Theorem 3.10. Hence it suffices to show that for large enough \(m\), \(g_{n,n+1}(m) < 1\), and it will follow that \(g_{i,n+1}(m) < 1\), \(1 \leq i \leq n-1\).
First, we have that

\[(3.26) \quad g_{1,2}(m) = \frac{1}{m},\]

since

\[a_{1,2}(y,m) = \int_0^\infty xh(x; y, m) \, dx = \int_0^\infty \frac{1}{\theta} \cdot \frac{\theta^m y^{m+1}}{m!} \cdot e^{-y\theta} \, d\theta.\]

Thus it follows inductively from (3.25) and (3.26) that \(g_{n,n+1}(m) \leq \frac{n}{m} < 1\) when \(m > n\). This completes the proof of (3.22).

Notice that if \(\theta\) has density \(q(\theta; y, m)\), then \(E \theta = \frac{(m+1)}{y}\) and \(\text{Var} \ \theta = \frac{(m+1)}{y^2} = \frac{E \theta}{y}\). Therefore priors on \(\theta\) with the same means, but successively larger values of \(m\), have successively smaller variances. Thus a large value for \(m\), with \(\frac{m+1}{y}\) fixed, means we are assuming we have more information on the real \(\theta\), in the sense that \(\text{Var} \ \theta\) is smaller.

As an example where \(m\) is not large enough for (3.22) to hold is the following. Let \(n = 2\) and \(m = 0\). To make a decision on the first job, whose value is \(x\), we need \(a_{1,2}(x+y, m+1) = x+y\). For \(y > 0\), however, \(x < x+y\) always, so we give the first job to man \(p_1\) no matter what \(x\) is. In the sense of the above paragraph, we do not have enough information to ever risk giving the best man away first.

However, assuming we have \(n\) men on hand and \(m\) is large enough (\(m \geq n\) will do) formula (3.22) - (3.26) give us, in explicit form, exactly what we need in order to use the optimal policy of Theorem 3.10.

For completeness, we include in this section another result from Derman, Lieberman, and Ross [3]. Suppose the successive job values \( X_1, X_2, \ldots, X_n \) form a sub-martingale, that is,

\[
E(X_j | X_{j-1}, \ldots, X_1) \geq X_{j-1}, \quad \text{for all } j \geq 2.
\]

Furthermore, suppose there are \( n \) men \( p_1 \leq \cdots \leq p_n \), and each man can do each job. Then we have the following theorem, whose proof may be found in [3].

**Theorem 3.11:** If the successive job values form a sub-martingale, then the optimal policy is to use \( p_1 \), then \( p_2, \ldots, \) and finally \( p_n \). If the \( X \)'s form a super-martingale, that is

\[
E(X_j | X_{j-1}, \ldots, X_1) \leq X_{j-1}, \quad \text{for all } j \geq 2,
\]

then the optimal policy is to use \( p_n \), then \( p_{n-1}, \ldots, \) and finally \( p_1 \).

Notice that throughout this chapter we have assumed the reward from assigning a \( p \) to an \( x \) is \( r(p,x) = px \). However, the results still hold if \( r(p,x) \) is of the form

\[
r(p,x) = h_1(p) g_1(x) + h_2(p) + g_2(x),
\]
with $h_1(p)$ non-negative and monotone, and $g_1(x)$ non-negative.

This follows because we can define a new set of $X$'s, $g_1(X_1), \ldots, g_1(X_n)$, and a new set of $p$'s, $p_1' = h_1(p_1), \ldots, p_n' = h_1(p_n)$, and use the fact that the optimal policies for the models we have considered are independent of the $p$'s. The terms $g_2(p)$ and $h_2(x)$ cause no trouble because their contribution to the total expected reward is the same no matter which decisions are made.

3.7. **Choosing the Best** $p$'s.

In [3], the authors look at various models where the $p$'s can be chosen from a given set $P$, and they attempt to choose the $p$'s in an optimal manner, before the assigning takes place. For instance, suppose $P = \{p|0 \leq p \leq 1\}$, and the cost of choosing a $p$ is $c(p)$. Since the optimal $n$ man reward from the assignment problem is

$$\sum_{i=1}^{n} p_i a_i, n+1,$$

for any $p_1 \leq \cdots \leq p_n$, they attempt to choose the $p$'s from the set $P$ so as to maximize the expected reward

$$\sum_{i=1}^{n} (p_i a_i, n+1 - c(p_i)).$$

The implication is that once this optimizing has been done and the optimal $p$'s have been found, the optimal assigning is then done using this fixed set of $p$'s.

We now look at the problem of choosing the best $p$'s from a different point of view, and we show that in several special cases,
our results are identical with the results from using the above approach.

Suppose we have some constraints on the various p's we may use, say $\sum_{i=1}^{n} p_i = N$. However, rather than choosing the p's all at once, we wish to choose them one at a time after observing the successive x's. After choosing a given p, we must then change the constraint. For example, if the first p chosen is $p_1$, then the constraint for the n-1 man problem is $\sum_{i=2}^{n} p_i = N - p_1$.

We now give several models where this formulation is quite natural, and where it is, at least conceptually, superior to the method in [3]. However, in all but one of the models we consider, the results from the two different formulations are identical.

Model 1: In this model we assume that n i.i.d. random variables, $X_1, \ldots, X_n$, arrive sequentially. After observing $X_i$, we must choose $p_i$ from some constraint set so as to maximize the expected reward from there on. (Note that we have changed notation: $p_i$ is now the ith p chosen, and we do not necessarily have $p_1 \leq \cdots \leq p_n$.) The constraints which the n p's must satisfy are

$$\sum_{i=1}^{n} p_i = N, \quad p_i \geq 0, \quad 1 \leq i \leq n.$$ 

As before, if a p is assigned to an x, the reward is $px$. We now give several interpretations of this model.

In the first interpretation we assume that various proposals are being given to a central planner, say, a central figure in the
State Department. These proposals arrive sequentially, and their values, the \( X_i \)'s, are random variables which represent their values per dollar spent, as assessed by the planner. The planner has a fixed amount of money \( N \) which he must allocate to \( n \) proposals so as to maximize the total expected value received from his money.

Another interpretation of this model is that a company decides that in the next \( n \) days, it must purchase a quantity \( N \) of some item. The unit price of this item, however, is a random variable which changes from day to day. Hence the company must observe the price each day and decide how much to buy on that day. Here the objective is to minimize the expected amount spent.

The following theorem shows that because of linearity in the reward function, the optimal policy is the intuitive one: make only one of the \( p \)'s positive.

**Theorem 3.12:** Let \( \{a_{i,n}\} \) be the optimal critical numbers from [3] (or from our Theorem 3.1, with \( r = 1 \)). Suppose there are \( n \) \( X \)'s remaining, and that the constraint on the \( p \)'s is

\[
\sum_{i=1}^{n} p_i = N, \quad p_i \geq 0, \quad 1 \leq i \leq n.
\]

Then if \( X_1 = x \), it is optimal to allocate \( p_1 = N \) if \( x > a_{n-1,n} \) and to allocate \( p_1 = 0 \) if \( x \leq a_{n-1,n} \). The optimal expected reward from this policy is \( N a_{n,n+1} \).
Proof: We proceed by induction on n. For n = 1, we must allocate
p_1 = N, and we receive an expected reward N\mu = Na_{1,2}', where \mu = E X_1.

Suppose the theorem is true for n-1. By defining R_n(N) to
be the optimal expected reward from n X_1's, we have R_{n-1}(N) = Na_{n-1,1}'.
Now let R_n(N|x) be the optimal expected reward from n X's given that
X_1 = x. Hence

\[ R_n(N|x) = \max_{0 \leq p_1 \leq N} \{ p_1 x + R_{n-1}(N - p_1) \} \]

\[ = \max_{0 \leq p_1 \leq N} \{ p_1 (x - a_{n-1,1}) + Na_{n-1,1} \}. \]

But

\[ \max_{0 \leq p_1 \leq N} \{ p_1 (x - a_{n-1,1}) + Na_{n-1,1} \} = \begin{cases} \{ \begin{array}{ll} Nx & \text{if } x > a_{n-1,1} \\ Na_{n-1,1} & \text{if } x \leq a_{n-1,1} \end{array} \end{cases} \]

That is, p_1 = N is best if x > a_{n-1,1} and p_1 = 0 is best if
x \leq a_{n-1,1}. Hence

\[ R_n(N) = N \int_0^{a_{n-1,1}} F(dx) + N \int_{a_{n-1,1}}^\infty x F(dx) = Na_{n,n+1}, \]

and the proof is complete.

Thus in the first interpretation we spend all our money on
one proposal, and in the second, we buy the whole quantity N on a
single day.
Model 2: This is the same model as Model 1, except we limit the amount of any one allocation by $K$. That is, the constraints are now

$$\sum_{i=1}^{n} p_i = N, \quad 0 \leq p_i \leq K, \quad 1 \leq i \leq n.$$  

This might be appropriate in the "allocating money to proposals" model, since we may only be allowed to spend $K$ dollars on any one proposal. It might also be appropriate when buying a quantity of $N$ with random prices, because size or weight considerations might limit us from buying any more than $K$ on any one day.

First we rule out several trivial cases. If $K > N$, this is the same model as Model 1, and if $nK < N$, then no policy is feasible. Hence we assume $K \leq N \leq nK$. Under this assumption the optimal policy is given below.

Theorem 3.13: Let $\{a_{i,n}\}$ be as before. Suppose there are $n$ $X$'s left and the constraint on the $p$'s is

$$\sum_{i=1}^{n} p_i = N, \quad 0 \leq p_i \leq K, \quad 1 \leq i \leq n,$$

with $K \leq N \leq nK$. Then if $d = \lfloor \frac{N}{K} \rfloor$, the greatest integer in $\frac{N}{K}$, and $X_i = x$, it is optimal to allocate $p_1 = 0$ if $x \leq a_{n-d-1,n}$, to allocate $p_1 = N - dK$ if $a_{n-d-1,n} < x \leq a_{n-d,n}$, and to allocate $p_1 = K$ if $a_{n-d,n} < x$. Furthermore, the optimal expected reward from the $n$ $X$'s is

$$(3.27) \quad (N - dK) a_{n-d,n+1} + K \sum_{i=0}^{d-1} a_{n-i,n+1}.$$
Proof: We induct on \( n \). For \( n = 1 \), we must have \( N = K \) and \( d = 1 \), so that the theorem is trivial. Assume the theorem is true for \( n-1 \), and define \( R_n(N, d) \) and \( R_n(N, d|x) \) to be the optimal expected reward from \( n \) \( X \)'s, and the optimal expected reward from \( n \) \( X \)'s, given \( X_1 = x \), respectively.

Now suppose there are \( n \) \( X \)'s remaining and \( K \leq N \leq nK \). If \( N = nK \), the theorem is trivially true, so assume \( K \leq N < nK \). Thus \( 1 \leq d \leq n-1 \), where \( d = \lfloor \frac{N}{K} \rfloor \). We treat the cases \( d = n-1, 1 < d < n-1, \) and \( d = 1 \) separately. In each case let \( N' = N - p_1 \) and \( d' = \lfloor \frac{N'}{K} \rfloor \) for any initial allocation \( p_1 \).

(1) \( d = n-1 \). In this case we must have \( N - (n-1)K \leq p_1 \leq K \), for if \( 0 \leq p_1 < N - (n-1)K \), the condition \( \sum p_i = N \) cannot be met. Thus we have

\[
R_n(N, n-1|x) = \max_{N-(n-1)K \leq p_1 \leq K} \left\{ p_1 x + (N-p_1-(n-2)K)a_{1,n} + K \sum_{i=0}^{n-3} a_{n-1-i,n} \right\}
\]

\[
= \begin{cases} 
Kx + (N-(n-1)K)a_{1,n} + K \sum_{i=0}^{n-3} a_{n-1-i,n} & \text{if } x > a_{1,n} \\
(N-(n-1)K)x + K \sum_{i=0}^{n-2} a_{n-1-i,n} & \text{if } x \leq a_{1,n}
\end{cases}
\]

Integrating over \( x \) and using the recursion formulas for the \( a_{1,n} \)'s, we obtain

\[
R_n(N, n-1) = (N - (n-1)K)a_{1,n+1} + K \sum_{i=0}^{n-2} a_{n-i,n+1}
\]

which was to be shown.
(2) \(1 < d < n-1\). Now we may allocate \(0 \leq p_1 \leq N-dK\) or \(N-dK < p_1 \leq K\). In the first case \(d' = d\), and in the second, \(d' = d-1\). Hence

\[
R_n(N, d|\{x\}) = \max\{ \max_{0 \leq p_1 \leq N-dK} (p_1 x + R_{n-1}(N-p_1, d)), \max_{N-dK < p_1 \leq K} (p_1 x + R_{n-1}(N-p_1, d-1)) \}
\]

By using (3.27), the inductive hypothesis, and comparing, we have

\[
R_n(N,d|\{x\}) = \begin{cases} 
Kx + (N-dK)a_{n-d,n} + K \sum_{i=0}^{d-2} a_{n-1-i,n} & \text{if } x > a_{n-d,0} \\
(N-dK)x + K \sum_{i=0}^{d-1} a_{n-1-i,n} & \text{if } a_{n-d,0} < x \leq a_{n-d,n} \\
(N-dK)a_{n-d-1,n} + K \sum_{i=0}^{d-1} a_{n-1-i,n} & \text{if } x \leq a_{n-d-1,n}
\end{cases}
\]

We again integrate over \(x\) and using the recursion formulas for the \(a_{i,n}\)'s to obtain

\[
R_n(N,d) = (N-dK)a_{n-d,n+1} + K \sum_{i=0}^{d-1} a_{n-i,n+1},
\]

which agrees with (3.27).

(3) \(d = 1\). In this case we may allocate \(0 \leq p_1 \leq N-K\) and \(N-K < p_1 \leq K\). In the first case \(d' = 1\) and in the second case \(d' = 0\), i.e., \(N' < K\) and we must use the result of Theorem 3.12. Thus we have
\[ R_n(N, 1|x) = \max \left( \max_{0 \leq p_1 \leq N-K} (p_1x + R_{n-1}(N-p_1, 1)) \right) \]

\[ \max_{N-K \leq p_1 \leq K} \left( p_1x + R_{n-1}(N-p_1, 0) \right) \]

where

\[ \max_{N-K \leq p_1 \leq K} \left( p_1x + R_{n-1}(N-p_1, 0) \right) = \begin{cases} 
Kx + (N-K)a_{n-1,n} & \text{if } x > a_{n-1,n} \\
(N-K)x + Ka_{n-1,n} & \text{if } x \leq a_{n-1,n}
\end{cases} \]

by Theorem 3.12. Hence we again use (3.27), the inductive hypothesis, and compare to obtain

\[ R_n(N, 1|x) = \begin{cases} 
Kx + (N-K)a_{n-1,n} & \text{if } x > a_{n-1,n} \\
(N-K)x + Ka_{n-1,n} & \text{if } a_{n-2,n} < x \leq a_{n-1,n} \\
(N-K)a_{n-2,n} + Ka_{n-1,n} & \text{if } x \leq a_{n-2,n}
\end{cases} \]

Integrating over \( x \) gives

\[ R_n(N) = (N-K)a_{n-1,n+1} + Ka_{n,n+1} \]

as desired.

Hence, all three cases have been proved, and the theorem follows.

Model 3: In this model, we assume that any \( p \) which is positive must be greater than \( K \), and that at least \( M \) \( p \)'s must be positive. That is,
\[ \sum_{i=1}^{n} p_i = N, \quad p_i \geq 0, \quad 1 \leq i \leq n, \]

\[ p_i > 0 \text{ implies } p_i \geq K, \text{ and } p_i > 0 \text{ for at least } M \text{ indices } i. \]

The motivation for this model comes from the "allocating money to proposals" interpretation. If the money is actually being allocated by a government agency, it may be necessary, for political reasons, to fund money for at least \( M > 1 \) proposals. Also, if the planner is going to spend any money at all on a given proposal, he must spend at least \( K \) dollars on it to "get it off the ground".

Again, we dispose of the trivial case when \( MK > N \), for there are no feasible policies if this is the case. If \( MK \leq N \), the optimal theorem is given below.

**Theorem 3.14:** Let \( \{a_{i,n}\} \) be as in Theorems 3.12 and 3.13. If there are \( n \) \( X \)'s remaining, \( MK \leq N \), and \( X_1 = x \), then the optimal policy is to allocate \( p_1 = 0 \) if \( x \leq a_{n-M,n} \), to allocate \( p_1 = K \) if \( a_{n-M,n} < x \leq a_{n-1,n} \), and to allocate \( p_1 = N-(M-1)K \) if \( x > a_{n-1,n} \). Furthermore, the optimal expected reward from \( n \) \( X \)'s is

\[
(3.28) \quad K \sum_{i=1}^{M-1} a_{n-i,n+1} + (N-(M-1)K) a_{n,n+1}.
\]

**Proof:** The proof is trivial for \( n = 1 \), so assume it is true for \( n-1 \). Define \( R_n(N,M) \) and \( R_n(N,M|x) \) to be the optimal expected reward from \( n \) \( X \)'s, and the optimal expected reward from \( n \) \( X \)'s, given \( X_1 = x \), respectively.

81
Now suppose there are \( n \) \( X \)'s remaining. Since the case \( M = n \) is trivial, we assume \( 1 \leq M < n \) and \( MK \leq N \). Since the first allocation may either be \( p_1 = 0 \) or \( K \leq p_1 \leq N - (M-1)K \), we have

\[
R_n(N, M | x) = \max(R_{n-1}(N,M), \max_{K \leq p_1 \leq N - (M-1)K} (p_1 x + R_{n-1}(N-p_1,M-1)))
\]

By using (3.28), the inductive hypothesis, and comparing, we obtain

\[
R_n(N, M | x) = \begin{cases} 
(N-(M-1)K)x + K \sum_{i=0}^{M-2} a_{n-1-i,n} & \text{if } x > a_{n-1,n} \\
Kx + K \sum_{i=1}^{M-2} a_{n-1-i,n} + (N-(M-1)K) a_{n-1,n} & \text{if } a_{n-M,n} < x \leq a_{n-1,n} \\
K \sum_{i=1}^{M-1} a_{n-1-i,n} + (N-(M-1)K) a_{n-1,n} & \text{if } x \leq a_{n-M,n}
\end{cases}
\]

Integrating over \( x \) and using the properties of the \( a_{i,n} \)'s, we obtain

\[
R_n(N,M) = K \sum_{i=1}^{M-1} a_{n-i,n+1} + (N-(M-1)K) a_{n,n+1}
\]

which agrees with (3.28). Thus the proof is complete.

Notice that in all three of these models we would have obtained the same results if we had proceeded as in [3], that is, if we have originally maximized the expression

\[
\sum_{i=1}^{n} p_i a_{i,n+1}
\]
over all possible allocations with $p_1 \leq \cdots \leq p_n$, and had then assigned 
these fixed set of $p$'s according to the optimal assignment. (Note 
that we are now reverting back to the previous notation for the $p$'s, 
so that $p_i$ is again the $i$th smallest $p$, not the $i$th $p$ allocated. 
Please excuse any ambiguity.)

To show, however, that there are cases where we really can do 
better by choosing the $p$'s sequentially rather than all at once, consider 
Model 1, except with the constraint there replaced by

$$\sum_{i=1}^{n} p_i^2 = N, \quad p_i \geq 0.$$ 

Let $n = 2$, $N = 1$. Using the method in [3], we will allocate the fixed 
amounts 0 and 1, and the expected reward is $a_{2,3}$. If we use our 
method and choose the $p$'s sequentially, we will observe the first $x$ 
and find

$$R_2(1|x) = \max_{0 \leq p_1 \leq 1} \left\{ p_1 x + \sqrt{1-p_1^2} \mu \right\}.$$

Thus the optimal $p_1$ is never 0 or 1. Hence we always follow a 
different, and certainly better, policy than if we use the allocations 
0 and 1.

We remark finally that Theorems 3.12 - 3.14 of Models 1-3 have 
obvious analogs if the $X$'s come from $r$ different distributions and the 
successive distributions are governed by a Markov chain (as in the first 
model of this chapter). For example, the analog of Theorem 3.12 is
Theorem 3.12': Let \( \{a_{j,n}^i\} \) be the optimal critical numbers from Theorem 3.1. Suppose there are \( n \) \( X \)'s remaining, and that the constraint on the \( p \)'s is

\[
\sum_{i=1}^{n} p_i = N, \quad p_i \geq 0, \quad 1 \leq i \leq n .
\]

Then if \( X_1 = x \) is from distribution \( F_i \), it is optimal to allocate \( p_1 = N \) if \( x > a_{n-1,n}^i \) and to allocate \( p_1 = 0 \) if \( x \leq a_{n-1,n}^i \). The optimal expected reward from this policy, assuming the previous \( X \) (the \((n+1)st from last) was from distribution \( F_k \), is \( N a_{n,n+1}^k \).

The proof of this and the analogs of Theorems 3.13 and 3.14 are straightforward, using the properties of the \( a_{j,n}^i \)'s of Theorem 3.1.

One possibility under this more general assumption on the \( X \)'s is that distribution \( F_i \) is degenerate at \( x_i \). Then the actual values of the successive \( X \)'s form a Markov chain. This might be an appropriate model when a company must buy a quantity \( N \) of some item, say stock, and the price varies from day to day.
CHAPTER 4

ASSIGNMENT PROBLEMS WITH A VARIABLE NUMBER
OF JOBS ARRIVING

In this chapter we take the original assignment problem
(with \( r = 1 \), i.e., only one job value distribution) and generalize
in the following way. Since jobs are arriving sequentially, it may be
unreasonable to assume that we know ahead of time that exactly \( n \) jobs
will arrive. Instead, we will assume there is some randomness associated
with the number of jobs and, possibly, with the times of arrivals.

We will deal with two overall models. In the first we con-
sider only policies which assign each incoming job if there is at least
one man left to assign it to. In the second, we may bypass incoming
jobs, and receive value \( 0 \), but in this case we impose the restriction
that bypassed jobs may not be assigned later on.

In each of these two models we treat two cases. The first is
where there is no time variable; we care only about how many jobs will
ever arrive and not about when they will arrive. In the second case,
time is a very definite factor, and it enters either in the form of a
limited time horizon or a discount factor. These ideas will be made
precise as we describe the particular models.
4.1. **All Jobs Must be Assigned; no Time Factor.**

We treat first the case where all jobs must be assigned if there is anyone to assign them to, and where time is no explicit factor. In particular, we suppose at the beginning of the decision problem that exactly $k$ jobs will arrive with probability $q_k$, where

$$q_k \geq 0, \quad k \geq 0 \quad \text{and} \quad \sum_{k=0}^{\infty} q_k = 1.$$ 

We define

$$q_k|_{\ell} = P(\text{at least } k \text{ jobs arrive} | \ell \text{ have already arrived}),$$

for $k > \ell$. That is, $q_k|_{\ell}$ is the probability that at least $k - \ell$ more jobs will arrive, given that $\ell$ have already arrived, so that

$$q_k|_{\ell} = \frac{\sum_{i=k}^{\infty} q_i}{\sum_{i=\ell}^{\infty} q_i}.$$ 

This problem can be solved very much like the deterministic case where $q_n = 1$, as shown by the following theorem.

**Theorem 4.1:** Suppose we start with $n$ men and we assume

$$q_k = P(\text{exactly } k \text{ jobs will arrive}), \quad k \geq 0.$$ 

If $n-j$ jobs then arrive, $0 \leq n-j \leq n-1$, and $j$ men are left with values $p_1 \leq \cdots \leq p_j$, then there exist numbers $a_{1,j}^{(n)} \leq \cdots \leq a_{j-1,j}^{(n)}$,
with \( a_{0,j}^{(n)} = -\infty \), \( a_{j,j}^{(n)} = +\infty \), such that if another job arrives with value

\( a_{i-1,j}^{(n)} < x \leq a_{i,j}^{(n)} \)

then it is best to assign man \( p_i \) to this job. These \( a_{i,j}^{(n)} \)'s are independent of the \( p_i \)'s.

Furthermore, suppose \( n-j+1 \) jobs have arrived, \( 0 \leq n-j+1 \leq n-1 \), have been paired with men, and \( j-1 \) men remain, with values

\( P_1 \leq \cdots \leq P_{j-1} \). Then \( a_{i,j}^{(n)} \) is the expected value of the job to which man \( p_i \) is assigned during the rest of the problem, under an optimal policy.

**Proof:** We assume we started with \( n \) men and at that time we assumed that

\[ q_k = P(\text{exactly } k \text{ jobs will arrive}), \quad k \geq 0. \]

Now we proceed by induction on \( j \), the number of men left after \( n-j \) jobs have already arrived, \( 0 \leq n-j \leq n-1 \). For \( j = 1 \), there is one man left, since \( n-1 \) jobs have already arrived. If another job arrives, with probability \( q_{n|n-1} \), we assign it to this last man, so the theorem is trivial in this case.

Suppose the theorem is true for \( j-1 \), where \( 1 \leq j-1 \leq n-1 \).

Before proceeding, we introduce the following notation. Let

\[ f_{j}(P_1, \ldots, P_j) = \text{optimal expected reward from men} \]

\[ P_1 \leq \cdots \leq P_j \text{ after } n-j \text{ of the original } n \text{ men have been assigned.} \]
Also let

\[ f^{(n)}(p_1, \ldots, p_j | x) = \text{optimal expected reward from men} \]

\[ p_1 \leq \cdots \leq p_j \text{ after n-j of the} \]

original men have been assigned, and

given that the (n-j+1)st job has

just arrived with value x.

From here on, the proof is entirely similar to the proof of

Theorem 3.1. By the inductive hypothesis, it follows that the optimal

policy for the remaining j-1 men, after the first of j has been

assigned, is independent of the j-1 p's. Hence we may define \[ a^{(n)}_{i,j}, \]

\[ 1 \leq i \leq j-1, \]

as the expected value, under an optimal policy, of the job

to which the ith smallest man is assigned in the rest of the problem.

Then the total expected reward from these men, whose values we label

\[ \tilde{p}_1 \leq \cdots \leq \tilde{p}_{j-1}, \]

is

\[ f^{(n)}(\tilde{p}_1, \ldots, \tilde{p}_{j-1}) = \sum_{i=1}^{j-1} \tilde{p}_i a^{(n)}_{i,j}, \tag{4.1} \]

true for any \[ \tilde{p}_1 \leq \cdots \leq \tilde{p}_{j-1}. \]

Furthermore, since \[ a^{(n)}_{i,j} \]

is independent of the p's and other policies may be obtained by permuting the p's, any

sum of the form

\[ \sum_{i=1}^{j-1} \tilde{p}_0(i) a^{(n)}_{i,j}, \]

88
where \( a(1), \ldots, a(j-1) \) is a permutation of \( 1, \ldots, j-1 \), can be obtained for the expected reward from the last \( j-1 \) men. By Hardy’s lemma this means that

\[
(4.2) \quad a_{1,j}^{(n)} \leq \cdots \leq a_{j-1,j}^{(n)},
\]

since (4.1) is optimal.

Using (4.1) we may write

\[
\begin{align*}
    f^{(n)}(p_1, \ldots, p_j | x) &= \max_{1 \leq k \leq j} \left( p_k x + f^{(n)}(p_1, \ldots, p_{k-1}, p_{k+1}, \ldots, p_j) \right) \\
    &= \max_{1 \leq k \leq j} \left( p_k x + \sum_{i=1}^{k-1} p_i a_{i,j}^{(n)} + \sum_{i=k+1}^{j} p_i a_{i-1,j}^{(n)} \right)
\end{align*}
\]

We again use Hardy’s lemma and (4.2), to obtain

\[
\begin{align*}
    f^{(n)}(p_1, \ldots, p_j | x) &= p_{k^*} x + \sum_{i=1}^{k^*-1} p_i a_{i,j}^{(n)} + \sum_{i=k^*+1}^{j} p_i a_{i-1,j}^{(n)}
\end{align*}
\]

where \( k^* \) is such that

\[
a_{k^*-1,j}^{(n)} < x \leq a_{k^*,j}^{(n)}
\]

with \( a_{0,j}^{(n)} = -\infty, \ a_{j,j}^{(n)} = +\infty \).

Hence the first choice among the remaining \( j \) men is the \( i \)th smallest if and only if another job comes with value \( a_{i-1,j}^{(n)} < x \leq a_{i,j}^{(n)} \). This completes the proof.
Corollary 4.1: We may calculate the \( a_{i,j}^{(n)} \)'s by the recursion

\[
a_{i,j+1}^{(n)} = q_{n-j+1|n-j} \left\{ \int_{a_{i-1,j}^{(n)}} xF(dx) + a_{i-1,j}^{(n)} F(a_{i-1,j}^{(n)}) + a_{i,j}^{(n)} (1 - F(a_{i,j}^{(n)})) \right\},
\]

for \( 1 \leq j \leq n, 1 \leq i \leq j \), where \( a_{0,j}^{(n)} = -\infty, a_{j,j}^{(n)} = +\infty \).

Proof: This follows directly from the interpretation given to \( a_{i,j+1}^{(n)} \) in Theorem 4.1 and the fact that \( q_{n-j+1|n-j} \) is the probability that at least one more job will arrive after \( n-j \) jobs have already arrived.

Corollary 4.2: Suppose \( q_k \) is of the form

\[
q_k = q^k (1-q), \quad 0 < q < 1.
\]

Then no superscripts are needed on \( a_{i,j}^{(n)} \), and we have the recursion

\[
a_{i,j+1} = q\left\{ \int_{a_{i-1,j}^{(n)}} xF(dx) + a_{i-1,j}^{(n)} F(a_{i-1,j}^{(n)}) + a_{i,j}^{(n)} (1 - F(a_{i,j}^{(n)})) \right\}.
\]

Proof: The reason for the superscript \( n \) in Theorem 4.1 is to show that the \( \{q_k\} \) distribution is valid only at beginning of the decision problem, that is, only when \( n \) men remain. This had to be modified to \( \{q_{n-j+k|n-j}\} \) after \( n-j \) jobs arrived. But if \( q_k = q^k (1-q) \), the process is memoryless in the sense that
\[ q_{k+\ell | k} = q_k, \quad k, \ell \geq 0. \]

Thus the distribution of the number of remaining jobs is the same no matter how many jobs have already arrived. Thus the superscript \( n \) may be removed; we need know only how many men remain.

The recursion follows immediately from Corollary 4.1 and the fact that

\[
q_{k+1 | k} = \sum_{i=k+1}^{\infty} q^i (1-q) / \sum_{i=k}^{\infty} q^i (1-q) = q.
\]

4.2. **All Jobs Must be Assigned; Time is a Factor.**

We now bring time into the picture, but we continue to assume that all jobs must be assigned. Suppose \( n \) men are available at the beginning of a time period, say a month, and assigning of these men must be done during this time period. Jobs are known to arrive at random times, with times between jobs constituting a renewal process. Suppose the times between jobs have distribution \( G \), and that this renewal process is independent of the values of the jobs, which are still i.i.d. random variables with distribution \( F \). Then the optimal policy is shown in the following theorem.

**Theorem 4.2:** In the above problem, suppose there are \( n \) men left, \( \tau \) time left, and a job of value \( x \) arrives. Then there exists numbers \( a_{1,n}(\tau) \leq \cdots \leq a_{n-1,n}(\tau) \), with \( a_{0,n}(\tau) = -\infty \), \( a_{n,n}(\tau) = +\infty \), such that if
\[
\int_0^\tau a_{i-1,n}(\tau-y) G(dy) < x \leq \int_0^\tau a_{i,n}(\tau-y) G(dy)
\]

it is best to assign the \(i\)th smallest \(p\) to this job. These \(a_{i,n}(\tau)\)'s are independent of the \(p\)'s, the values of the men.

The interpretation of \(a_{i,n}(\tau)\) is as follows. If there are \(n-1\) men left, \(\tau\) time left, and a job arrives, then \(a_{i,n}(\tau)\) is the expected value, under the optimal policy, of the job to which the \(i\)th smallest man is assigned during the rest of the problem. (This is before the value of the present job is observed.)

Letting

\[
\bar{a}_{i,n}(\tau) = \int_0^\tau a_{i,n}(\tau-y) G(dy)
\]

we have the recursion

\[
a_{i,n+1}(\tau) = \int_{\bar{a}_{i-1,n}(\tau)} x F(dx) + \bar{a}_{i-1,n}(\tau) \left( F(\bar{a}_{i-1,n}(\tau)) + a_{i,n}(\tau)(1 - F(\bar{a}_{i,n}(\tau))) \right)
\]

Proof: The proof proceeds as in Theorem 4.1. We let

\[
f(p_1, \ldots, p_n; \tau) = \text{optimal expected reward from men}
\]

\(p_1 \leq \cdots \leq p_n\) with \(\tau\) time left, given we're starting just after a job arrived and was assigned.
Also let

\[ f(p_1, \ldots, p_n; \tau | x) = \text{optimal expected reward from men} \]

\[ p_1 \leq \cdots \leq p_n, \text{ given that a job of} \]

\[ \text{value } x \text{ just arrived and } \tau \]

\[ \text{time is left}. \]

Then we have

\[ f(p_1, \ldots, p_n; \tau | x) = \max_{1 \leq k \leq n} \left( p_k x + f(p_{k-1}, \ldots, p_{k+1}, \ldots, p_n; \tau) \right). \]

Again we proceed by induction on \( n \), the number of men left.

For \( n = 1 \), the theorem is trivially true. Suppose it is true for \( n-1 \).

Next, suppose there are \( n \) men left with values \( p_1 \leq \cdots \leq p_n \), \( \tau \) time left, and a job arrives with value \( x \). By the inductive hypothesis, the optimal policy is independent of the remaining \( n-1 \) \( p \)'s, after the first man has been assigned. Therefore define \( a_{i,n}(\tau-y), \) \( 1 \leq i \leq n-1, \) \( 0 < y \leq \tau \), to be the expected value, under an optimal policy, of the job to which the \( i \)th smallest of the \( n-1 \) remaining men is assigned, given that the next job arrives \( y \) units of time later than the present job. This means that the expected value, under an optimal policy, of the job to which the \( i \)th smallest of the \( n-1 \) remaining men is assigned, right after the present assignment is made, is

\[ \ddot{a}_{i,n}(\tau) = \int_0^\tau a_{i,n}(\tau-y) G(dy), \quad 1 \leq i \leq n-1, \]

93
Therefore, we have

\[ f(p_1, \ldots, p_n; \tau | x) = \max_{1 \leq k \leq n} \left( p_k x + \sum_{i=1}^{k-1} p_i \tilde{a}_{i,n}(\tau) + \sum_{i=k+1}^{n} p_i \tilde{a}_{i-1,n}(\tau) \right) \]

From here on, we may proceed in a straightforward manner as in the proof of Theorem 4.1.

Note that the calculations needed to find \( \tilde{a}_{i,n}(\tau) \) are quite tedious, since whole functions must be found. For example, suppose \( F \) is uniform on \([0,1]\), \( G(x) = 1 - e^{-x} \), and we want \( \tilde{a}_{1,2}(\tau), \tilde{a}_{2,3}(\tau) \), the numbers needed when a job arrives with \( \tau \) time left and 3 men remaining. Then the following calculations are necessary:

\[ a_{1,2}(\tau) = \int_0^1 x \, dx = \frac{1}{2} ; \quad \tilde{a}_{1,2}(\tau) = \int_0^\tau \frac{1}{2} e^{-y} \, dy = \frac{1}{2} (1 - e^{-\tau}) ; \]

\[ a_{1,3}(\tau) = \int_0^{\frac{1}{2} (1 - e^{-\tau})} x \, dx + \frac{1}{2} (1 - e^{-\tau}) \left( \frac{1}{2} (1 + e^{-\tau}) \right) \]

\[ = \frac{3}{8} - \frac{1}{4} e^{-\tau} - \frac{1}{8} e^{-2\tau} ; \]

\[ a_{2,3}(\tau) = \int_{\frac{1}{2} (1 - e^{-\tau})}^1 x \, dx + \frac{1}{2} (1 - e^{-\tau}) \left( \frac{1}{2} (1 - e^{-\tau}) \right) \]

\[ = \frac{5}{8} - \frac{1}{4} e^{-\tau} + \frac{1}{8} e^{-2\tau} ; \]

\[ \tilde{a}_{1,3}(\tau) = \int_0^\tau a_{1,3}(\tau - y) \, e^{-y} \, dy \]

\[ = \frac{3}{8} - \frac{1}{2} e^{-\tau} - \frac{1}{4} \tau e^{-\tau} + \frac{1}{8} e^{-2\tau} ; \]

\[ \tilde{a}_{2,3}(\tau) = \int_0^\tau a_{2,3}(\tau - y) \, dy \]

\[ = \frac{5}{8} - \frac{1}{2} e^{-\tau} - \frac{1}{4} \tau e^{-\tau} - \frac{1}{8} e^{-2\tau} . \]
Note that we had to find $a_{1,3}(\tau-y)$ and $a_{2,3}(\tau-y)$ for all $0 \leq \tau-y \leq \tau$ before we could find $\tilde{a}_{1,3}(\tau)$ and $\tilde{a}_{2,3}(\tau)$.

4.3. Jobs May be Bypassed; Time is a Factor.

The limitation of the above models is that we require that every job be assigned to a man if a man is available. We might also like to be able to bypass a job if it does not look good. The following models allow this possibility. However, dynamic programming in the form of backward induction no longer works, and we must resort to other methods.

The first model is a generalization of a model introduced in a paper by Elfving [4]. In it there are $n$ men, all with value $p = 1$. Jobs arrive at random times according to a variable time Poisson process, that is, the arrival of a job during the time interval $(t, t+h)$ is independent of the previous arrivals, and the probability of its occurrence is $\lambda(t)h + o(h)$. We assume $\lambda(t)$ is continuous. The values of the jobs are assumed to be non-negative i.i.d. random variables from a continuous distribution $F$ with mean $0 < \mu < \infty$, and they are independent of the arrival time process.

When a job arrives, we must decide whether to assign it to one of the $n$ men or not to assign it at all. If a job arrives at time $t$ with value $x$, a reward $r(t)x$ is obtained if the job is assigned, and $0$ is obtained if it is not assigned. Here $r(t)$ is a piece-wise continuous, non-negative, non-increasing discount function with $r(0) = 1$. 
The policies we consider are all of the following form. If there are \( n \) men left and a job arrives at time \( t \) with value \( x \), we assign it if and only if \( x > y_n(t) \), where we assume \( y_n(t) \) is non-negative for \( t \geq 0 \). We refer to \( y_n(t) \) as the critical curve at time \( t \) when there are \( n \) men left to assign. Let \( A = \sup\{t \mid r(t) > 0\} \). Then we need to determine a policy \( y_n(t) \) only for \( t < A \); for \( t \geq A \) it does not make any difference what \( y_n(t) \) is. Therefore, in what follows, it will simplify matters to set \( y_n(t) = 0 \) for \( t \geq A \).

Before proceeding, we suggest several interpretations of the above model, besides the men and jobs interpretation. Suppose we have \( n \) identical houses to sell, and offers of random amounts arrive at random times. (Ehrling uses this interpretation, with \( n = 1 \).) Then we must decide which offers to accept. Another interpretation is that we have various (identical) positions to fill, say teaching jobs, and applicants of random quality arrive at random times. Here a reasonable \( r \) function might be \( r(t) = 1, t \leq A \), and \( r(t) = 0, t > A \), which indicates that we only have a period of time \( A \) in which to fill the positions.

We impose two restrictions on the above models which make them quite different mathematically from similar models. First we assume the men never finish a job once they are assigned. Thus, an assigned man may never be used again. Secondly, a job which is bypassed may not be used later on. In the house interpretation above, this means that if a particular offer is rejected, it may not be accepted later on. Similarly, in the model where we are trying to fill positions, if an applicant is originally rejected, he may not be hired at a later date.
We now proceed to derive a differential equation for $y_n(t)$.

First, let

$$E_n(t; y_n, \ldots, y_1) = \text{expected reward from } n \text{ men, starting at}$$

a time $t$ and using the critical curve $y_n$

until the first man is assigned, then $y_{n-1}$

until the second man is assigned, and so on.

Also, let $E_n(t) = E_n(t; y_n, \ldots, y_1)$ if $y_n, \ldots, y_1$ are optimal for the $n$ man problem. Then we have the following optimality criterion.

**Theorem 4.3:** $y_n$ is optimal if and only if it satisfies

$$E_n(t) = r(t) y_n(t) + E_{n-1}(t), \quad n \geq 1, \ t \geq 0.$$  \hspace{1cm} (4.3)

**Proof:** Suppose $y_n$ satisfies (4.3), and a job of value $x$ arrives at time $t$. If $x > y_n(t)$, then by assigning the job we can get an expected value of $r(t)x + E_{n-1}(t)$, which is more than the optimal expected value, $E_n(t)$, we can get if we bypass the job. Therefore, we should assign the job. Similarly, if $x \leq y_n(t)$ we should bypass the job and get an expected value of $E_n(t)$, greater than $r(t)x + E_{n-1}(t)$, the expected reward from assigning the job. This means $y_n(t)$ is the optimal critical curve.

Conversely, suppose $y_n$ is the optimal critical curve. Then the same reasoning as above shows that (4.3) must hold.

97
Now let \( s_n(t) = \sum_{i=1}^{n} y_i(t) \). Then by iterating (4.3), we see that \( y_1, \ldots, y_n \) are optimal if and only if they satisfy

\[
(4.4) \quad E_n(t) = r(t) s_n(t), \quad n \geq 1, \ t \geq 0.
\]

Now we find \( E_n(t; y_n, \ldots, y_1) \) for arbitrary critical curves \( y_n, \ldots, y_1 \). Let \( T_t(y_n) \) be the time, starting from \( t \), when the first man is assigned. That is,

\[
T_t(y_n) = \inf\{\tau > t : \text{a job arrives at } \tau \text{ with value } x > y_n(\tau)\}.
\]

Then \( T_t(y_n) \) is the first occurrence of an event, after time \( t \), in a variable time Poisson process with intensity \( \lambda(\tau) \bar{F}(y_n(\tau)) \) at time \( \tau \), where \( \bar{F} = 1 - F \). From the theory of variable time Poisson processes, we have

\[
(4.5) \quad P(T_t(y_n) \leq \tau) = 1 - \exp(-\int_t^\tau \lambda(\sigma) \bar{F}(y_n(\sigma)) d\sigma).
\]

Therefore,

\[
E_n(t; y_n, \ldots, y_1) = \int_t^\infty \left\{ r(\tau) \frac{H(y_n(\tau))}{\bar{F}(y_n(\tau))} + E_{n-1}(\tau; y_{n-1}, \ldots, y_1) \right\} d\tau P(T_t(y_n) \leq \tau),
\]

where \( H(y) = \int y F(dx) \). From (4.5) this gives
\[(4.6) \quad E_n(t; y_n, \ldots, y_1) \]

\[= \int_t^\infty \left\{ \lambda(\tau) \ r(\tau) \ H(y_n(\tau)) + \lambda(\tau) \ \bar{F}(y_n(\tau)) \ E_{n-1}(\tau; y_{n-1}, \ldots, y_1) \right\} \]

\[\cdot \ \exp(- \int_t^\tau \lambda(\sigma) \ \bar{F}(y_n(\sigma)) \ d\sigma) \ d\tau \ .\]

Hence from \((4.4)\), \(y_1, y_2, \ldots\) are optimal only if they satisfy

\[(4.7) \quad r(t) \ s_n(t) = \int_t^\infty \left\{ \lambda(\tau) \ r(\tau) \ H(y_n(\tau)) + \lambda(\tau) \ \bar{F}(y_n(\tau)) \ r(\tau) \ s_{n-1}(\tau) \right\} \]

\[\cdot \ \exp(- \int_t^\tau \lambda(\sigma) \ \bar{F}(y_n(\sigma)) d\sigma) d\tau, \quad n \geq 1 .\]

Note from \((4.7)\) that if \(r(t) \ s_{n-1}(t)\) is continuous and differentiable, then so is \(r(t) \ y_n(t)\), and hence \(r(t) \ s_n(t)\). But since \(r(t) \ s_0(t) = 0\), this means that if \(y_1, \ldots, y_n\) exist, then \(r y_1\) is continuous and differentiable, \(i \geq 1\).

For \(n = 1\), we differentiate \((4.7)\) to obtain

\[\frac{d r y_1}{dt} = -\lambda(t) \ r(t) \ H(y_1(t)) + \lambda(t) \ \bar{F}(y_1(t)) \int_t^\infty (\cdot) \ d\tau\]

\[= -\lambda(t) \ r(t) \ H(y_1(t)) + \lambda(t) \ \bar{F}(y_1(t)) \ r(t) \ y_1(t)\]

\[= -\lambda(t) \ r(t) \ \varphi(y_1(t)) ,\]

where \(\varphi(y) = H(y) - y \bar{F}(y)\). Now assume inductively that
\[
\frac{d r_n}{d t} = -\lambda(t) r(t) \phi(y_{n-1}(t)) .
\]

Next, differentiate (4.7) to obtain

\[(4.8) \quad \frac{d r_n}{d t} = -\lambda(t) r(t) \mathcal{H}(y_n(t)) - \lambda(t) \bar{F}(y_n(t)) r(t) s_{n-1}(t) + \lambda(t) \bar{F}(y_n(t)) \int_{t}^{\infty} \, dt - \frac{d r_{n-1}}{d t}
\]

\[= -\lambda(t) r(t) \left( \mathcal{H}(y_n(t)) + \bar{F}(y_n(t)) s_{n-1}(t) \right) + \lambda(t) \bar{F}(y_n(t)) \left( r(t) y_n(t) + r(t) s_{n-1}(t) \right) + \lambda(t) r(t) \phi(y_{n-1}(t))
\]

\[= -\lambda(t) r(t) \left( \phi(y_n(t)) - \phi(y_{n-1}(t)) \right) .
\]

From this it follows that

\[
\frac{d r_{n}}{d t} = -\lambda(t) r(t) \phi(y_n(t)) ,
\]

so that the induction is complete. Thus (4.8) gives us a differential equation for the optimal \( y_n \) in terms of the optimal \( y_{n-1} \), where \( y_0 = 0 \).

**Proposition 4.1:** The optimal \( y_n, \ldots, y_1 \) satisfy

\[(4.9) \quad y_{n+1}(t) \leq y_n(t), \quad n \geq 1, \quad t \geq 0
\]
and also

$$\frac{dr(t) y_n(t)}{dt} \leq 0, \quad n \geq 1, \ t \geq 0. \quad (4.10)$$

Proof: First we show that $\phi(y)$ is decreasing in $y$. Let $y_1 \leq y_2$.

Then

$$\phi(y_1) - \phi(y_2) = \int_{y_1}^{y_2} x F(dx) - y_1 \bar{F}(y_1) - \int_{y_1}^{y_2} x F(dx) + y_2 \bar{F}(y_2)$$

$$= \int_{y_1}^{y_2} x F(dx) + (y_2 - y_1) - y_2 F(y_2) + y_1 F(y_1)$$

$$\geq y_1 (F(y_2) - F(y_1)) + y_2 - y_1 - y_2 F(y_2) + y_1 F(y_1)$$

$$= (y_2 - y_1) (1 - F(y_2)) \geq 0.$$

In fact, $\phi$ decreases from $\mu$ to 0 as $y$ increases from 0 to $+\infty$.

We already have

$$\frac{dry_1}{dt} = -\lambda(t) r(t) \phi(y_1(t)) \leq 0,$$

so (4.10) holds for $n = 1$. Now we proceed by induction. Suppose (4.10) holds for $n$. Then from (4.7) we have
\[ r(t) s_{n+1}(t) \]
\[ = \int_r^\infty \left\{ \frac{r(\tau) \cdot H(y_{n+1}(\tau))}{\bar{H}(y_{n+1}(\tau))} + r(\tau) \cdot s_{n-1}(\tau) \right\} d\tau \cdot P(T_t(y_{n+1}) \leq \tau) \]
\[ + \int_t^\infty r(\tau) y_n(\tau) d\tau \cdot P(T_t(y_{n+1}) \leq \tau) \]
\[ = E_n(t; y_{n+1}, y_{n-1}, \ldots, y_1) + \int_t^\infty r(\tau) y_n(\tau) d\tau \cdot P(T_t(y_{n+1}) \leq \tau). \]

But since \( y_{n+1}, y_{n-1}, \ldots, y_1 \) are not optimal for the \( n \) man problem,

\[ E_n(t; y_{n+1}, y_{n-1}, \ldots, y_1) \leq E_n(t; y_n, \ldots, y_1) \]
\[ = E_n(t) = r(t) s_n(t). \]

Also, since (4.10) holds for \( n \),

\[ \int_t^\infty r(\tau) y_n(\tau) d\tau \cdot P(T_t(y_{n+1}) \leq \tau) \leq r(t) y_n(t) \int_t^\infty d\tau \cdot P(T_t(y_{n+1}) \leq \tau) \]
\[ \leq r(t) y_n(t). \]

Hence

\[ r(t) s_{n+1}(t) \leq r(t) s_n(t) + r(t) y_n(t). \]

If \( t < A \), then \( r(t) > 0 \), and hence \( y_{n+1}(t) \leq y_n(t) \). If \( t \geq A \),
\[ y_{n+1}(t) = y_n(t) = 0. \]

This proves (4.9), providing we can complete the induction. But
\[ \frac{d \phi_n^{n+1}}{dt} = - \lambda(t) r(t) (\phi(y_{n+1}(t)) - \phi(y_n(t))) \leq 0 \]

since \( y_{n+1}(t) \leq y_n(t) \) and \( \phi \) is decreasing. This proves (4.10) for \( n+1 \), and the proof is complete.

**Proposition 4.2:** Suppose \( y_{n-1}, \ldots, y_1 \) are optimal, and let \( z_n \) be any non-negative function which vanishes on \( [A, \infty) \). Define \( \tilde{z}_n \) by

\[ r(t) \tilde{z}_n(t) + r(t) s_n(t) = E_n(t; z_n, y_{n-1}, \ldots, y_1) \]

Then \( \tilde{z}_n(t) \leq y_{n-1}(t), t \geq 0 \).

**Proof:** As in the proof of Proposition 4.1, we have

\[ E_n(t; z_n, y_{n-1}, \ldots, y_1) \]

\[ = E_{n-1}(t; z_n, y_{n-2}, \ldots, y_1) + \int_t^\infty r(\tau) y_{n-1}(\tau) d\tau P(T_t(z_n) \leq \tau) \]

\[ \leq E_{n-1}(t) + r(t) y_{n-1}(t) \int_t^\infty d\tau P(T_t(z_n) \leq \tau) \]

\[ \leq r(t) s_{n-1}(t) + r(t) y_{n-1}(t) , \]

which means that

\[ r(t) \tilde{z}_n(t) \leq r(t) y_{n-1}(t) , \]
or

\[ z_n(t) \leq y_{n-1}(t), \quad t \geq 0. \]

This completes the proof.

Now we show the relationship between solutions of (4.7) and (4.8) under the assumption that

\[ \int_0^\infty \lambda(t) r(t) \, dt < \infty. \quad (4.11) \]

**Theorem 4.4:** Assume (4.11) holds. Let \( y_1, y_2, \ldots \) be the optimal critical curves, so that they satisfy (4.7) for each \( n \geq 1 \). Then \( r(t) y_n(t) \to 0 \) as \( t \to \infty \), \( n \geq 1 \). Conversely, suppose \( y_1, \ldots, y_{n-1} \) are optimal and that \( y_n(t) \) satisfies

(a) \( r(t) y_n(t) \to 0 \) as \( t \to \infty \)

(b) \[
\frac{dy_n}{dt} = -\lambda(t) r(t) (\phi(y_n(t)) - \phi(y_{n-1}(t))) .
\]

Then \( y_n \) satisfies (4.7).

**Proof:** Suppose \( y_1, y_2, \ldots \) are optimal. Then they satisfy (4.7) for \( n \geq 1 \), and by Proposition 4.1, \( y_{n+1}(t) \leq y_n(t) \) for \( n \geq 1 \). But

\[
r(t) y_1(t) = \int_t^\infty \lambda(\tau) r(\tau) H(y_1(\tau)) \exp(- \int_t^\tau \lambda(\sigma) \tilde{F}(y_1(\sigma)) \, d\sigma) \, d\tau
\]

\[
\leq \mu \int_t^\infty \lambda(\tau) r(\tau) \, d\tau \to 0 \quad \text{as} \quad t \to \infty ,
\]

104
by (4.11). Hence \( r(t) y_n(t) \leq r(t) y_1(t) \to 0 \) as \( t \to \infty \).

Conversely, suppose \( y_1, \ldots, y_{n-1} \) are optimal, and \( y_n \) satisfies (a) and (b) above. Define \( \tilde{y}_n(t) \) by

\[
(4.12) \quad r(t) \tilde{y}_n(t) = \int_{\tau}^{t} \left\{ \lambda(\tau) r(\tau) H(y_n(\tau)) + \lambda(\tau) \bar{F}(y_n(\tau)) r(\tau) s_{n-1}(\tau) \right\} \exp(- \int_{t}^{\tau} \lambda(\sigma) \bar{F}(y_n(\sigma)) d\sigma) d\tau - r(t) s_{n-1}(t). 
\]

We wish to show that \( \tilde{y}_n \equiv y_n \).

By Proposition 4.2, \( r(t) \tilde{y}_n(t) \leq r(t) y_{n-1}(t) \to 0 \) as \( t \to \infty \).

Differentiating (4.12) we get

\[
\frac{d\tilde{y}_n}{dt} = -\lambda(t) r(t) (H(y_n(t)) - \tilde{y}_n(t) \bar{F}(y_n(t))).
\]

Hence from (b),

\[
(4.13) \quad \frac{d(r(\tilde{y}_n - y_n))}{dt} = \lambda(t) r(t) \bar{F}(y_n(t)) (\tilde{y}_n(t) - y_n(t)).
\]

Suppose \( \tilde{y}_n(t_0) \neq y_n(t_0) \) for some \( t_0 \), and let \( t_1 \) be the smallest \( t \) larger than \( t_0 \) for which \( r(t) (\tilde{y}_n(t) - y_n(t)) = 0 \).

\( t = \infty \) works if nothing else will, since \( r\tilde{y}_n \to 0 \) and \( ry_n \to 0 \) as \( t \to \infty \). Then integrating (4.13) from \( t_0 \) to any \( t \in (t_0, t_1) \) gives

\[
\ln(r(t)|\tilde{y}_n(t) - y_n(t)|) = \ln(r(t_0)|\tilde{y}_n(t_0) - y_n(t_0)|) + \int_{t_0}^{t} \lambda(\tau) \bar{F}(y_n(\tau)) d\tau.
\]
By letting \( t \to t_1 \), the left side goes to \(-\infty\), whereas the right side remains bounded or goes to \(+\infty\), which is a contradiction. Hence \( \tilde{y}_n = y_n \) and the proof is complete.

Thus our problem of finding a solution \( y_n \) of the integral equation (4.7), given the optimal \( y_{n-1}, \ldots, y_1 \), is reduced to finding a solution of the differential equation (4.8) which satisfies \( r(t) y_n(t) \to 0 \) as \( t \to \infty \). We now show that such a solution exists, and that it is not only unique but also optimal.

**Theorem 4.5:** Suppose the optimal \( y_{n-1}, \ldots, y_1 \) exist and are unique. Then a solution \( y_n \) to

\[
\frac{d r_n}{dt} = -\lambda(t) r(t) (\phi(y_n(t)) - \phi(y_{n-1}(t)))
\]

exists with \( r(t) y_n(t) \to 0 \) as \( t \to \infty \), and it is unique. Furthermore, it is optimal for the \( n \) man problem.

**Proof.** Suppose \( y_n \) and \( \tilde{y}_n \) are two such solutions. Then for \( t < A \),

\[
(4.14) \quad \frac{dr(\tilde{y}_n - y_n)}{dt} = -\lambda(t) r(t) (\phi(\tilde{y}_n(t)) - \phi(y_n(t)))
\]

\[
= -\lambda(t) r(t) (\tilde{y}_n(t) - y_n(t)) \phi'(\xi_t),
\]

where \( \phi'(\xi_t) \) is the derivative of \( \phi \) evaluated at some point \( \xi_t \) between \( \tilde{y}_n(t) \) and \( y_n(t) \). Now suppose \( \tilde{y}_n(t_0) \neq y_n(t_0) \) for some \( t_0 < A \), and let \( t_1 \) be the smallest \( t \) larger than \( t_0 \) for which
\( \tilde{y}_n(t) = y_n(t). \) (Again \( t = \infty \) works if nothing else will.) Then by integrating (4.14) from \( t_0 \) to some point \( t \in (t_0, t_1) \), we get

\[
\ln(r(t) | \tilde{y}_n(t) - y_n(t)|) = \ln(r(t_0) | \tilde{y}_n(t_0) - y_n(t_0)|) - \int_{t_0}^{t} \lambda(\tau) \phi'(\xi(\tau)) \, d\tau.
\]

Letting \( t \to t_1 \), the left side goes to \(-\infty\), whereas the right side remains bounded or goes to \( +\infty \), since \( \phi' \leq 0 \). This is a contradiction, and hence \( \tilde{y}_n = y_n \).

Therefore, by Theorem 4.4, \( y_n \) is a solution to (4.7). Also, by Proposition 4.2 and the first part of this theorem, \( y_n \) is the unique solution to (4.7). Thus, \( y_n \) is optimal, if an optimal \( y_n \) exists. However, by Theorem 4.3, an optimal \( y_n \), finite for all \( t \geq 0 \), will exists if we can show that \( E_n(t) \) is finite for all \( t \geq 0 \). We do this inductively.

Suppose \( E_m(t) \leq mB, 0 \leq m \leq n-1, t \geq 0 \), where

\[
B = \int_{0}^{\infty} \lambda(t) r(t) \, dt < \infty.
\]

Then for any \( y_{n'}, \ldots, y_1 \), we have

\[
E_n(t; y_{n'}, \ldots, y_1)
= \int_{0}^{\infty} \lambda(\tau) r(\tau) H(y_n(\tau)) \exp(-\int_{t}^{\tau} \lambda(\sigma) \tilde{F}(y_n(\sigma)) \, d\sigma) \, d\tau
+ \int_{t}^{\infty} E_{n-1}(\tau; y_{n-1}, \ldots, y_1) \, d\tau P(T_t(y_n) \leq \tau)
\leq \mu \int_{0}^{\infty} \lambda(\tau) r(\tau) \, d\tau + \int_{t}^{\infty} (n-1) \mu B \, d\tau P(T_t(y_n) \leq \tau)
\leq \mu B + (n-1) \mu B
= n\mu B.
\]

107
Thus the inductive step is proved, an optimal $y_n$ exists, and it satisfies the differential equation (4.8), with $r(t) y_n(t) \to 0$ as $t \to \infty$. This completes the proof of the theorem.

**Examples:**

1. Suppose $\lambda(t)$ is a constant, which we take to be 1 for simplicity. Let $r(t) = e^{-\alpha t}$. Then the critical curves are obtained from

$$
\frac{dy_n(t)}{dt} = e^{-\alpha t} - \alpha e^{-\alpha t} y_n(t) = e^{-\alpha t} (\rho(y_n(t)) - \rho(y_{n-1}(t))) ,
$$
or

$$(4.15) \quad \frac{dy_n(t)}{dt} - \alpha y_n(t) = -\rho(y_n(t)) + \rho(y_{n-1}(t)) .$$

Since $\rho(y) = \int_{y}^{\infty} F(x) dx$ (after integration by parts), it is easy to see that the constants determined by the recursion

$$
\alpha y_1 = \int_{y_1}^{\infty} F(x) dx ,
$$

$$
\alpha(y_1 + \cdots + y_n) = \int_{y_n}^{\infty} F(x) dx , \quad n > 1
$$
satisfy (4.15). Furthermore, since $r(t) y_n(t) \to 0$ as $t \to \infty$, these constant solutions are the optimal critical curves. More will be said about this particular $r(t)$ function later on.
2. Again, suppose \( \lambda(t) = 1 \), and let

\[
\begin{align*}
    r(t) = \begin{cases} 
        1 & \text{if } 0 \leq t \leq T \\
        0 & \text{if } t > T
    \end{cases}
\end{align*}
\]

Since \( r(t) y_n(t) \) is continuous everywhere when \( y_n \) is optimal, we must impose the restriction that \( \lim_{t \uparrow T} y_n(t) = y_n(T) = 0 \), for all \( n \geq 1 \).

Now suppose \( n = 1 \) and \( t \in [0, T] \). Then we must solve

\[
\frac{dy_1(t)}{dt} = -\phi(y_1(t)) ,
\]

which, after imposing the restriction \( y_1(T) = 0 \), has the implicit solution

\[
y_1(t) = \int_0^t \frac{1}{\phi(x)} \, dx = T - t .
\]

For example, if \( F(x) = 1 - e^{-x} \), then \( \phi(x) = e^{-x} \), and

\[
\int_0^y \frac{1}{\phi(x)} \, dx = e^y - 1 ,
\]

so that

\[
y_1(t) = \ln(T - t + 1) .
\]

If \( F(x) = x, 0 \leq x \leq 1 \), then \( \phi(x) = (1-x)^2/2 \), and
\[
\int_0^y \frac{dx}{\varrho(x)} = \frac{2y}{1-y}.
\]

This gives the solution

\[y_1(t) = \frac{T-t}{2tT-t}.
\]

Note that for \( n = 1 \), the differential equation (4.16) is easily solved. The only difficulty is in evaluating the integral (4.17). For \( n > 1 \), the problem becomes much more difficult, however, since the differential equation cannot in general be solved by elementary methods.

Again let \( F(x) = 1 - e^{-x} \), so that the differential equation (4.8) becomes

\[
(4.18) \quad \frac{dy_n(t)}{dt} + e^{-y_n(t)} + n y_n(t) = \varrho(y_{n-1}(t)), \quad 0 \leq t \leq T
\]

Suppose that \( y_{n-1} \) has been found. Then the above non-linear differential equation can be made linear by letting \( y_n(t) = \ln u(t) \), with \( u(T) = 1 \). Then (4.18) becomes

\[
\frac{du}{dt} + g(t) u(t) = -1, \quad 0 \leq t \leq T,
\]

where \( g(t) = -\varrho(y_{n-1}(t)) \). This equation has an integrating factor

\[
h(t) = \exp(- \int_T^t g(x) \, dx),
\]

which implies

110
\[ \frac{d}{dt} (h(t) u(t)) = -h(t) , \]

or

\[ h(T) u(T) - h(t) u(t) = - \int_t^T h(x) \, dx . \]

Since \( h(T) = u(T) = 1 \), and \( y_n(t) = \ln u(t) \), we end up with

\[ y_n(t) = \ln \left( \frac{1}{h(t)} (1 + \int_t^T h(x) \, dx) \right) . \]

For example, if \( n = 2 \), then

\[ h(t) = \exp(\int_t^T \phi(y_1(x)) \, dx) \]

\[ = \exp(\int_t^T \frac{1}{T-x+1} \, dx) \]

\[ = \exp(\ln(T-t+1)) \]

\[ = T-t+1 , \]

which gives

\[ y_2(t) = \ln \left( \frac{1}{T-t+1} (1 + \int_t^T \frac{1}{T-x+1} \, dx) \right) \]

\[ = \ln \left( 1 + \frac{1}{2} \frac{(T-t)^2}{T-t+1} \right) . \]

Unfortunately, the integration becomes unwieldy for larger \( n \).

The situation is even worse if \( F(x) = x \), \( 0 \leq x \leq 1 \). In that case the differential equation (4.6) becomes
\[ \frac{dy_n(t)}{dt} + \frac{1}{2} (1 - y_n(t))^2 = \frac{1}{2} (1 - y_{n-1}(t))^2 \, . \]

Even for \( n = 2 \), if we let \( u = 1 - y_n \), the equation becomes

\[ \frac{du}{dt} - \frac{1}{2} u^2 = -\frac{1}{2} \left( \frac{2}{2n-1} \right)^2 \, . \]

This is not easy, but can be solved by several more substitutions. Then for larger \( n \), the right side becomes more complex, which makes an explicit solution that much more difficult to find.

3. As a last example, suppose \( \lambda(t) \) is not a constant.

It is then convenient to change the time scale by the transformation

\[ u(t) = \int_0^t \lambda(\tau) \, d\tau \, . \]

Then as \( t \) goes from \( 0 \) to \( \infty \), \( u \) goes from \( 0 \) to \( U = \int_0^\infty \lambda(\tau) \, d\tau \). Note that \( U \) is the expected number of jobs arriving in \([0, \infty)\), and that \( U \) may be infinite or finite.

Now if we let \( \tilde{y}_n(u) = y_n(t) \) and \( \tilde{r}(u) = r(t) \), the differential equation (4.8) becomes

\[ \frac{d\tilde{y}_n}{du} = -\tilde{r}(u) \left( \phi(\tilde{y}_n(u)) - \phi(\tilde{y}_{n-1}(u)) \right) \, , \]

for \( 0 \leq u \leq U \). Also condition (4.11) becomes
\[ \int_{0}^{\infty} r(u) \, du < \infty. \]

As a result of this transformation, we are back in the case where \( \lambda(t) = 1 \). In particular, the results of Example 2 hold with \( t \) replaced by \( u \) and \( T \) replaced by \( u(T) \).

We now turn to the problem we are really interested in, namely, where the men are not all identical. As usual, we assume we have men \( p_1 \leq \cdots \leq p_n \) (\( p_i > 0 \), \( 1 \leq i \leq n \)), and if we assign man \( p_i \) to a job of value \( x \) at time \( t \), we receive a reward \( p_i r(t)x \). This problem seems to be much more complicated than the problem with identical men, because we not only have to decide whether to assign a job, but we must also decide which man to assign it to. However, it will be shown how this problem is directly related to the problem with identical men.

We first note that by linearity, the optimal critical curve to use when there are \( n \) identical men of value \( p > 0 \) (\( p \) not necessarily equal to 1) is still the optimal \( y_n \) from the last section. Also, the optimal expected reward, starting from time \( t \), with these men is

\[ pE_n(t) = pr(t) (y_n(t) + \cdots + y_1(t)). \]

Now suppose the \( n \) men available at time \( t \) have values \( p_1 \leq \cdots \leq p_n \). By Lemma 2.3 we know there exist numbers

\[ 0 = \tilde{z}_{0,n+1}(t) \leq \tilde{z}_{1,n+1}(t) \leq \cdots \leq \tilde{z}_{n,n+1}(t) \leq \tilde{z}_{n+1,n+1}(t) = +\infty, \]

113
such that if the next job arrives at time $t$ with value
\[ x < z_{i+1, n+1}(t), \quad i \geq 1, \]
then it is optimal to assign man $p_i$ to the job, and if $0 \leq x \leq z_{1, n+1}(t)$, it is optimal not to assign the job to anyone. Therefore, we consider only the policies of this form.

From here on, it will be convenient to index the critical curves backwards, i.e., we will follow policies of the form given above, except with $z_{i, n+1}(t) = z_{n+1-l, 1}(t)$, so that
\[ 0 = z_{n+1, n+1}(t) \leq z_{n, n+1}(t) \leq \cdots \leq z_{1, n+1}(t) \leq z_{0, n+1}(t) = +\infty. \]

Let $z_n(t) = (z_{n, n+1}(t), \ldots, z_{1, n+1}(t))$ and let $p = (p_1, \ldots, p_n)$. Now suppose we follow a policy of the above form, which we may denote by $(z_n(t), \ldots, z_1(t))$. This means we use the vector $z_n$ until the first man is assigned, then $z_{n-1}$ until the second man is assigned, and so on. Then the expected reward, starting from time $t$ with the vector of men $p$, will be denoted by $E_n(t; p, z_n, \ldots, z_1)$, and the optimal $n$ man reward will be denoted by $E_n(t; p)$.

We now proceed in two stages. First, we show that if $z^*_i = (y_1, \ldots, y_i), \quad 1 \leq i \leq n,$ where the $y_i$'s are the optimal critical curves from the previous section, then for any vector $p$,
\[ E_n(t; p, z^*_n, \ldots, z^*_1) = \sum_{i=1}^{n} p_i r(t) y_{n+1-i}(t). \]

Secondly, we show that for any $p, z_n, \ldots, z_1$. 

114
\[ E_n(t; p, z_1, \ldots, z_1) \leq \sum_{i=1}^{n} p_i \ r(t) \ y_{n+1-i}(t), \]

which means that \( E_n(t; p) = \sum_{i=1}^{n} p_i \ r(t) \ y_{n+1-i}(t) \) and that the policy \( z_1^*, \ldots, z_1^* \) is optimal, independent of the vector \( p \).

**Theorem 4.6:** If \( z_1^*, \ldots, z_1^* \) are defined as above, and \( p = (p_1, \ldots, p_n) \) is any vector with \( 0 < p_1 \leq \cdots \leq p_n \), then

\[ E_n(t; p, z_1^*, \ldots, z_1^*) = \sum_{i=1}^{n} p_i \ r(t) \ y_{n+1-i}(t). \]

**Proof:** We modify our assignment problem by assigning \( n \) \( p_1 \)'s, \( n-1 \) \( (p_2-p_1) \)'s, \( n-2 \) \( (p_3-p_2) \)'s, \ldots, and \( 1 \) \( (p_n-p_{n-1}) \). We assign the \( n-i \) \( (p_{i+1}-p_i) \)'s as follows. First, they are assigned independently of the other \( (p_{j+1}-p_j) \)'s, \( i \neq j \). In particular, they are assigned as \( n-i \) identical men would optimally be assigned, that is, according to the critical curves \( y_{n-i} \), then \( y_{n-i-1} \), \ldots, and finally \( y_1 \). Therefore, from the previous section, the expected reward from these \( (p_{i+1}-p_i) \)'s, starting from time \( t \), is

\[ (p_{i+1}-p_i) \ E_{n-i}(t) = (p_{i+1} - p_i) \ r(t) \ s_{n-i}(t). \]

This means the expected reward from all of \( (p_{i+1}-p_i) \)'s, \( 0 \leq i \leq n-1 \) (with \( p_0 = 0 \)), starting from time \( t \), is

\[ \sum_{i=0}^{n-1} (p_{i+1}-p_i) \ r(t) \ s_{n-i}(t) = \sum_{i=1}^{n} p_i \ r(t) \ y_{n+1-i}(t). \]
Now we notice that for every realization of the process, the reward from the policy $z^*_n, \ldots, z^*_1$ is exactly the same as the reward from the modified assignment policy. For instance, if the first job after time $t$ arrives at time $t+\tau$ with value $y_{n-1}(t+\tau) < x < y_{n-2}(t+\tau)$, then in the modified policy a $p_1$ and a $p_2 - p_1$ are assigned (simultaneously), with reward $r(t+\tau)(p_1x + (p_2 - p_1)x) = r(t+\tau)p_2x$, whereas in the $z^*$ policy, a $p_2$ is assigned, also with reward $r(t+\tau)p_2x$.

This means the expected rewards from the two policies are the same, and the theorem follows.

**Theorem 4.7**: For any vector $p = (p_1, \ldots, p_n)$ with $0 < p_1 \leq \cdots \leq p_n$,

$$E_n(t; p) = \sum_{i=1}^{n} p_i r(t) y_{n+1-i}(t).$$

Hence the policy $z^*_n, \ldots, z^*_1$ is optimal.

**Proof**: The theorem is true for $n = 1$ from the results of the last section. We now proceed by induction. Assume the theorem true for $n-1$. Then we need to show that for any

$$z^*_n(t) = (z^*_{n,n+1}(t), \ldots, z^*_{1,n+1}(t)), \quad z^*_{i+1,n+1}(t) \leq z^*_{i,n+1} \quad \text{all } i,$$

we have, for any $p$,

$$E_n(t; p, z^*_n, x^*_{n-1}, \ldots, z^*_1) \leq E_n(t; p, z^*_n, \ldots, z^*_1).$$

116
Again we consider a modified policy which gives a reward the same as the reward from following policy $z_n, z_{n-1}^*, \ldots, z_1^*$. As before, we split $p$ into $n$ $p_1$'s, $n-1$ $(p_{2-p_1})$'s, $\ldots$, and $1$ $(p_n-p_{n-1})$.

Then the $n-1$ $(p_{i+1} - p_i)$'s are assigned as $n-1$ identical men, using the critical curves $z_{n-i}, y_{n-i-1}, \ldots, y_1$. Using the notation from the last section, the expected reward from these men is,

$$(p_{i+1} - p_i) E_{n-i}(t; z_{n-i}, y_{n-i-1}, \ldots, y_1).$$

But this is not the optimal way to assign these men. In fact,

$$E_{n-i}(t; z_{n-i}, y_{n-i-1}, \ldots, y_1) \leq (p_{i+1} - p_i) E_{n-i}(t; y_{n-i}, y_{n-i-1}, \ldots, y_1)$$

$$= (p_{i+1} - p_i) E_{n-i}(t)$$

$$= (p_{i+1} - p_i) r(t) s_{n-i}(t).$$

Thus an upper bound on the total expected reward from all of the men in the modified process, starting at time $t$, is

$$\sum_{i=0}^{n-1} (p_{i+1} - p_i) r(t) s_{n-i}(t) = \sum_{i=1}^{n} p_i r(t) y_{n-i-1}(t).$$

However, the reward from the modified process is again the same, for every realization, as the reward from using the policy $z_n^*, z_{n-1}^*, \ldots, z_1^*$. Hence, we have
\[ E_n(t; p, z_n, z_{n-1}^*, \ldots, z_1^*) \leq \sum_{i=1}^{n} p_i \ r(t) \ y_{n+1-i}(t) \]

for each \( z_n \) and each \( p \). The fact that \( z_n^*, \ldots, z_1^* \) are optimal then follows from Theorem 4.6.

Note that there is no double subscripting on the optimal critical curves. That is, if we have \( n+1 \) men, the optimal curves are \( y_{n+1}, \ldots, y_1 \), and after we assign one of the men, the optimal curves are simply \( y_n, \ldots, y_1 \). Therefore, the whole difficulty lies in solving the differential equations for the \( y_i \)'s.

An important aspect of the two preceding theorems is that they make no use of the arrival time distribution of the jobs. They use only the facts that (1) the \( y_i \)'s are the optimal critical curves for the identical men problem, and (2) the reward function is linear in \( p \). Thus we may alter the arrival time distribution of the jobs, so that it is no longer a variable time Poisson process, and, providing the \( y_i \)'s can be found, the results of the two preceding theorems will still hold.

In particular, suppose jobs arrive according to a renewal process with interarrival distribution \( G \). Now the problems involved in finding the optimal \( y_i \)'s are much greater than before. In fact, suppose \( n = 1, p_1 = 1 \), and we follow an arbitrary critical curve \( y(t) \). Then it is not even clear how to find the expected reward from this policy, whereas this was easy in the Poisson case. Therefore, we make the simplifying assumption that \( r(t) = e^{-\alpha t} \).
With this assumption we may find the optimal $y_1$'s as follows. First we define $E_n(t)$ to be the optimal expected reward from $n$ identical men (with $p_1$'s = 1), starting at time $t$, where $t$ is a time right after a job arrived, and where $E_n(t)$ does not include the reward from this job. Another way of saying this is that $E_n(t)$ is the reward from $n$ men, starting at time 0 and using the discount function $r(t') = e^{-\alpha(t + t')}$. Then by the form of $r(t)$ we have for any $t_0$, $t \geq 0$,

\[(4.19) \quad e^{-\alpha(t - t_0)} E_n(t_0) = E_n(t) \]

Also, let $y_n(t)$ be the critical curve at time $t$ when there are $n$ men left. Then as before, $y_n$, $\ldots$, $y_1$ are optimal if and only if they satisfy

\[(4.20) \quad e^{-\alpha t}(y_m(t) + \ldots + y_1(t)) = E_m(t), \quad 1 \leq m \leq n, \quad t \geq 0 \]

Using relations (4.19) and (4.20), we may derive the optimal policy explicitly.

**Theorem 4.8:** Let $\beta = \tilde{G}(\alpha)$, where $\tilde{G}(\alpha)$ is the Laplace transform of $G$, evaluated at $\alpha$. Then the optimal critical curves $y_n(t), \ldots, y_1(t)$ are constants, independent of $t$, and are found by the implicit recursion

\[(1 - \beta) (y_m + \ldots + y_1) = \beta \int_{y_m}^{\infty} \tilde{F}(x) dx, \quad 1 \leq m \leq n \]
Proof: We show that the optimal \( y_i \)'s are constants by induction.

From (4.19) and (4.20) we have

\[
e^{-\alpha t} y_1(t) = E_1(t) = e^{-\alpha(t-t_0)} E_1(t_0) = e^{-\alpha(t-t_0)} e^{-\alpha t_0} y_1(t_0),
\]

which implies

\[
e^{-\alpha t} y_1(t) = e^{-\alpha t} y_1(t_0),
\]

or

\[
y_1(t) = y_1(t_0) \quad \text{for all } t, t_0 \geq 0.
\]

Now assume the optimal \( y_i \) is a constant, \( 1 \leq i \leq n-1 \). Then from (4.19) and (4.20), we have

\[
e^{-\alpha t} (y_n(t) + y_{n-1} + \cdots + y_1)
= e^{-\alpha(t-t_0)} e^{-\alpha t_0} (y_n(t_0) + y_{n-1} + \cdots + y_1),
\]

which implies

\[
e^{-\alpha t} y_n(t) = e^{-\alpha t} y_n(t_0),
\]

or

\[
y_n(t) = y_n(t_0), \quad \text{for all } t, t_0 \geq 0.
\]

Therefore, in the rest of the proof we need only consider policies whose critical curves are constants.
Now let \( y \) be any positive constant and let \( T_y \) be the amount of time, starting from right after a job arrives, until another job arrives with value greater than \( y \). We need the distribution of \( T_y \). Let \( P_y(t) = P(T_y \leq t) \). Then by conditioning on the time of the next job,

\[
1 - P_y(t) = \bar{G}(t) + \int_0^t (1 - P_y(t-\tau)) F(y) G(d\tau),
\]
or

\[
P_y(t) = \bar{F}(y) G(t) + \int_0^t P_y(t-\tau) F(y) G(d\tau).
\]

Let \( K(t) = F(y) G(t) \). Then \( P_y \) satisfies

\[
P_y(t) = \bar{F}(y) G(t) + (P_y * K)(t),
\]
which has the solution

\[
P_y(t) = \bar{F}(y) G(t) + (\bar{F}(y) G * M_K)(t),
\]

where

\[
M_K(t) = \sum_{n=1}^\infty (F(y))^n G^{(n)}(t)
\]

and \( G^{(n)}(t) \) is the \( n \)-fold convolution of \( G \) with itself.

Taking Laplace transforms this gives

\[
\bar{P}_y(s) = \bar{F}(y) \bar{G}(s) + \bar{F}(y) \bar{G}(s) \bar{M}_K(s).
\]

But
\[ \hat{M}_K(s) = \sum_{n=1}^{\infty} (F(y) \bar{G}(s))^n = \frac{F(y) \bar{G}(s)}{1 - F(y) \bar{G}(s)}. \]

Therefore,

\[ \bar{y}_y(s) = \frac{\bar{F}(y) \bar{G}(s)}{1 - F(y) \bar{G}(s)}. \]

Now for \( n = 1 \), the optimal policy \( y_1 \) satisfies

\[ e^{-\alpha t} y_1 = E_1(t) = \int_0^\infty e^{-\alpha(t+\tau)} \frac{H(y_1)}{\bar{F}(y_1)} \frac{d\tau}{\bar{F}(y_1)} P(T_{y_1} \leq \tau) \]

\[ = e^{-\alpha t} \frac{H(y_1)}{\bar{F}(y_1)} \bar{y}_1(\alpha), \]

which gives

\[ y_1 = \frac{H(y_1) \bar{G}(\alpha)}{1 - F(y_1) \bar{G}(\alpha)}, \]

or

\[ (4.21) \quad y_1 = \bar{G}(\alpha) \left( H(y_1) + y_1 F(y_1) \right), \]

where as before

\[ H(y) = \int_y^\infty xF(dx). \]

Integrating by parts, (4.21) becomes

122
\[ y_1 = G(\alpha) \left( y_1 + \int_{y_1}^{\infty} \bar{F}(x) \, dx \right), \]

or

\[ (1-\beta) \, y_1 = \beta \int_{y_1}^{\infty} \bar{F}(x) \, dx, \quad \text{with} \quad \beta = C(\alpha). \]

Now suppose that the optimal \( y_{n-1}, \ldots, y_1 \) are determined by

\[ (1-\beta) \, (y_1 + \cdots + y_m) = \beta \int_{y_m}^{\infty} \bar{F}(x) \, dx, \quad 1 \leq m \leq n-1. \]

Then from (4.20), the optimal \( y_n \) satisfies

\[ e^{-\alpha t}(y_1 + \cdots + y_n) = \int_{0}^{\infty} e^{-\alpha(t+\tau)} \left( \frac{H(y_n)}{\bar{F}(y_n)} + (y_1 + \cdots + y_{n-1}) \right) \, d\tau \, P(T_{y_n} < \tau) \]

or

\[ y_1 + \cdots + y_n = \left( \frac{H(y_n)}{\bar{F}(y_n)} + (y_1 + \cdots + y_{n-1}) \right) \bar{F}(y_n) (\alpha). \]

Letting \( s_i = y_1 + \cdots + y_i \), we have

\[ y_n + s_{n-1} = \frac{H(y_n)\beta}{1 - \beta \bar{F}(y_n)} + s_{n-1} \frac{\bar{F}(y_n)\beta}{1 - \beta \bar{F}(y_n)}, \]

which reduces to

\[ (1-\beta) \, s_{n-1} + y_n = \beta (H(y_n) + y_n \bar{F}(y_n)) \]
Again we integrate the right side by parts and manipulate slightly to obtain

\[(1-\beta) s_n = \beta \int_{y_n}^{\infty} \bar{F}(x) \, dx.\]

Now it is easy to check that a (finite) optimal \( y_n, \ldots, y_1 \) exists, and also that (4.22) has a unique solution. Therefore this solution is optimal, and the proof is complete.

**Theorem 4.2:** Suppose men \( 0 < p_1 \leq \cdots \leq p_n \) are available and a job of value \( x \) arrives at time \( t \). Then under the same reward structure as above, with \( r(t) = e^{-\alpha t} \), it is best to assign man \( p_i \), \( i \geq 1 \) if and only if \( y_{n-i+1} < x \leq y_{n-i} \), and it is best not to assign the job at all if and only if \( x < y_n \). The reward from this policy, starting at a time \( t \) just after a job has arrived (but not counting the reward from this job) is

\[ e^{-\alpha t} \sum_{i=1}^{n} p_i y_{n+1-i}. \]

Also these \( y_i \)'s are independent of the \( p_i \)'s. Furthermore, they are determined by the recursion

\[(1-\beta) (y_1 + \cdots + y_m) = \beta \int_{y_m}^{\infty} \bar{F}(x) \, dx, \quad 1 \leq m \leq n,
\]

where \( \beta = \bar{G}(\alpha) \).
Proof: The proof follows directly from Theorems 4.6, 4.7, and 4.8, and from the comments following Theorem 4.7.

The optimal $y_i$'s may be found graphically from the following diagram:

![Graph showing the optimal $y_i$'s](image)

This graph shows that the sequence $(y_i)$ is decreasing in $i$. We now show that the $y_i$'s have another intuitive property.

**Proposition 4.3:** For $0 \leq \beta_1 \leq \beta_2 \leq 1$, $y_n(\beta_1) \leq y_n(\beta_2)$, $n \geq 1$, where $y_n(\beta_1)$ is the optimal $y_n$ determined from $\beta = \beta_1$.

**Proof:** We proceed by induction on $n$. For $n = 1$, we have

$$(1-\beta) y_1(\beta) = \beta \int_{y_1(\beta)}^{\infty} \tilde{F}(x) \, dx.$$ 

Differentiating both sides with respect to $\beta$ gives
\[(1-\beta) \ y_1'(\beta) - y_1(\beta) = \int_{y_1(\beta)}^{\infty} \tilde{F}(x) \ dx - \beta \ y_1'(\beta) \ \tilde{F}(y_1(\beta)) \ ,\]

or

\[
y_1'(\beta) = \frac{y_1(\beta) + \int_{y_1(\beta)}^{\infty} \tilde{F}(x) \ dx}{(1-\beta) + \beta \ \tilde{F}(y_1(\beta))} \geq 0 .
\]

Thus \( y_1 \) is increasing in \( \beta \).

Now suppose \( y_i'(\beta) \geq 0, \ i = 1, \ldots, n-1 \). Then

\[(1-\beta) \ s_n'(\beta) - s_n(\beta) = \int_{y_n(\beta)}^{\infty} \tilde{F}(x) \ dx - \beta \ s_n'(\beta) \ \tilde{F}(y_n(\beta)) \ ,\]

or

\[
y_n'(\beta) = \frac{s_n(\beta) + \int_{y_n(\beta)}^{\infty} \tilde{F}(x) \ dx - (1-\beta) \ s_{n-1}'(\beta)}{(1-\beta) + \beta \ \tilde{F}(y_n(\beta))}
\]

\[
= \frac{s_{n-1}(\beta) + \int_{y_{n-1}(\beta)}^{\infty} \tilde{F}(x) dx \ - (1-\beta) s_{n-2}'(\beta)}{(1-\beta) + \beta \ \tilde{F}(y_{n-1}(\beta))}
\]

But since

\[-\frac{(1-\beta)}{(1-\beta) + \beta \ \tilde{F}(y_{n-1}(\beta))} \geq -1 ,\]
we have

\[ y_n'(\beta) \geq \frac{s_n(\beta) + \int y_n(\beta) \bar{F}(x) \, dx - s_{n-1}(\beta) - \int y_{n-1}(\beta) \bar{F}(x) \, dx + \frac{(1-\beta)^2 s'_{n-2} (\beta)}{(1-\beta) \bar{F}(y_{n-1}(\beta))}}{(1-\beta) + \beta \bar{F}(y_n(\beta))} \cdot \gamma \cdot s_n(\beta) \]

where

\[ \gamma = \frac{(1-\beta)^2}{(1-\beta) + \beta \bar{F}(y_{n-1}(\beta))} \geq 0 \]

This completes the induction, and hence \( y_n \) is increasing in \( \beta, n \geq 1 \).

4.4. Jobs May be Bypassed; Time no Factor.

We now look at another "timeless" model which is both a strengthening and a weakening of the first model of this chapter. It is a weakening in the sense that we assume

\[ q_k = P(\text{exactly } k \text{ jobs arrive}) = q^k (1-q), \text{ some } 0 < q < 1. \]

That is, we assume that the process which governs the number of jobs is memoryless. However, this model is stronger than the previous
model because we allow jobs to be bypassed, even if there are men available to do them.

This model might be appropriate, for instance, when the "men" are houses to be sold, the jobs are "offers", the number of offers in the past does not affect our thinking as to how many offers will occur in the future, and time is of no importance.

The reward structure is as before. If a p man is assigned to an x job, a reward of px is obtained. We first assume all of the p's are the same, say, equal to 1. The next theorem gives the optimal policy in this case.

Theorem 4.10: Suppose the job values are non-negative i.i.d. random variables from a continuous distribution F, and that the men are identical with p_i's equal to 1. Then there exist numbers

\[ 0 \leq \cdots \leq a_n \leq \cdots \leq a_1 < \infty \]

such that if there are n men left and the next job has value x, then it is best to assign one of them to the job if and only if \( x > a_n \). The optimal expected reward from the n man problem is \( a_n + \cdots + a_1 \).

Furthermore, the \( a_i \)'s are determined by the recursion

\[
(4.23) \quad (1-q) (a_1 + \cdots + a_n) = q \int_{a_n}^{\infty} F(x) \, dx, \quad n \geq 1.
\]

128
Proof: We need only consider policies of the form described in the statement of the theorem, i.e., stationary critical number policies, because of Lemma 2.3 and the memoryless quality of the \(q \kappa\) distribution. We now proceed by induction to show that the \(a_1\)'s determined by (4.23) are the optimal critical numbers.

First, define \( E_n(a_n, \ldots, a_1) \) to be the expected \( n \) man reward from using the critical numbers \( a_n \), then \( a_{n-1}, \ldots \), and finally \( a_1 \). Also let \( E_n \) be the optimal expected \( n \) man reward. Then exactly as in Theorem 4.3, \( a_n, \ldots, a_1 \) are optimal if and only if they satisfy

\[
(4.24) \quad s_i = E_i, \quad 1 \leq i \leq n,
\]

where \( s_i = a_1 + \cdots + a_i \).

Now we calculate \( E_n(a_n, \ldots, a_1) \) for any \( a_n, \ldots, a_1 \).

We have

\[
E_n(a_n, \ldots, a_1) = q \left[ \int_0^{a_n} E_n(a_n, \ldots, a_1) F(dx) \right. \\
+ \left. \int_{a_n}^{\infty} (x + E_{n-1}(a_{n-1}, \ldots, a_1)) F(dx) \right],
\]

or

\[
E_n(a_n, \ldots, a_1) = \frac{q(\int_{a_n}^{\infty} x F(dx) + F(a_n) E_{n-1}(a_{n-1}, \ldots, a_1))}{1 - q F(a_n)}.
\]

129
The factor $q$ appears since $q$ is the probability that at least one more job will arrive. Thus from (4.24), the optimal $a_n, \ldots, a_1$ must satisfy

$$s_m = \frac{q \left( \int_{a_m}^{\infty} xF(dx) + \bar{F}(a_m) s_{m-1} \right)}{1 - qF(a_m)}, \quad 1 \leq m \leq n.$$  \hspace{1cm} (4.25)

For $n=1$, (4.25) becomes

$$a_1 = \frac{q \int_{a_1}^{\infty} xF(dx)}{1 - qF(a_1)},$$

or after integrating by parts and simplifying,

$$(1-q) a_1 = q \int_{a_1}^{\infty} \bar{F}(x) \, dx.$$  

Now suppose the optimal $a_{n-1}, \ldots, a_1$ satisfy

$$s_m = \frac{q \int_{a_m}^{\infty} \bar{F}(x) \, dx}{a_m}, \quad 1 \leq m \leq n-1.$$  \hspace{1cm} (4.26)

Then from (4.25), the optimal $a_n, \ldots, a_1$ satisfy

$$a_n + s_{n-1} = \frac{q \left( \int_{a_n}^{\infty} xF(dx) + \bar{F}(a_n) s_{n-1} \right)}{1 - qF(a_n)}.$$  

Again we integrate by parts and, using (4.26), we can simplify to obtain
\[(1-q)s_n = q \int_{a_n}^{\infty} \tilde{F}(x) \, dx .\]

Since this equation has a unique solution \(a_n\) and since it is easy to see that a (finite) optimal policy exists, the solution we have found is optimal.

We now proceed with the non-identical men problem exactly as we did in the preceding model. In fact the proofs of Theorem 4.6 and 4.7 still apply, and we have the following.

**Theorem 4.11:** Suppose men \(0 < p_1 \leq \cdots \leq p_n\) are left, and \(a_n', \ldots, a_1'\) are the optimal critical numbers from the Theorem 4.10. Then if the next job arrives with value \(a_{n-i+1} < x \leq a_{n-i}', \ 1 \leq i \leq n, \ a_0' = +\infty\), it is best to assign man \(p_i\); if \(x \leq a_n'\), it is best not to assign any of the men to this job. The reward from this optimal policy is

\[E_n(p) = \sum_{i=1}^{n} p_i a_{n+1-i}.\]

Before concluding this chapter, we bring attention to the remarkable similarity between Theorems 4.9 and 4.11. In Theorem 4.9 we consider a model with an explicit time variable, and we find that the optimal critical curves are really constants which must satisfy a certain recursion relation. In Theorem 4.11 we consider a model with
no explicit relation to time, and we find that the optimal critical numbers must satisfy the exact same recursion relation, with \( \bar{\alpha} \) replaced by \( q \).

This means that in the latter model we act in the same way as we would in the former model if \( G \) and \( \alpha \) were, respectively, any interarrival distribution and any discount rate which satisfied \( q = \bar{\alpha} \). Conversely, in the former model we act exactly as we would in the latter model with \( q = \bar{\alpha} \), that is, as if the probability of at least one more job ever arriving were always \( q \). Also, the optimal expected rewards, starting from time \( 0 \), are the same in the two models, namely
\[
\sum_{i=1}^{n} p_i a_{n+1-i} = \sum_{i=1}^{n} p_i y_{n+1-i}, \quad \text{when} \quad q = \bar{\alpha},
\]
and this is true for any positive \( p_i \)'s.
CHAPTER 5
SEVERAL MORE ASSIGNMENT PROBLEMS

In this chapter we consider several more stochastic assignment models which are in some sense less restrictive than those already studied. One feature of these will be that men become reavailable in some, possibly random, manner, so that the problem can continue indefinitely. This suggests looking for asymptotic results rather than the transitive results we obtained in Chapters 3 and 4. One unfortunate feature of the present models is, however, that because of their complexity, the results are usually not of such a nice form as those of the previous chapters. In fact, the optimal policies are usually not even independent of the p's, the values of the men. Hence, the optimal policies must often be found by dynamic programming or linear programming.

5.1. Model 1: Men Reavailable After Finishing Jobs.

The first model is as follows. We have a fixed number of men n, with values $\bar{p}_1 \leq \cdots \leq \bar{p}_n$. Jobs arrive according to a renewal process, with interarrival distribution G, and take on values $X_1, X_2, \ldots$, i.i.d. random variables from a distribution F. If a job arrives and $1 \leq k \leq n$ men are available, a choice is made,
depending on the value of the job $x$, as to which man to assign. We require that one of the men be assigned. If man $\tilde{p}_i$ is assigned, a reward $r(\tilde{p}_i, x)$ is obtained, where the assumptions regarding $r$ will be given later on.

Each man takes an exponential amount of time, with mean $1/\mu$, to do any job. This amount of time is independent of the value of the job, of the particular man doing the job, and of all other aspects of the problem. When he completes the job, he again becomes available to do another job. If a job arrives and no man is available, then this job is lost. The object is to maximize the expected reward per unit time.

If we view jobs as "customers" and men as "servers", the above problem has a structure exactly like a GI/M/$\infty$ queue, except that there is no waiting room, i.e., no queue is allowed to form. The important similarity, however, is in the sequence of busy and idle periods, where these are defined by the following diagram:

```
               busy cycle
               
               idle period    busy period
               
---+---+---
      |   |   |
      t_1 t_2 t_3
      ↑   ↑   ↑
all men become idle  next job arrives  first time after $t_2$
when all men are again idle
```

134
Notice that since all men complete jobs in (stochastically) the same amount of time, the length of a busy cycle \((t_2 - t_1)\) is the same no matter which assignments are made. This, plus the fact that no rewards are earned during an idle period, allows us to maximize expected reward per unit time by maximizing expected reward during a busy period.

Now a busy period can be viewed as consisting of two stages. The first stage is the beginning of the busy period. At this time an \(x\) is observed, and a decision between all \(n\) men must be made. This is the only time during the busy period when a decision must be made between all \(n\) men.

The second stage is the rest of the busy period, where all decisions, if any, must be made between \(k\) men, where \(0 \leq k \leq n-1\). (If \(k = 0\) or \(1\), there is really no decision to be made.) We first find the optimal decisions for all possible states, i.e., combinations of available men, in the second stage of the busy period. This will require no knowledge of how to make a decision in the first stage, for the following reason. Given that a point is reached where an incoming job of value \(x\) finds men \(\tilde{p}_{i_1} \leq \cdots \leq \tilde{p}_{i_k}\), \(0 \leq k \leq n-1\), idle, it is immaterial to the decision-maker how to make a decision at the beginning of the busy period, because he knows he will not have to make a first stage decision again until the next busy period. Having found the optimal decisions for the second stage, we then find the optimal decisions for the first stage.

We suppose that the reward function \(r(p, x)\) satisfies
(5.1) \[ r(p_2, x_2) + r(p_1, x_1) \geq r(p_2, x_1) + r(p_1, x_2) \] for all \( x_1 \leq x_2, \ p_1 \leq p_2 \)

Then we may apply Lemma 2.3 to deduce the form of the optimal policy.

**Lemma 5.1:** Suppose a job of value \( x \) arrives and men \( \tilde{p}_{i_1} \leq \cdots \leq \tilde{p}_{i_k} \) are available, \( 1 \leq k \leq n \). Then there exist numbers \( b_1 \leq \cdots \leq b_{k-1} \), with \( b_0 = -\infty \), \( b_k = +\infty \), such that it is best to assign man \( \tilde{p}_{i_j} \) if and only if \( b_{j-1} < x \leq b_j \).

**Proof:** Suppose man \( \tilde{p}_{i_j} \) is chosen. Then let \( a_{i_j} \) be the optimal expected reward from just after this decision until the end of the busy period. This \( a_{i_j} \) will depend on which men are working, which men are idle, and the parameters of the problem, but it is some number. Then we make our decision by finding

\[ \max_{1 \leq j \leq k} (r(\tilde{p}_{i_j}, x) + a_{i_j}) \]

By Lemma 2.3 there exist numbers \( b_1 \leq \cdots \leq b_{k-1} \), with \( b_0 = -\infty \), \( b_k = +\infty \), such that the above expression is maximized by \( r(\tilde{p}_{i_j}, x) + a_{i_j} \) if and only if \( b_{j-1} < x \leq b_j \). This completes the proof.

We now set up the second stage of the busy period as a dynamic programming problem. First of all, since there may be equality among some of the \( \tilde{p} \)'s, we suppose that for some \( 1 \leq m \leq n \), there are numbers \( p_1 \leq \cdots \leq p_m \) and \( n_1, \ldots, n_m \), with \( 1 \leq n_j \leq n \), \( n_1 + \cdots + n_m = n \).
such that \( n_j \) of the \( \mathbf{p} \)'s equal \( p_j \), \( 1 \leq j \leq m \). Thus there is no
difference between two men with the same \( p_i \).

During the second stage of the busy period, we focus our
attention on an embedded Markov chain, with time points \( \tau_1, \tau_2, \ldots \)
corresponding to times just after the successive jobs arrive. The
states of the Markov chain are of the form

\[
(i_1, \ldots, i_j),
\]

where \( 0 \leq j \leq m, 1 \leq r_k \leq n_k \) for \( 1 \leq k \leq j \), and \( 0 \leq \sum_{k=1}^{j} r_k \leq n-1 \).
This notation means that \( r_k \) men of value \( p_{i_k} \), \( 1 \leq k \leq j \), are idle
when we observe the chain, that is, just after a job arrives. In the
special case when \( j = 0 \) (everyone is busy), we will denote the state
by \( (0) \).

Now suppose a job of value \( x \) arrives and \( s = (i_1, \ldots, i_j), 2 \leq j \leq m \), is the state of the chain. Then by Lemma 5.1, we need only
consider policies of the form: assign man \( p_{i_k} \) if and only if

\[
b_s^{k-1} < x \leq b_s^k,
\]

where \(-\infty = b_s^0 \leq b_s^1 \leq \ldots \leq b_s^{j-1} \leq b_s^j = +\infty \). Furthermore, define

\[
 u_s^k = F(b_s^k) - F(b_s^{k-1}) = F(\text{assigning job to a } p_{i_k} \text{ man}),
\]

and

\[
 v_s^k = \int_{b_s^{k-1}}^{b_s^k} r(p_{i_k}, x) F(dx),
\]

137
for $1 \leq k \leq j$. We assume $r$ and $F$ are such that $\psi_p^{k_s}$ is finite for each $p_i$, $b_s^{k-1}$ and $b_s^k$.

Before proceeding, we need some further notation. Suppose $Z_1, \ldots, Z_{j+k}$ are i.i.d. exponential random variables with mean $1/\mu$. Also, suppose $Y$ has distribution $G$ and is independent of the $Z$'s.

Then let

$$q_{j,k} = P(\max_{1 \leq i \leq j} Z_i < Y < \min_{j+1 \leq i \leq j+k} Z_i).$$

In what follows, the $Z$'s will be times to complete jobs and $Y$ will be the time from right after the last job was assigned until the next job arrives. Then $q_{j,k}$ is the probability that, starting with $j+k$ men working, the first $j$ of them finish and the last $k$ are still working when the next job arrives, for a given permutation of the $j+k$ men. To compute $q_{j,k}$, we use the following lemma.

**Lemma 5.2:**

$$q_{j,k} = \int_{0}^{\infty} \int_{0}^{\infty} (G(y_2) - G(y_1)) \mu e^{-\mu y_1} (1 - e^{-\mu y_1})^{j-1} k \mu e^{-k \mu y_2} dy_1 dy_2,$$

for $j,k \geq 1$;

$$q_{0,k} = \int_{0}^{\infty} G(y_2) k \mu e^{-k \mu y_2} dy_2,$$

for $k \geq 1$;

$$q_{j,0} = \int_{0}^{\infty} (1 - G(y_1)) \mu e^{-\mu y_1} (1 - e^{-\mu y_1})^{j-1} dy_1$$

for $j \geq 1$. 

138
Proof: Let $Y_1 = \max_{1 \leq i \leq j} Z_i$, $Y_2 = \min_{j+1 \leq i \leq j+k} Z_i$. Then it is easily shown that the densities of $Y_1$ and $Y_2$ are

$$f_{Y_1}(y_1) = j\mu e^{-j\mu y_1} (1 - e^{-\mu y_1})^{j-1},$$

and

$$f_{Y_2}(y_2) = k\mu e^{-k\mu y_2}.$$

Thus we may find $P(Y_1 < Y < Y_2)$ by conditioning on $Y_1$ and $Y_2$, using independence, and using the above densities of $Y_1$ and $Y_2$. The formulas of the lemma then follow immediately.

Examples:

1. Suppose $G$ is exponential with mean $1/\lambda$. To evaluate the above formulas for $q_{j, k}$, we use the binomial expansion

$$(1 - e^{-\mu t})^{j-1} = \sum_{i=0}^{j-1} (-1)^i \binom{j-1}{i} e^{-i\mu t}.$$

Then the integrals may be evaluated explicitly to give

$$q_{j, k} = \frac{j\lambda \mu}{(\lambda + k\mu)} \sum_{i=0}^{j-1} (-1)^i \frac{(j-1)}{(\lambda + (i+1)k\mu)}, \quad j \geq 1, \ k \geq 0,$$

$$q_{0, k} = \frac{\lambda}{\lambda + k\mu}, \quad k \geq 1.$$

2. $G$ is degenerate, with unit mass at $a$. Then

$$q_{j, k} = (1 - e^{-\mu a})^j e^{-k\mu a}, \quad j, k \geq 0.$$
3. \( G \) is uniform on \([a,b]\). Then

\[
(b-a) q_{j,k} = (b-a) \left(1 - e^{-\alpha \mu j}\right) e^{-kb \mu}
+ \left(\int_a^b (y_2-a) \kappa \mu e^{-\kappa y_2} dy_2 \right) \left(1 - e^{-\mu a j}\right)
+ \left(\int_a^b \int_a^b (y_2-y_1) \kappa \mu e^{-\kappa y_2} j \mu e^{-\mu y_1} e^{-y_1} dy_1 dy_2 \right),
\]

for \(j,k \geq 0\). These integrals may be evaluated using the above binomial
expansion for \((1 - e^{-\mu y_1})^{j-1}\), and the fact that

\[
\int_a^b s e^{-cs} ds = \frac{1}{c} (a e^{-ca} - b e^{-cb}) + \frac{1}{c^2} (e^{-ca} - e^{-cb}).
\]

Now that this notation has been introduced, the second stage
of the busy period may be formulated as a Markov decision problem as
follows. Each time a job arrives, a decision \(a_s\) is made depending
upon the state of the system \(s\). Then an expected reward \(\tilde{w}(s,a_s)\)
is received from that time until the next job arrives, and the process
moves to state \(t\) the next time a job arrives with probability
\(\tilde{\pi}(t|s,a_s)\). From any state there is positive probability, independent
of \(a_s\), of going to the "stopping state", that is, where all men are
idle. When this occurs the busy period is over and the process comes
to an end; no further rewards are received. The stopping state will
not be shown explicitly in the transition matrix. Rather, its presence
will be seen implicitly because the transition matrix will have row
sums strictly less than 1.
Lemma 5.3: Suppose a job arrives and the state of the embedded Markov chain is \( s = (r_1^i, \ldots, r_j^i) \), where \( 1 \leq j \leq m, 2 \leq \sum_{k=1}^{1} r_k^i \leq n-1 \), and \( 1 \leq r_k^i \leq n_k i_k \) for \( 1 \leq k \leq j \). Also suppose we follow a critical number policy, as described above, which is determined by the vector \( (b_1^s, \ldots, b_j^s) \), where \( b_1^s \leq \cdots \leq b_j^s \). Then we have the following transition probabilities and expected rewards. If

\[
\begin{align*}
t = (i_1^l, \ldots, i^\sigma, \ldots, i_j^l, \ell_1^l, \ldots, \ell_j^l),
\end{align*}
\]

where \( 0 \leq s_k \leq n_k - r_k \) for \( 1 \leq k \leq j, k \neq \sigma; 1 \leq \sigma \leq j; 0 \leq j' \leq m-j; 1 \leq t_k \leq n_{k} \), for \( 1 \leq k \leq j' \); and the \( i_k \)'s are distinct from the \( \ell_k \)'s,

then

\[
\begin{align*}
\tilde{n}(t | s, a_s) = u_0^i \prod_{k=1}^{j} \left( \begin{array}{c} n_{i_k} - r_k \\ s_k \end{array} \right)^i \prod_{k=1}^{j'} \left( \begin{array}{c} n_{\ell_k} \\ t_k \end{array} \right)^{j'} q_i^\sigma \tau_1^\sigma \tau_2^\sigma
\end{align*}
\]

where

\[
\begin{align*}
\tau_1^\sigma &= \sum_{k=1}^{i} s_k + \sum_{k=1}^{j} t_k, \\
\tau_2^\sigma &= n+1 - \tau_1^\sigma - \sum_{k=1}^{i} r_k.
\end{align*}
\]

If \( t = (i_1^l, \ldots, i_j^l, \ell_1^l, \ldots, \ell_j^l) \), where \( 0 \leq s_k \leq n_k - r_k \) for \( 1 \leq k \leq j; 0 \leq j' \leq m-j; 1 \leq t_k \leq n_{k} \), for \( 1 \leq k \leq j' \);

\[
\sum_{k=1}^{i} (r_k + s_k) + \sum_{k=1}^{j} t_k \leq n-1; \text{ and the } i_k \text{'s are distinct from the } \ell_k \text{'s},
\]

then
\[
\pi(t | s, a_s) = \sum_{\sigma=1}^{i} \left( \begin{array}{c} n_{i_{\sigma}} + 1 - r_{\sigma} \\ s_{\sigma} + 1 \end{array} \right) \prod_{k=1}^{j} \left( \begin{array}{c} n_{i_k} - r_k \\ s_k \end{array} \right) \cdot \prod_{k=1}^{j'} \left( \begin{array}{c} n_{s_{l_k}} \\ t_k \end{array} \right) q_{\tau_1, \tau_2}.
\]

where

\[\tau_1 = \sum_{k=1}^{i} s_k + \sum_{k=1}^{j'} t_k + 1, \quad \tau_2 = n + 1 - \sum_{k=1}^{i} r_k - \tau_1.\]

If $t$ is the stopping state, then

\[
\pi(t | s, a_s) = q_{\tau, 0}, \quad \text{where} \quad \tau = n+1 - \sum_{k=1}^{i} r_k.
\]

Otherwise $\pi(t | s, a_s) = 0$. Furthermore, the expected rewards are given by

\[
\bar{w}(s, a_s) = \sum_{k=1}^{i} v_{s_{l_k}}^{k}.
\]

Suppose $s = (1^1)$ or $(0)$. If $t = (t_1^1, \ldots, t_{j'}^{j'})$, where $0 \leq j' \leq m$, $1 \leq t_k \leq \ell_{s_{l_k}}$ for $1 \leq k \leq j'$, and $\sum_{k=1}^{j'} t_k \leq n-1$, then

\[
\pi(t | s, a_s) = \prod_{k=1}^{j'} \left( \begin{array}{c} n_{s_{l_k}} \\ t_k \end{array} \right) q_{\tau_1, \tau_2},
\]
where \( \tau_1 = \sum_{k=1}^{j'} t_k \), \( \tau_2 = n - \tau_1 \). If \( t \) is the stopping state, then

\[
(5.7) \quad \tilde{\pi}(t|s, a_g) = q_{n,0}.
\]

Otherwise \( \tilde{\pi}(t|s, a_g) = 0 \). Finally, if \( s = (i_1^1) \), then

\[
(5.8) \quad \tilde{\omega}(s, a_g) = v_{i_1^1},
\]

where \( v_{i_1^1} \) is the mean of the \( r(p_{i_1^1}, x) \) with respect to the distribution \( F \). If \( s = (0) \), then \( \tilde{\omega}(s, a_g) = 0 \).

\textbf{Proof:} Suppose \( s \) is a state where at least two men are idle when the next job arrives, that is, \( s = (i_1^1, \ldots, i_j^j) \), where \( \sum_{k=1}^{j} r_k \geq 2 \).

Suppose we assign the job to a \( p_{i_{\sigma}} \) man (with probability \( u_{\sigma} \)). Then there are two possibilities before the arrival of the next job:

(a) No \( p_{i_{\sigma}} \)'s finish, \( s_k \) of the \( n_{i_k} - r_k \) busy \( p_{i_k} \)'s finish,

\( 1 \leq k \leq j, k \neq \sigma \), and for some \( 0 \leq j' \leq m-j, t_k \) of the \( n_{i_{k'}} \) busy \( p_{i_{k'}} \)'s finish, \( 1 \leq k \leq j' \). Thus, \( \tau_1^j \) of the \( \tau_1^\sigma + \tau_2^\sigma \) busy men finish, and the next state is

\[
t = (i_1^{r_1+s_1}, \ldots, i_{\sigma}^{r_{\sigma}-1}, \ldots, i_j^{r_j+s_j}, t_1, \ldots, t_j').
\]

This proves (5.2).
(b) $s_\sigma + 1$ of the $n_i\sigma + 1 - r_\sigma$ busy $p_i\sigma$'s finish, $s_k$ of
the busy $p_{1k}$'s finish, $1 \leq k \leq j$, $k \neq \sigma$, and for some $0 \leq j' \leq m-j$,
$t_k$ of the $n_k\ell_k$ busy $p_{\ell_k}$'s finish, $1 \leq k \leq j'$. This means $\tau_1$ of
the $\tau_1 + \tau_2$ busy men finish, and the next state is

$$t = (i_1^{1+s_1}, \ldots, i_j^{1+s_j}, t_1, \ldots, t_j)$$

Equation (5.3) follows.

The only other possibility is that the stopping state is
reached. This happens if and only if all $n - \sum_{k=1}^{m} r_k + 1$ busy men
finish before the next job arrives, and this event has probability
$q_{1,0}$, where $\tau = n - \sum_{k=1}^{m} r_k + 1$.

Now let $X$ be the value of the present job, a random variable.
Since the reward from this job is the only reward received before the
next job arrives, we condition on $X$ to obtain

$$\tilde{w}(s, a_s) = \sum_{k=1}^{b_k} \int_{b_k-1}^{b_k} r(p_{1k}, x) F(dx) = \sum_{k=1}^{v_k} v_s$$

proving (5.5).

The rest of the lemma follows easily by noting that if
$s = (i_1)$ or $(0)$, there is only one action possible: assign man $p_{i1}$
if $s = (i_1)$, and do nothing if $s = (0)$. In either case, all $n$
men are busy once the action has been taken, and the formulas for the
transition probabilities and expected rewards follow.
There are two important things to notice about the above lemma. First, the assumption that times to do jobs are exponential is crucial, since we need to know that each time a job arrives, all the busy men essentially start their jobs all over again. However, the interarrival times need not be exponential; the embedded chain is Markov regardless of the form of $G$ because we observe at times just after jobs arrive.

Secondly, if the stopping state is not included as a state, then all row sums of $\tilde{\pi}$ are strictly less than 1, no matter what actions are taken. In fact, if $s$ is a state with $1 \leq k \leq n-1$ idle men, the associated row sum is $1 - q_{n+1-k,0}$; if $s = (0)$, the associated row sum is $1 - q_{n,0}$.

Now suppose we use a stationary policy $\bar{\pi}$, which means each time we are in state $s = (i_1, \ldots, i_j)$, we use the same decision rule $a_s = (b_{1s}, \ldots, b_{js-1})$. Then we need $\tilde{V}_s(\bar{\pi})$, the expected value of the reward received by the end of the busy period, starting in state $s$. Now $\tilde{V}_s(\bar{\pi})$ satisfies

\begin{equation}
\tilde{V}_s(\bar{\pi}) = \tilde{w}(s, a_s) + \sum_{t \in \text{states}} \tilde{\pi}(t|s, a_s) \tilde{V}_t(\bar{\pi})
\end{equation}

Since the state space is finite and $\tilde{\pi}$ is strictly substochastic for all decision rules, (5.9) implies, in matrix notation, that

\begin{equation}
\tilde{V}(\bar{\pi}) = \tilde{w}(\bar{\pi}) + \sum_{k=1}^{\infty} \tilde{\pi}_k(\bar{\pi}) \tilde{w}(\bar{\pi}) = (I - \tilde{\pi}(\bar{\pi}))^{-1} \tilde{w}(\bar{\pi})
\end{equation}
where $\tilde{V}$ and $\tilde{w}$ are, respectively, the vectors of total rewards and one-period expected rewards, and $\tilde{\pi}$ is the matrix of transition probabilities. Note that $\tilde{V}(f)$ is well-defined for each $f$. This follows since the state space is finite, $\tilde{w}(s, a_s)$ is finite, and $(I - \tilde{\pi}(f))^{-1}$ exists and is finite.

We now show that the total expected reward from any policy is finite and in fact that there is a stationary policy which is optimal. First define the operator

$$Ty = \sup_{f}(\tilde{w}(f) + \tilde{\pi}(f)y),$$

where $y$ is a vector with as many components as there are states in the state space, and the supremum is taken over all decision rules $f$. This supremum, then, is really done component by component, so that

$$(Ty)_s = \sup_{a_s}(\tilde{w}(s, a_s) + \sum_{t} \tilde{\pi}(t|s, a_s) y_t).$$

For the following theorem we will need to assume that for each $s$, the above supremum is actually a maximum, that is, it is attained by some $a_s$. This will be the case, for instance, if $F$ puts its mass at only finitely many points. It will also be the case if $\tilde{w}(s, a_s)$ is continuous in the decision variables $(u_{s}^{1}, \ldots, u_{s}^{j})$ (where $s$ is state of the form $(i_{r}^{1}, \ldots, i_{r}^{j})$). Then since $\tilde{\pi}(t|s, a_s)$ is clearly continuous in the variables $(u_{s}^{1}, \ldots, u_{s}^{j})$, the function

$$\tilde{w}(s, a_s) + \sum_{t} \tilde{\pi}(t|s, a_s) y_t$$
will be continuous over the compact set

\[ 0 \leq u_s^1, \ldots, u_s^j \leq 1, \quad \sum_{k=1}^{i_s} u_s^k = 1, \]

and hence will achieve its maximum. In turn, \( \tilde{w}(s, a_s) \) will be continuous in \( (u_s^1, \ldots, u_s^j) \) if \( F(x) \) is continuous and \( r(p, x) \) is continuous in \( x \) for each \( p \).

**Theorem 5.1:** There exists a unique fixed point \( V^* \) of the operator \( T \).

Furthermore, if for each \( s \), there exists \( a_s^* \) with

\[
(5.11) \quad (TV^*_s) = \tilde{w}(s, a_s^*) + \sum_t \pi(t|s, a_s^*) V^*_t,
\]

then the stationary policy \( f^* \) which takes action \( a_s^* \) in state \( s \) is optimal over all policies, and the total expected reward, \( \bar{V}(f^*_s) \), from this policy is \( V^* \).

**Proof:** Since the row sums of \( \tilde{\pi}(f) \) are uniformly bounded above by \( 1 - q_n \), it follows that \( T \) is a contraction mapping with modulus \( \alpha \leq 1 - q_n \). Therefore, by the Banach fixed point theorem, \( T \) has exactly one fixed point \( V^* \), and for any \( y, T^k y \to V^* \) as \( k \to \infty \).

Now let \( f^*_s \) be the policy which, in state \( s \), takes action \( a_s^* \), where \( a_s^* \) satisfies \( (5.11) \). Then

\[ V^* = TV^*_s = \tilde{w}(f^*_s) + \tilde{\pi}(f^*_s) V^* , \]

or

\[ V^* = TV^*_s = \tilde{w}(f^*_s) + \tilde{\pi}(f^*_s) V^* , \]
\[ V^* = (I - \tilde{\pi}(f_\ast))^{-1} \tilde{w}(f_\ast) = \tilde{v}(f_\ast). \]

That is, the total expected reward from \( f_\ast^\infty \) is \( V^* \). Furthermore, let \( g = (f_1, f_2, \ldots) \) be any other policy. That is, \( g \) uses decision rule \( f_1 \), then \( f_2 \), and so on. Then we have

\[ V^* = TV^* = \tilde{w}(f_\ast) + \tilde{\pi}(f_\ast) V^* \]

\[ \geq \tilde{w}(f_1) + \tilde{\pi}(f_1) V^* \]

\[ \geq \tilde{w}(f_1) + \tilde{\pi}(f_1) (\tilde{w}(f_2) + \tilde{\pi}(f_2) V^*) \]

\[ \vdots \]

\[ \geq \tilde{w}(f_1) + \tilde{\pi}(f_1) \tilde{w}(f_2) + \tilde{\pi}(f_1) \tilde{\pi}(f_2) \tilde{w}(f_3) + \cdots + \tilde{\pi}(f_k) \cdots \tilde{\pi}(f_1) V^* , \]

and this last expression, as \( k \to \infty \), goes to the expected total reward from \( g \). Thus \( f_\ast^\infty \) is optimal.

Before discussing methods of finding \( V^* \) and the optimal \( f_\ast \), we see how the problem may be somewhat simplified. Note that in any state of the form \((0)\) or \((i_1^r)\), where \( 1 \leq r_1 \leq n_1 \), there is only one decision possible. Therefore, we may "remove" any or all of these states from the state space. Removal of any subset \( C \) of these states amounts to observing the embedded Markov chain only at times right after a job arrives and the state is not a member of \( C \). However,
when the set $C$ is removed, the transition probabilities and expected rewards associated with the remaining states must be modified to $\pi$'s and $\tilde{w}$'s.

To this purpose, define the following. Let $\tilde{\pi}_C$ be the square submatrix of $\tilde{\pi}$ corresponding to $C$; let $\tilde{w}_C$ be the expected reward (column) subvector of $\tilde{w}$ corresponding to $C$; let $\tilde{\pi}_{s,C}$ be the (row) subvector of transition probabilities from a state $s \notin C$ into states in $C$; finally, let $\tilde{\pi}_{C,t}$ be the (column) subvector of transition probabilities from the states of $C$ to a state $t \notin C$. Then it is easily seen that the transition probabilities $\pi(t|s, a_s)$ and expected rewards $\tilde{w}(s, a_s)$ for the new embedded Markov chain are, for $s, t \notin C$,

\[(5.12) \quad \pi(t|s, a_s) = \tilde{\pi}(t|s, a_s) + \sum_{k=0}^{\infty} \tilde{\pi}_{s,C} \tilde{\pi}_C^{-k} \tilde{\pi}_{C,t} = \tilde{\pi}(t|s, a_s) + \tilde{\pi}_{s,C}(I - \tilde{\pi}_C)^{-1} \tilde{\pi}_{C,t},\]

and

\[(5.13) \quad \tilde{w}(s, a_s) = \tilde{\tilde{w}}(s, a_s) + \sum_{k=0}^{\infty} \tilde{\pi}_{s,C} \tilde{\pi}_C^{-k} \tilde{w}_C = \tilde{\tilde{w}}(s, a_s) + \tilde{\pi}_{s,C}(I - \tilde{\pi}_C)^{-1} \tilde{w}_C .\]

One special set $C$ is $\tilde{C} = \{(0); (1^1), \ldots, (m^1)\}$. There is only one possible decision in each state of $\tilde{C}$, even when $m = n$, i.e., even when all of the $\tilde{p}$'s are distinct. Thus, $\tilde{C}$ may always be removed from the state space. The resulting transition probabilities and expected rewards are particularly simple, as shown below.
Proposition 5.1: Suppose the set \( \mathcal{C} \) is removed from the state space, and \( \pi(t \mid s, a_s) \) and \( w(s, a_s) \) are the resulting transition probabilities and expected rewards. Let \( t = (t_1^1, \ldots, t_j^j) \) where \( 2 \leq \sum_{k=1}^{i} t_k \leq n-1 \). Then if \( s \) is a state with \( 3 \leq k \leq n-1 \) idle men, we have

\[
\pi(t \mid s, a_s) = \pi(t \mid s, a_s)
\]

(5.14) and

\[
w(s, a_s) = \tilde{w}(s, a_s)
\]

If \( s = (i_1^1, i_2^1) \), then letting \( \tau = \sum_{k=1}^{i} t_j \), we have

\[
\pi(t \mid s, a_s) = \pi(t \mid s, a_s) + \prod_{k=1}^{j} \left( \begin{array}{c} n_k \\ t_k \end{array} \right) \frac{q_{0,n-1} q_{\tau,n-\tau}}{(1-q_{0,n} q_{1,n-1})}
\]

(5.15) and

\[
w(s, a_s) = \tilde{w}(s, a_s) + q_{0,n-1} u_s^2 v_{i_1} + u_s^1 v_{i_2}
\]

\[
+ \frac{q_{0,n-1} q_{1,n-1}}{(1-q_{0,n} q_{1,n-1})} \sum_{k=1}^{m} n_k v_k,
\]

where as before,

\[
v_k = \int_{-\infty}^{\infty} r(p_k, x) F(dx)
\]

Finally, if \( s = (i_1^2) \), then
\[ \pi(t|s, a_s) = \pi(t|s, a_s) + \prod_{k=1}^{j} \left( \frac{\sum_{l_k} q_{0,n-l} q_{r,n-r}}{1-q_{0,n-\sum_{l_k} q_{r,n-r}}} \right) \]

(5.16) and

\[ w(s, a_s) = \tilde{w}(s, a_s) + q_{0,n-l} \nu_1 + \frac{q_{0,n-l} q_{1,n-1}}{1-q_{0,n-\sum_{l_k} q_{r,n-r}}} \sum_{k=1}^{m} n_k \nu_k. \]

**Proof:** (5.14) follows from (5.12) and (5.13), because if \( s \) is a state with \( k \geq 3 \) idle men, then \( \tilde{\pi}_{s,i} = 0. \)

To prove (5.15) and (5.16) we need \((I - \tilde{\pi}_{C})^{-1}\). Since \( I - \tilde{\pi}_{C} \) has the special structure

\[
I - \tilde{\pi}_{C} = \begin{bmatrix}
1 - q_{0,n} & -n_1 q_{1,n-1} & \cdots & -n_m q_{1,n-1} \\
-q_{0,n} & 1 - n_1 q_{1,n-1} & \cdots & -n_m q_{1,n-1} \\
\vdots & \vdots & \ddots & \vdots \\
-q_{0,n} & -n_1 q_{1,n-1} & \cdots & 1 - n_m q_{1,n-1}
\end{bmatrix}
\]

it is easy to show that all row sums of \((I - \tilde{\pi}_{C})^{-1}\) equal \(1/(1-q_{0,n} - \sum_{l_k} q_{r,n-r})\). In fact,

\[
(I - \tilde{\pi}_{C})^{-1} = \begin{bmatrix}
1 + \frac{a}{1-anb} & \frac{n_1 b}{1-anb} & \cdots & \frac{n_m b}{1-anb} \\
\frac{a}{1-a nb} & 1 + \frac{n_1 b}{1-a nb} & \cdots & \frac{n_m b}{1-a nb} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{a}{1-a nb} & \frac{n_1 b}{1-a nb} & \cdots & 1 + \frac{n_m b}{1-a nb}
\end{bmatrix}
\]

151
where \( a = q_{0,n} \), \( b = q_{1,n-1} \).

Also, if \( s = (i_1^1, i_2^1) \), we have

\[
\tilde{\pi}_{s, c} = [0, 0, \ldots, 0, u_s^2 q_{0,n-1}, 0, 0, \ldots, 0, u_s^1 q_{0,n-1}, 0, \ldots, 0];
\]

component \( i_1 \)

component \( i_2 \)

if \( s = (i_1^2) \), we have

\[
\tilde{\pi}_{s, c} = [0, 0, \ldots, 0, q_{0,n-1}, 0, \ldots, 0].
\]

component \( i_1 \)

Finally,

\[
\tilde{\pi}_{c, t} = \begin{bmatrix}
\prod_{k=1}^{j} \left( \begin{array}{c}
\frac{n_x}{k} \\
\frac{t_x}{k}
\end{array} \right) q_{\tau,n-\tau} \\
\vdots \\
\prod_{k=1}^{j} \left( \begin{array}{c}
\frac{n_x}{k} \\
\frac{t_x}{k}
\end{array} \right) q_{\tau,n-\tau}
\end{bmatrix}
\]

and

\[
\tilde{w}_c = \begin{bmatrix}
0 \\
v_1 \\
\vdots \\
v_m
\end{bmatrix}
\]

From these, (5.15) and (5.16) follow from (5.12) and (5.13) by direct matrix multiplication.
When \( m < n \), we will want to remove more states than \( \tilde{C} \), namely, all states of the form \((0)\) and \( (i_1^{r_1}, \ldots, i_j^{r_j}) \), \( 1 \leq r_1 \leq n_i \). This can be done using (5.12) and (5.13). Note, of course, that the calculations needed to do this have to be done only once since they do not depend on the policy being used.

With this modified state space, we can now state that when the process starts in state \( s = (i_1^{r_1}, \ldots, i_j^{r_j}) \), \( 2 \leq j \leq m \), \( 2 \leq \sum_{k=1}^{j} r_k \leq n-1 \), and a policy \( f^{\infty} \) is used, we receive an expected total reward

\[
V_s(f) = \bar{V}_s(f) = (I - \pi(f))^{-1}_s W(f)
\]

until the end of the busy period. Of course, the same \( f^{\infty}_\pi \) which was optimal in Theorem 5.1 still maximizes the total expected reward vector \( V(g) \) over all policies \( g = (f_1, f_2, \ldots) \).

We now turn to the first stage of the decision process, that is, the beginning of the busy period. Once we have found the optimal decision in each state of the second stage of the busy period, it is a relatively simple matter to find the optimal decision at the beginning of the busy period. If a \( p_i \) man is assigned initially, then with probability \( q_{1,0} \), the busy period will end before the next job arrives, and with probability \( q_{0,1} \), the state of the system will be

\[
s_1 = (1^{n_1}, \ldots, (i-1)^{n_i-1}, i^{n_i-1}, (i+1)^{n_{i+1}}, \ldots, m^m)
\]
when the next job arrives. That is, the second stage of the busy period will have started, with \( n-1 \) idle men.

Therefore, if the job which begins a busy period has value \( x \), we must find

\[
\max_{1 \leq i \leq m} \left[ r(p_i, x) + q_0, l V_{s_i} (f_\pi^*) \right],
\]

where \( V_{s_i} (f_\pi^*) \) is the total expected reward starting in state \( s_i \) and using the optimal policy \( f_\pi^* \). Again from Lemma 2.3, there exist numbers \( -\infty = b_0 \leq b_1 \leq \cdots \leq b_{m-1} \leq b_m = +\infty \), such that it is best to assign a \( p_i \) man if and only if \( b_{i-1} < x \leq b_i \), and these \( b \)'s can be found using that lemma.

We now turn our attention to finding \( V^* \) and \( f_\pi^* \). As usual, the two main techniques are successive approximations and policy improvement. For the first method, we use the fact that

\[
T^k y \to V^* \quad \text{as} \quad k \to \infty,
\]

for each \( y \). The modulus of \( T \) depends on which of the original states have been removed from the state space. For instance, if none of the states have been removed, an upper bound for the modulus is \( 1 - q_{n, 0} \). However, no matter which of the states have been removed, the modulus is still some number \( 0 < \alpha < 1 \). From this fact we have the inequality
\[ \|\pi^k_y - v^*\| \leq \frac{\alpha^k}{1 - \alpha} \|\pi y - y\| , \]

where \( \|y\| = \max_s |y_s| \). This shows us how close we are to the actual \( V^* \) at the kth iteration.

Note that it may be difficult to find the \( a_s^* \) which maximizes

\[ w(s, a_s) + \sum_t \pi(t|s, a_s) y_t \]

over all \( a_s \). This may be practical only if there are only finitely many actions possible, i.e., only if \( V \) puts its mass at finitely many points.

The second method, policy improvement, may be done by linear programming [2], when there are again only finitely many actions in each state. We then solve

\[
\begin{align*}
\text{maximize} & \quad \sum_f x^T(f) w(f) \\
\text{subject to} & \quad \sum_f [I - \pi(f)]^T x(f) = c, \\
& \quad x(f) \geq 0.
\end{align*}
\]

(5.17)

Here \( c \) is any \( S \times 1 \) vector whose components are all strictly positive and sum to 1, the summations are over all decision rules \( f \), and \( S \) is the number of states, after any removals.

The notation in (5.17) may be somewhat misleading. Actually there is one column for each action in each state, so that the number
of distinct columns corresponds to the number of actions in a given state, summed over all the states. This number may be very large, but the columns do not all need to be listed in the memory of a computer. Instead, we make use of the following fact. If we consider state \( s = (i_1, \ldots, i_j) \), then all possible \( s \)-rows of \( \pi(f) \), which generate all possible \( s \)-columns of \( (I - \pi(f))^T \), are of the form

\[
\begin{bmatrix}
a_{01} + \sum_{k=1}^{i} a_{k1} u_s^k, & \ldots, & a_{0S} + \sum_{k=1}^{i} a_{kS} u_s^k
\end{bmatrix},
\]

where the coefficients, the \( a_{kl} \)'s, are fixed, and only the \( u_s \)'s vary from policy to policy. This follows from Lemma 5.3. The same thing is true for the possible \( s \)-components of \( w(f) \), except that there are also fixed coefficients of the \( v_s^k \)'s. Thus, once these fixed coefficients have been calculated, new columns of \( (I - \pi(f))^T \) and new components of \( w(f) \) may be generated easily by varying the \( u_s \)'s and \( v_s \)'s.

Probably the best way to solve the above linear programming problem is by the (primal) simplex method. If this is done, it can be shown [2] that all basic solutions are of the form

\[
x(\tilde{\gamma}) = [I - \pi(\tilde{\gamma})]^T c
\]

for some \( \tilde{\gamma} \),

\[
x(f) = 0
\]

for some \( f \neq \tilde{\gamma} \).

Since we can write \( [I - \pi(\tilde{\gamma})]^T c = [c^T \sum_{k=0}^{\infty} \pi^k(\tilde{\gamma})]^T \), the \( s \)-component of the basic solution can be interpreted as the expected number of times...

156
the process is in state \( s \), given the initial distribution is specified by the vector \( c \).

When the usual optimality criterion of the simplex method is achieved, the final basis corresponds to the optimal policy. Furthermore, the corresponding optimal solution to the dual of (5.18) is the optimal total expected reward vector \( V^* \).

**Special cases when** \( r(p,x) = px \).

When \( r(p,x) = px \) we would like the optimal policy to be independent of the \( p \)'s, as it was in other problems we have considered. Unfortunately, it is not in general, and we use Lemma 2.4 to show why.

Suppose policy \( f_{j}^{\infty} \) is used. Then the expected reward until the end of the busy period, starting from state \( s \), is of the form

\[
V_s(f_j) = \sum_{k=1}^{m} p_k x_k^j, 
\]

since every reward given is of the form \( p_i x \). Each policy \( f_{j}^{\infty} \) gives a different vector \( x_s^j = (x_1^j, s', \ldots, x_m^j, s) \), but for each \( j \) we have

\[
\sum_{k=1}^{m} x_k^j, s = c_s, 
\]

where \( c_s \) is a constant independent of \( j \). This follows from the fact that if all the \( p \)'s were equal to 1, then the same expected reward \( c_s \) would be obtained no matter what policy was used. Thus,
by Lemma 2.4, the only way the optimal policy can be independent of the $p$'s is that for each $s$, there must be some $f_j^\infty$ such that $x_j^s$ fulfills the role of $x$ in Lemma 2.4. That is,

\begin{equation}
\sum_{i=k}^{m} (x_j^i, s - x_j^{i'}, s) \geq 0, \quad 2 \leq k \leq m, \quad \text{for all } j \neq j'.
\end{equation}

Note that this is always the case when $m = 2$ and the conditions are such that an optimal stationary policy $f_\pi^\infty$ exists (see Theorem 5.1). This means, for instance, that if there are $n$ men, each of which has value either $p_1$ or $p_2$, and the job distribution $F$ puts its mass at only finitely many points, then the optimal policy is independent of $p_1$ and $p_2$.

However, several computer runs which found $x_1^j$ for each policy when $F$ was discrete, $G$ was exponential, and $m = n = 3$, showed that no $x_1^j$ satisfied (5.18), and thus that the optimal policy was not independent of the $p$'s. Therefore, it seems that if $m > 2$, the optimal policy will, in general, depend upon the $p$'s.

One very special case is when $m = n = 2$ and $r(p, x) = px$. Since $m = 2$, we know from the above that if an optimal policy exists, it is independent of $p_1$ and $p_2$. However, more can be said. Since $n = 2$, there are no states in the second stage of the busy period where more than one action is possible. The only time when a real choice must be made is at the beginning of a busy period. Then if we assign man $p_1$ initially, the busy period ends with probability $q_{1,0}$ before the next job arrives, or, with probability $q_{0,1}$, man $p_1$ is
still busy when the next job arrives and man \( p_2 \) must be assigned. In this latter case we receive an expected reward 
\( p_2 \nu = p_2 \int_{-\infty}^{\infty} xF(dx) \),
plus an expected reward \( V_0 \) from then until the end of the busy period.
Similar statements hold if we initially assign man \( p_2 \). Thus if the first job has value \( x \), we must find

\[
\max\{p_1 x + q_{0,1} (p_2 \nu + V_0), p_2 x + q_{0,1} (p_1 \nu + V_0)\}.
\]

From this it follows that the optimal policy is to assign man \( p_1 \) if and only if \( x \leq q_{0,1} \nu \).

Note that the above reasoning holds (1) even if the service times are not exponential, although the 2 men's service times are distributed the same, and (2) if a queue is allowed to form, that is, jobs which arrive when both men are busy are not turned away.

**A queue of \( N \) allowed to form.**

Suppose we allow a waiting line of jobs, of length at most \( N < \infty \), to form when all men are busy. The jobs are serviced on a FCFS basis as men become available to do them. This means that no decisions are necessary except when at least 2 men, with different \( p \)'s, are idle.

This problem can be handled as before, except that new states, new transition probabilities, and new one-period expected rewards must be added to account for times when all men are busy. For the time being, let the process be observed at all times just after the successive jobs arrive. As before, the state of the system can be of the form
\((i^1_1, \ldots, i^j_j), j \geq 1 \text{ and } 1 \leq r_k \leq n_i^j_k \text{ for } 1 \leq k \leq j, \text{ but now it can also be of the form (0,k), } 1 \leq k \leq N, \text{ which indicates that 0 men are idle and } k \text{ jobs, including the one which just arrived, are in the waiting line.}

Let \(C = \{(0,1), \ldots, (0,N)\} \) and let \(C'\) be the rest of the states. Now the only possible one-step transitions from \(C'\) into \(C\) are from \((i^1_1) \rightarrow (0,1)\) for some \(i^1_1\), and each of these has probability \(q_{0,n}^i\). For other new transition probabilities we need the following notation.

Let \(\{Z^i_k\}\) be i.i.d. service times, that is, each \(Z\) is exponential with mean \(1/\mu\). Let \(Y\) be the time between arrivals of two successive jobs. Also let \(U^i_j = \min_{1 \leq k \leq j} z^i_k\). Then define

\[
\alpha_{\ell,k} = P\left( \sum_{i=1}^{k+1} U^i_n + \sum_{i=2}^{\ell} U^i_{n-(i-1)} < Y < \sum_{i=1}^{k+1} U^i_n + \sum_{i=2}^{\ell+1} U^i_{n-(i-1)} \right)
\]

for \(0 < \ell \leq n, k \geq 1\), and

\[
\beta_k = P\left( \sum_{i=1}^{k} U^i_n < Y < \sum_{i=1}^{k+1} U^i_n \right), \quad k \geq 0 .
\]

Now consider a transition from \(s = (0, k_1)\) to \(t = (0, k_2)\) where \(1 \leq k_1 < N\). First of all, we must have \(k_2 \leq k_1 + 1\), that is, the number which will be waiting can not be more than the number waiting now plus the one which will arrive. Then it is easy to see that if \(k_2 \leq k_1 + 1\), we have \(\pi(t|s, a_s) = \beta_{k_1-k_2+1}\). The reasoning is that exactly \(k_1 - k_2 + 1\) of the \(n + k_1\) jobs being serviced or
waiting in line must be finished when the next job arrives. If \( k_1 = N \),

\[
\pi(t|s, a_s) = \begin{cases} 
\beta_{N-k_2+1} & 1 \leq k_2 < N \\
\beta_0 + \beta_1 & k_2 = N
\end{cases}
\]

Finally consider a transition from \( s = (0, k) \) to \( t = (i_1, \ldots, i_j) \), where \( 1 \leq j \leq m \) and \( \ell = \sum_{i=1}^{j} r_i > 0 \). Then we have

\[
\pi(t|s, a_s) = \prod_{k=1}^{j} \binom{n_i}{k} \alpha_{\ell, k},
\]

since exactly \( \ell+k \) of the \( n+k \) jobs being serviced or waiting in line must be finished by the time the next job arrives.

Of course, all one-step transition probabilities between states in \( C' \) are the same as when no queue was allowed.

Next we need expected rewards when in \( C \). Suppose a job arrives and there are no idle men. Then assuming this job may wait in line, that is, the queue is not too long, it will eventually be serviced by one of the \( n \) men. It is equally likely who will end up doing this job, so the expected reward from the job is

\[
\tilde{w}(s, a_s) = \frac{1}{n} \sum_{i=1}^{m} n_i v_i,
\]

where \( s \) is any of the states \((0, k)\), \( 1 \leq k \leq N \). (The actual reward from this job will not be realized until the job is actually started, 161
but its expected value is known to be the above expression as soon as it arrives and sees that it must wait in line.) The rest of the rewards are as before.

Again, the states $\mathcal{C}$, or even the states $\bar{\mathcal{C}} = \{(0,1), \ldots, (0,N); (1), \ldots, (m)\}$ may be removed as before, using (5.12) and (5.13), since there is only one action possible in each of these states.

The only problem left, then, is how to calculate the $\alpha$'s and $\beta$'s. We look at the $\beta$'s first. Since $U^i_j$ is a minimum of $j$ exponentials with rate $\mu$, it is exponential with rate $j\mu$. Thus

$$\sum_{i=1}^{k} U^i_n$$

is gamma with density

$$f_{k,n}(x) = \frac{x^{k-1}(n\mu)^k}{(k-1)!} e^{-n\mu x} , \quad x > 0 .$$

This means that

$$\beta_k = \int_0^\infty \int_0^\infty (G(x+y) - G(x)) \frac{x^{k-1}(n\mu)^k}{(k-1)!} e^{-n\mu x} (n\mu) e^{-n\mu y} dx \, dy ,$$

by conditioning on $\sum_{i=1}^{k} U^i_n$ and $U^{k+1}_n$. For example, if $G$ is exponential with mean $1/\lambda$, then $\beta_k$ becomes

$$\beta_k = \left( \frac{\rho}{1+\rho} \right)^k - \left( \frac{\rho}{1+\rho} \right)^{k+1} = \frac{\rho^k}{(1+\rho)^{k+1}}$$

where $\rho = \lambda / n\mu$.

There are no such simple formulas for $\alpha_{j,k}$ because

$$\sum_{i=2}^{k} U^i_n(n^{-i-1})$$

is the sum of exponentials with unequal parameters and
hence has no simple form for its density. Given, however, that through
tedious calculations we can find the density $f(x)$ of $\sum_{i=1}^{k-1} U_i^n + \sum_{i=2}^{n} U_{n-(i-1)}$, then $\alpha_{\ell,k}$ is given by

$$\alpha_{\ell,k} = \int_{0}^{\infty} \int_{0}^{\infty} (G(x+y) - G(x)) f(x) (n-\ell) \mu \ e^{-(n-\ell)\mu y} \ dx \ dy.$$ 

As a final note, suppose the system is in light traffic,
that is, $EY > 1/\mu$. Then a large value of $N$ may be considered
a good approximation to the case where $N = \infty$, i.e., the case where a
queue of arbitrary length is allowed to form.

**Cost of being idle imposed.**

In the above model we have not imposed an explicit cost for
keeping our best men idle, which we may want to do. Suppose that a
cost of $c(p_i)$, where $c$ is increasing in $p_i$, is incurred per unit
time that man $p_i$ is idle. Fortunately this new aspect of the model
causes no new difficulties. Lemma 5.1 still applies, so that the
same type of critical number policies considered there are the only
type we need to consider now. Furthermore, the same embedded Markov
chain structure still applies with the same transition probabilities.

The only thing we need to change are the one-period expected
rewards. If we let $s = (i'_1, \ldots, i'_j)$ then the new expected
reward in this state is given by

163
\begin{align*}
\tilde{w}(s, a_s) &= \sum_{k=1}^{m} v^k_s - \frac{1}{\lambda} \sum_{k=1}^{m} u^k_s (r_k - 1) c(p_{ik}) + \sum_{\ell \neq k} \frac{1}{\lambda'} \left( \sum_{k=1}^{m-j} (n_{ik} - r_k) c(p_{ik}) + \sum_{k=1}^{m-j} n_{ik} c(p_{ik}) \right) \\
&\quad - \frac{1}{\lambda'} \left( \sum_{k=1}^{m-j} (n_{ik} - r_k) c(p_{ik}) + \sum_{k=1}^{m-j} n_{ik} c(p_{ik}) \right)
\end{align*}

where \( \{\ell_1, \ldots, \ell_{m-j}\} = \{1, \ldots, m\} - \{i_1, \ldots, i_j\} \), \( \frac{1}{\lambda} \) is the expected time between jobs, and \( \frac{1}{\lambda'} \) is the expected amount of time a busy man will be idle before the next job arrives. This means

\[
\frac{1}{\lambda'} = \int_0^\infty \int_0^y (y-z) \mu e^{-\mu z} \, dz \, G(dy) .
\]

If a queue of length \( N \) is allowed to form and \( s = (0,k), 1 \leq k \leq N \), the new expected reward in this state is given by

\[
\tilde{w}(s, a_s) = \frac{1}{n} \sum_{i=1}^{n} n_i v_i - \frac{1}{\lambda'} \sum_{i=1}^{n} n_i c(p_i) .
\]

From here we may proceed as before with successive approximations or policy improvement to obtain the optimal policy.

**Examples when** \( r(p,x) = px \).

1. Let \( n = 5, m = 2, n_1 = 2, n_2 = 3, N = 0 \). Suppose \( G \) is exponential with mean 1, and \( \mu = 1 \). Also suppose the job values are distributed as \( X \), where
\[ X = \begin{cases} 
1 & \text{with pr. .5} \\
2 & \text{with pr. .25} \\
3 & \text{with pr. .25} 
\end{cases} \]

Since \( m = 2 \), the optimal policy is independent of \( p_1 \) and \( p_2 \), so we may as well take these to be 0 and 1, respectively.

We set up the problem and use policy improvement, beginning with the policy: always give to a \( p_2 \) man if there is one available. This, however, turns out to meet the optimality criterion, so that it is the optimal policy in the second stage of the busy period.

This is not too surprising for the following reason. Since the maximal service rate is 5/unit time and the arrival rate is 1/unit time, there will usually be idle men. Also since \( n_2 = 3 > n_1 = 2 \), there are enough \( p_2 \)'s that we may want to use them whenever possible, and indeed, this is the case.

For the first decision in the busy period, we need

\[
\max(q_{0,1} V_{s_1}(f_*), x + q_{0,1} V_{s_2}(f_*)),
\]

where \( s_1 = (1^1, 2^1), s_2 = (1^2, 2^2) \). Then since \( q_{0,1} = .5 \), \( V_{s_1}(f_*) = 5.67 \), and \( V_{s_2}(f_*) = 5.44 \), it turns out that it is also always optimal to assign a \( p_2 \) man initially.

2. Let \( n = 4, m = 2, n_1 = 2, n_2 = 2, p_1 = 0, p_2 = 1, N = 0 \). Suppose G is degenerate with unit mass at 1, and \( \mu = 1/3 \). Also, suppose the job values are distributed as \( X \), where
Again the policy improvement method is used. After eliminating the states $C = \{(0), (1^1), (2^1), (1^2), (2^2)\}$ the remaining states are $(1^1, 2^1), (1^1, 2^2), \text{ and } (1^2, 2^1)$. We start with the policy

$$f_1 : (u_{s_0}^1 = .5, u_{s_1}^1 = .5, u_{s_2}^1 = .5) ,$$

where $s_0, s_1, s_2$ are the states $(1^1, 2^1), (1^1, 2^2), \text{ and } (1^2, 2^1)$, respectively. This means that in state $s_1$, for instance, $f_1$ assigns man $p_1$ if and only if the next job has value 1 or 2, that is, with probability .5.

Starting with $f_1$ we need 4 pivots to arrive at the optimal policy. The succession is

$$f_2 : (u_{s_0}^1 = .5, u_{s_1}^1 = 0, u_{s_2}^1 = .5) ,$$

$$f_3 : (u_{s_0}^1 = .5, u_{s_1}^1 = 0, u_{s_2}^1 = .2) ,$$

$$f_4 : (u_{s_0}^1 = .2, u_{s_1}^1 = 0, u_{s_2}^1 = .2) ,$$

$$f_\# = f_5 : (u_{s_0}^1 = .2, u_{s_1}^1 = .2, u_{s_2}^1 = .2) .$$

166
The expected total rewards from $f_*$ are given by

$$(V_{s_0}(f_*), V_{s_1}(f_*), V_{s_2}(f_*)) = (17.13, 18.09, 16.63).$$

To get the optimal policy at the beginning of the busy period we need to find

$$\max(q_{0,1} V_{s_1}(f_*), x + q_{0,1} V_{s_2}(f_*)),$$

where $q_{0,1} = .719$. Using the above values for the $V(f_*)$'s, we find that

$$q_{0,1} V_{s_1}(f_*) \geq x + q_{0,1} V_{s_2}(f_*),$$

if and only if $x \leq 1.05$.

Thus, we see that every time we have a decision to make, it is best to assign a $p_1$ man if the job value is 1, and to assign a $p_2$ man if the job value is 2 or 3.

**Interpretations of Model 1.**

One interpretation of this model is when the "men" are repairmen (or even repair teams) and the "jobs" are machines, say airplanes, which come to be repaired. The values of the men, the $p_i$'s, are taken to be numbers between 0 and 1 which rate how close the repairmen are to being perfect. The values of the airplanes, the $x$'s, may have one of several interpretations.
(1) An airplane of value $x_i$, when repaired by a man of value $p$, has a lifetime (until the next repair is needed) with distribution $1 - p\tilde{F}_i(t)$ and mean $px_i$. Thus we attempt to maximize the mean lifetimes of the planes our repairmen service. Here we tend to assign good men (large $p$'s) to good planes (large $x$'s).

(2) An airplane of value $x_i$ means that its defect is rated $x_i$ on a 0 to 1 scale. The possible defects are rated $0 < x_L < \cdots < x_i < 1$, with $x_L$ being the most serious, $x_i$ the least serious. If a man of value $p$ works on a plane with defect $x$, he repairs it satisfactorily with probability $px$, which is taken as the reward. This means the sum of the rewards for the first k planes is the expected number repaired satisfactorily. Here we will tend to pair good men (large $p$'s) with less serious defects (large $x$'s). In order to get an optimal policy which does the opposite, that is, which tends to pair good men with more serious defects, we would have to change the reward structure.

In either of the above two interpretations, it is reasonable to impose a cost $c(p_i)$ per unit time for a $p_i$ man being idle, and also to allow a queue of $N$ to form. These additional aspects of the model may be handled as described earlier in this section.

5.2. Model 2: Random Replacements After Assignments.

In this model there are $n$ men on hand at the beginning of each period, where successive time periods have equal, deterministic lengths. Each man has one of $m$ possible values $p_1 < \cdots < p_m$. 

168
Jobs arrive sequentially, one per time period, and take on values $X_1, X_2, \ldots$, i.i.d. random variables with a common distribution $F$, and each job is assigned to one of the men on hand. If a $p_i$ man is assigned to an $x$ job, a reward $r(p_i, x)$ is received, where we again assume that

$$r(p_2, x_2) + r(p_1, x_1) \geq r(p_2, x_1) + r(p_1, x_2) \text{ for all } x_1 \leq x_2, \ p_1 \leq p_2.$$  

When a man is assigned, a replacement is then chosen randomly so that $n$ men will again be on hand at the beginning of the next period. The probability of choosing a $p_i$ man to be a replacement is $q_i$, $1 \leq i \leq m$, and these $q_i$'s remain constant, independent of any decisions made throughout the problem.

Our objective is to maximize the infinite horizon expected discounted reward. As in the previous model, we set the problem up as a Markov decision chain and show that an optimal stationary policy exists when certain restrictions are imposed. We also see that the optimal policy may be found by successive approximations or by policy improvement in the form of linear programming.

By exactly the same proof as in Lemma 5.1, we see that the optimal policy has the same structure as that of Model 1 above. That is, if men of values $p_{i1} < \cdots < p_{ij}$ are available and a job of value $x$ arrives, there exist numbers $-\infty = b^0 \leq b^1 \leq \cdots \leq b^{j-1} \leq b^j = +\infty$, such that it is optimal to assign a $p_{ij}$ man if and only if $b^{j-1} < x \leq b^j$. Thus we consider only policies of this form.
We now set up the problem as a Markov decision chain. As before, we label the states \( s = (i_1, \ldots, i_j), \, 1 \leq j \leq m, \sum_{k=1}^{j} r_k = n. \)

An action \( a_s \) in state \( s \) will be specified by the critical numbers \( -\infty = b_s^0 \leq b_s^1 \leq \cdots \leq b_s^{j-1} \leq b_s^j = +\infty \), and as before, we define the numbers

\[
\begin{align*}
u_s^k &= F(b_s^k) - F(b_s^{k-1}) = P(\text{a } p_{i_k} \text{ man is assigned}), \quad 1 \leq k \leq j, \\
\end{align*}
\]

and

\[
\begin{align*}
u_s^k &= \int_{b_s^{k-1}}^{b_s^k} r(p_{i_k}, x) F(dx), \quad \quad 1 \leq k \leq j.
\end{align*}
\]

We assume \( r \) and \( F \) are such that \( \nu_s^k \) is finite for all \( s, k \), and \( p_{i_k} \).

The following lemma gives the transition probabilities and one-period expected rewards.

**Lemma 5.4:** Suppose \( s = (r_1, \ldots, r_j), \, 1 \leq j \leq m. \) If

\[
t = (r_1, \ldots, r_{k-1}, r_k; i_{k+1}, \ldots, i_j), \quad \text{where } i_{k+1} \neq i_k \text{ but may be one of } i_1, \ldots, i_{k-1}, i_{k+1}, \ldots, i_j,
\]

then the transition probability from \( s \) to \( t \) is

\[
\pi(t|s, a_s) = q_{i_{k+1}} u_s^k; \tag{5.19}
\]

if \( t = s \), then
\[
\pi(t \mid s, a_s) = \sum_{k=1}^{i} q^k \nu^k_s;
\]

otherwise \( \pi(t \mid s, a_s) = 0. \)

Suppose \( s = (i_1^n) \). If \( t = (i_1^{n-1}, i_2^1), i_1 \neq i_2, \) then

\[
\pi(t \mid s, a_s) = q_{i_2};
\]

if \( t = s, \) then

\[
\pi(t \mid s, a_s) = q_{i_1};
\]

otherwise \( \pi(t \mid s, a_s) = 0. \)

If \( s = (i_1^1, \ldots, i_j^j), 1 \leq j \leq m, \) the expected one-period reward is given by

\[
w(s, a_s) = \sum_{k=1}^{i} \nu^k_s;
\]

if \( j = 1, \) this reduces to

\[
w(s, a_s) = \int_{-\infty}^{\infty} r(p_{i_1}, x) F(dx);
\]

Proof: If \( s = (i_1^1, \ldots, i_j^j), \) then one of the men \( p_{i_1}, \ldots, p_{i_j} \) must be assigned and a replacement is then obtained. Since a \( p_{i_k} \) man is assigned with probability \( u^k_s \) and a \( p_{i_\ell} \) man is the replacement with probability \( q_{i_\ell}, \) formulas (5.19) - (5.22) follow.
To obtain $w(s, a_s)$, we condition on the value of the job $X$ to obtain

$$w(s, a_s) = \sum_{k=1}^{b_s^k} \int_{b_s^{k-1}} r(p_{ik}, x) F(dx) = \sum_{k=1}^{b_s^k} v_s^k.$$  

Then (5.24) follows from (5.23) trivially.

We now show when an optimal stationary policy exists. Suppose the discount factor is $0 < \alpha < 1$, and define

$$(Ty)_s = \sup_{a_s} (w(s, a_s) + \alpha \sum_t \pi(t|s, a_s) y_t).$$

Here $Ty$ and $y$ are vectors with as many components as there are states in the state space, and $(Ty)_s$ is the component of $Ty$ corresponding to state $s$. Since $\alpha < 1$, $T_\alpha$ is a contraction mapping with modulus $\leq \alpha$. Hence we may again apply the Banach fixed point theorem to assert that $T$ has a unique fixed point $V^*$ and that

$$T^k y \to V^* \quad \text{as } k \to \infty, \quad \text{for all } y.$$  

We assume that $r$ and $F$ are such that for each $s$, there exists an $a^*_s$ with

$$(TV^*)_s = w(s, a^*_s) + \sum_t \pi(t|s, a^*_s) V^*_t.$$  

172
(The comments preceding Theorem 5.1 give sufficient conditions for this to hold.) Then the same proof as in Theorem 5.1 shows that the stationary policy \( f_\infty \), which uses action \( a^* \) in state \( s \), is optimal for the discounted problem. Furthermore, we may use successive approximations or policy improvement as in Model 1 to find \( V^* \) and \( f^* \). Note that if we start with \( y = 0 \), then successive approximations also finds the policy (non-stationary) which maximizes the N-period expected discounted cost, for each \( N \geq 1 \).

We now examine the ordering of the components of \( V^* \), assuming that \( r(p,x) \) is increasing in \( p \) for each \( x \). Let \( s \) and \( t \) be any states such that \( t \) has the same men available as \( s \), except for one man, who is better (has a higher \( p \)) than the corresponding man in \( s \). We then say that \( t \) is one man better than \( s \). Then it is easy to see that \( V^*_t > V^*_s \), for the following reason. Starting in state \( t \), we may use a policy which does exactly what we would do (optimally) if we started in state \( s \). Let \( V_t \) be the expected reward from this policy. Since it is non-optimal, \( V_t \leq V^*_t \). But since \( r(p,x) \) is increasing in \( p \) for each \( x \), \( V_t \geq V^*_t \). Hence \( V^*_t = V^*_s \).

Now suppose \( s_1, \ldots, s_N \) are such that \( s_{i+1} \) is one man better than \( s_i \), \( 1 \leq i \leq N-1 \). Then we say \( s_N \) is better than \( s_1 \), and we have \( V^*_s \leq V^*_s \), since \( V^*_s \leq V^*_s \) for each \( i \). This says what is intuitively clear, namely, if we start in a state which, in the above sense, has better men, we receive a higher expected reward.

Given any two states \( s_1 \) and \( s_N \), we now ask whether \( s_N \) is a better state than \( s_1 \). (Notice that a given pair of states may not even relate to each other under this "better than" ordering.)
To answer this question, it is convenient to denote each state by an
m-vector whose $k$th component is the number of $p_k$ men available in
that state. Then letting $s_1 = (r_1, \ldots, r_m)$ and $s_N = (r'_1, \ldots, r'_m)$,
we have the following criterion.

**Proposition 5.2:** $s_N$ is better than $s_1$ if and only if

$$\sum_{i=1}^{k} (r_i - r'_i) \geq 0, \quad \text{for } 1 \leq k \leq m-1.$$

**Proof:** Sufficiency is proved by considering the chain of states

$$s_1 = (r_1, \ldots, r_m), (r'_1, r_2 + (r_1 - r'_1), r_3, \ldots, r_m),$$

$$\ldots, (r'_1, \ldots, r'_k, r_{k+1} + \sum_{i=1}^{k} (r_i - r'_i), r_{k+2}, \ldots, r_m),$$

$$\ldots, (r'_1, \ldots, r'_{m-1}, r_m + \sum_{i=1}^{m-1} (r_i - r'_i)) = (r'_1, \ldots, r'_m) = s_N,$$

where the next-to-last equality follows because

$$r_m + \sum_{i=1}^{m-1} (r_i - r'_i) = r_m + (n - r_m) - (n - r'_m) = r'_m.$$

Since each state in this chain is obviously better than the one preceding
it, $s_N$ is better than $s_1$.

To show necessity, suppose

$$\sum_{i=1}^{k} (r_i - r'_i) < 0, \quad \text{some } 1 \leq k \leq m-1.$$
Now if \( s_1, s_2, \ldots, \) is any chain of states such that \( s_{i+1} \) is one man better than \( s_i \), \( i \geq 1 \), then

\[
\sum_{i=1}^{k} r_i^j \leq \sum_{i=1}^{k} r_i^i, \quad \text{for } j \geq 2,
\]

where \( s_j = (r_1^j, \ldots, r_m^j) \). Hence \( s_N \) cannot be reached by such a chain and therefore, by definition, \( s_N \) is not better than \( s_1 \).

This proposition gives us an easy way to check if one state is better than another. For example, if \( n = 20, m = 5 \), and

\[
s_1 = (7, 2, 1, 4, 6), \quad s_N = (3, 2, 3, 5, 7),
\]

then \( s_N \) is better than \( s_1 \), as can be seen from the chain

\[
(7,2,1,4,6); (3,6,1,4,6); (3,2,5,4,6); (3,2,3,6,6); \text{ and } (3,2,3,5,7).
\]

Hence we know that \( V^*_{s_1} \leq V^*_{s_N} \). However, if \( s_N \) is instead given by \( s_N = (3,4,3,5,5) \), then

\[
\sum_{i=1}^{5} (r_i^i - r_i^1) = 14 - 15 < 0,
\]

so that \( s_N \) is not better than \( s_1 \) (and \( s_1 \) is not better than \( s_N \)), and we do not know the relation between \( V^*_{s_1} \) and \( V^*_{s_N} \) unless we actually find these numbers.
Special cases when \( r(p,x) = px \).

As in Model 1, we may use Lemma 2.4 to give necessary and sufficient conditions for the optimal policy to be independent of the \( p \)'s. Several computer runs, however, again showed that these conditions were not met by several examples with \( m = 3 \). Therefore, if \( m \geq 3 \) the optimal policy will in general depend upon the particular values of the \( p \)'s. On the other hand, if \( m = 2 \) and \( F \) is such that an optimal stationary policy exists, then this optimal policy will as before be independent of \( p_1 \) and \( p_2 \).

As a special example, suppose \( F \) is continuous and \( m = n = 2 \). The states may be denoted by \( (2,0) \), \( (1,1) \), and \( (0,2) \), and only in state \( (1,1) \) must a decision be made. If we assign man \( p_1 \) with probability \( u \) when in state \( (1,1) \), the transition matrix is given by

\[
\pi(u) = \begin{bmatrix}
q & 1-q & 0 \\
(1-u)q & uq + (1-u)(1-q) & u(1-q) \\
0 & q & 1-q
\end{bmatrix},
\]

and the one-period expected reward vector is given by

\[
 w^T(u) = \left[ p_1 v, \quad p_1 \int_{-\infty}^{F^{-1}(u)} xF(dx) + p_2 \int_{F^{-1}(u)}^{\infty} xF(dx), \quad p_2 v \right],
\]

where \( v = \int_{-\infty}^{\infty} xF(dx) \). Thus the expected reward vector \( V(u) \) is given by

\[
 V(u) = [I - \mathcal{C}(u)]^{-1} w(u).
\]
By calculating the inverse of $I - \alpha \eta(u)$, we find that

$$V(1,1)(u) = \frac{1}{[(1-\alpha)^2 + \alpha(1-\alpha)(q(1-u)+(1-q)u)]} \left[ \alpha q(1-u) (1-\alpha(1-q)) p_1 v + (1-\alpha q)(1-\alpha(1-q))(p_1 \int_{-\infty}^{F^{-1}(u)} xF(dx) + p_2 \int_{F^{-1}(u)}^{\infty} xF(dx)) + \alpha(1-q) u(1-\alpha q) p_2 v \right].$$

By taking the derivative with respect to $u$ and setting it equal to 0, we see that $V(1,1)(u)$ is maximized by $u^*$, where $u^*$ satisfies

$$\alpha[(1-q)(\int_{-\infty}^{F^{-1}(u^*)} xF(dx) - u^*F^{-1}(u^*)) + q(\int_{F^{-1}(u^*)}^{\infty} xF(dx) + (1-u^*)F^{-1}(u^*))] = (1-\alpha) F^{-1}(u^*).$$

In particular, if $q = 1/2$, the above equation becomes

$$\frac{\alpha v}{2} = (1 - \frac{\alpha}{2}) F^{-1}(u^*),$$

or

$$F^{-1}(u^*) = \frac{\alpha v}{(2-\alpha)}.$$

Also, since we know an optimal policy maximizes each component of $V$, we know that this same $u^*$ maximizes $V(2,0)(u)$ and $V(0,2)(u)$. 

177
Interpretations of Model 2.

One interpretation of Model 2 is that the "men" are parts, say motors, and the "jobs" are machines which use these motors. We assume that \( n \) motors are always on hand, and that at the beginning of each time period, a machine arrives needing a motor. The values of the motors, the \( p \)'s, are numbers between 0 and 1 which indicate how good the motors are with respect to a perfect motor. For instance, a motor with \( p = .9 \) is in some sense .9 as good as a perfect motor.

The random replacement of motors (the \( p \)'s) occurs in the following way. When a machine arrives for service, we take the old motor out of it and put in one of the \( n \) available ones. We then repair the old one a random amount, so that after repair, its quality is a random variable \( p \) with distribution \( \{q_i\} \). After the repair has taken place, we observe this random variable and add the repaired motor to our stockpile. We assume the cost of repair is fixed, so that it may be neglected in the decision making, and that the time to repair is negligible.

The values of the machines, the \( x \)'s, may have one of several interpretations.

1. Suppose a machine of value \( x_1 \) means that the lifetime of this machine, using a perfect motor, has a distribution \( F_1(t) \), with mean \( x_1 \). If the motor is imperfect with value \( p \), this machine has a lifetime with distribution \( 1 - pF_1(t) \) and with mean \( px_1 \). Suppose we wish to maximize the total expected lifetime of the first \( N \) machines we service. Then we will try to maximize the total \( N \)-period undiscounted expected reward, with \( r(p,x) = px \). In this interpretation,
we will tend to pair good motors (large $p$'s) with good machines (large $x$'s).

(2) Suppose all machines have the same lifetime distribution $1 - p\bar{F}(t)$ when using a motor of value $p$. However, some are more important than others in the sense that their values, the $x$'s, are the lengths of time we wish them to perform properly. If a motor of value $p$ is assigned to a machine of value $x$, we would either expect to receive a reward $h(p\bar{F}(x))$, with $h$ increasing, or incur a cost $c(1 - p\bar{F}(x))$, with $c$ increasing. Note that if $h(y) = y$, then the sum of the rewards from the first $N$ machines is just the expected number which will perform properly for their allotted amounts of time.

If $r(p,x) = h(p\bar{F}(x))$, then assuming $F$ has a density $f$, we have

$$\frac{\partial^2 r(p,x)}{\partial p \partial x} = -f(x)[h''(p\bar{F}(x)) p\bar{F}(x) + h'(p\bar{F}(x))] .$$

Since we are interested in the sign of $\frac{\partial^2 r}{\partial p \partial x}$, we must examine the sign of

$$g(y) = y h''(y) + h'(y), \quad y > 0 .$$

Note that if $h$ is convex and increasing, then $g(y) \geq 0$ and hence $\frac{\partial^2 r}{\partial p \partial x} \leq 0$. This means we will tend to pair good motors (large $p$'s) with machines which do not have to perform as long (small $x$'s).

However, it is not necessary that $h$ be convex in order for $g$ to be positive. In fact, $h$ may even be concave, as in the case of
\[ h(x) = \ln(1+x) \quad \text{or} \quad h(x) = x^a, \quad 0 < a < 1. \] In these cases we will also have \[ \frac{\partial^2 r}{\partial p \partial x} \leq 0. \]

On the other hand, if \( h \) is a function like

\[ h(x) = \sin(\pi x/2), \quad 0 \leq x \leq 1, \]

then

\[ g(y) = \frac{\pi}{2} ( - \frac{\pi y}{2} \sin(\frac{\pi y}{2}) + \cos(\frac{\pi y}{2}) ) , \]

which is non-negative for \( x \leq a \) and non-positive for \( x \geq a \), where \( a \) is slightly larger than \( 1/2 \). In this case, if we are sure that all possible \( p \)'s and \( x \)'s are such that \( p\overline{F}(x) \geq a \), then \( \frac{\partial^2 r}{\partial p \partial x} \geq 0 \)

and we will tend to assign small \( p \)'s to small \( x \)'s and large \( p \)'s to large \( x \)'s.

The same type of remarks also apply when the cost criterion \[ c(1 - p\overline{F}(x)) \] is used. Again it depends on the function \( c \) what characteristics the optimal policy will have.

5.3. **Model 3: Issuing Batteries.**

In this model we have a storeroom of batteries which deteriorate with age. The batteries are not all alike, in the sense that either some are older than others or some deteriorate faster than others. Machines, say automobiles, arrive at random times and request these batteries, and we must find issuing policies which optimize a
given criterion. In particular, we seek conditions under which a LIFO policy (always issue newest battery) or a FIFO policy (always issue oldest battery) is optimal.

We will examine several variations of this model, but all of them have several characteristics in common. We assume a given battery of age $x$ has a value $p(x)$, where $0 \leq p(x) \leq 1$. As in Model 2, this value indicates how good the battery is with respect to a perfect battery. If an automobile receives this battery of age $x$, it will keep running at least $t$ units of time with probability $p(x) \bar{F}(t)$. If we take $p(0) = 1$, we may consider $\bar{F}(t)$ to be the probability an automobile will run at least $t$ units of time, starting with a new battery. We now look at several variations of Model 3.

**Variation 1:**

A storeroom has $n$ batteries, which all deteriorate according to the same function $p$, but which have different initial ages. At random points of time, according to a renewal process with interarrival distribution $G$, automobiles arrive and request batteries. If an automobile receives a battery of age $x$, it will keep running at least $t$ units of time with probability $p(x) \bar{F}(t)$ where $\bar{F}$ is some distribution function, the same for all automobiles. We attempt to find an issuing policy which maximizes the total expected lifetime the first $n$ automobiles receive from these $n$ batteries.
Theorem 5.2: If $p$ is convex, a LIFO policy is optimal; if $p$ is concave, a FIFO policy is optimal.

Proof: Suppose a battery of age $x$ is used for a certain car, and let $L(x)$ be the lifetime of the car with this battery. Then

$$E(L(x)) = \int_0^\infty p(x) \tilde{F}(t) dt = p(x) \mu,$$

where $\mu$ is the expected lifetime of a car using a new battery when $p(0) = 1$.

The proof of the theorem now proceeds by induction on $n$. Suppose $p$ is a convex function, and when $n-1$ batteries are left, a LIFO policy is optimal. Now suppose there are $n$ batteries of ages $t_1 \leq \cdots \leq t_n$ on hand, a car arrives, and the battery of age $t_i$ is issued. Then the expected reward from using an optimal policy thereafter (a LIFO policy) is

$$p(t_i) \mu + \int_0^\infty \int_0^\infty \cdots \int_0^\infty p(t_1 + x_1 + \cdots + x_{n-1}) G(dx_{n-1}) \cdots G(dx_1) \mu.$$

We show that this is less than or equal to the expected reward from issuing the battery of $t_{i-1}$ first and then using a LIFO policy. We have
\[
\left\{ p(t_i) \mu + \left( \int_0^\infty p(t_{i-1} + x_1) \right) + \cdots + \int_0^\infty p(t_i - x_1 - \cdots - x_{i-1}) + \int_0^\infty p(t_{i+1} + x_1 + \cdots + x_i) + \cdots + \int_0^\infty p(t_n + x_1 + \cdots + x_{n-1}) \right. \\
\left. G(dx_{n-1}) \cdots G(dx_1) \right. \mu \right\} \\
- \left\{ p(t_{i-1}) \mu + \left( \int_0^\infty p(t_{i-1} + x_1) \right) + \cdots + \int_0^\infty p(t_i + x_1 + \cdots + x_{i-1}) + \int_0^\infty p(t_{i+1} + x_1 + \cdots + x_i) + \cdots + \int_0^\infty p(t_n + x_1 + \cdots + x_{n-1}) \right. \\
\left. G(dx_{n-1}) \cdots G(dx_1) \right. \mu \right\} \\
= \left\{ p(t_i) - \int_0^\infty \cdots \int_0^\infty p(t_{i-1} + x_1 + \cdots + x_{i-1}) G(dx_{i-1}) \cdots G(dx_1) \right. \\
\left. \mu \right\} - \left\{ p(t_{i-1}) - \int_0^\infty \cdots \int_0^\infty p(t_{i-1} + x_1 + \cdots + x_{i-1}) G(dx_{i-1}) \cdots G(dx_1) \right. \\
\left. \mu \right\},
\]

and this is \( \leq 0 \) for all \( t_{i-1} \leq t_i \) if and only if the function

\[
g(t) \equiv p(t) - \int_0^\infty \cdots \int_0^\infty p(t + x_1 + \cdots + x_{i-1}) G(dx_{i-1}) \cdots G(dx_1)
\]

is a decreasing function of \( t \). But we can write

\[
g(t) = \int_0^\infty \cdots \int_0^\infty (p(t) - p(t + x_1 + \cdots + x_{i-1})) G(dx_{i-1}) \cdots G(dx_1),
\]

so that it suffices to show that

\[
\tilde{g}(t) \equiv p(t) - p(t+y)
\]
is a decreasing function of $t$ for all $y \geq 0$. This is an immediate consequence of convexity of $p$.

Thus it is better to use the battery of age $t_{i-1}$ first and then LIFO than the battery of age $t_i$ and then LIFO. This implies that it is best to use the battery of age $t_1$ first and then LIFO, i.e., a LIFO policy is optimal for issuing all $n$ batteries.

Similarly, the same proof shows that a FIFO policy is optimal when $p$ is concave.

Notice in the proof that nothing is needed about $p$ being a decreasing function. In fact, $p$ need not even be monotone. This is important in cases where parts may deteriorate during one segment of their lives but may get better with age during another segment.

**Variation 2:**

This model is the same as above except that replacements become available. In particular, if an automobile arrives to be serviced, its old battery is taken out, one of the batteries on hand is put in, and then this old battery is revamped a random amount, so that it can be used again. Suppose the successive revamped batteries, after revamping, are the same as batteries with ages $T_1, T_2, \ldots$, i.i.d. random variables from distribution $H$.

We seek an issuing policy which maximizes the total expected lifetime of the cars serviced in a given amount of time, starting with batteries of ages $t_1 \leq \cdots \leq t_n$ on hand.
Theorem 5.3: If $p$ is convex and decreasing, a LIFO policy is optimal.

Proof: Notice first that the expected reward in any time $\tau$ from any policy is finite since it is bounded above by

$$1 \cdot \mu E(\text{number of arrivals in } [0, \tau]).$$

Suppose now that there is $\tau$ time left in the problem, batteries of ages $t_1 \leq \cdots \leq t_n$ are on hand, and a car arrives. Let $R(i)$ be the expected reward from issuing the battery of age $t_i$ first and then using an optimal policy $S_i$. We show that $R(i) \geq R(i+1)$, $i = 1, \ldots, n-1$, which is sufficient to prove the theorem.

We get a lower bound on $R(i)$ by computing the expected reward from issuing the battery of age $t_i$, which we label battery $i$, first and then following the non-optimal policy $\tilde{S}_i$ which issues exactly like $S_{i+1}$ except in one instance. When $S_{i+1}$ says to issue battery $i$, $\tilde{S}_i$ issues battery $i+1$. Let $\tilde{R}(i)$ be the expected reward from $\tilde{S}_i$. Then $\tilde{R}(i) \leq R(i)$ since $\tilde{S}_i$ is non-optimal.

Now $R(i+1)$ and $\tilde{R}(i)$ differ only in terms involving rewards from batteries $i$ and $i+1$. In fact, we have

$$\tilde{R}(i) - R(i+1) = (p(t_i) - p(t_{i+1}) + \sum_{j=1}^{\infty} \int_{A_j} p(t_{i+1} + s_j) - p(t_i + s_j) \mu,$$

where $s_j$ is the time elapsed until the $j$th car after the present one arrives; the integration is done with respect to the product measure.
\[
\bigcup_{k=1}^{\infty} G(dx_k) \, H(dy_k)
\]

where \( x_k \)'s are times between arrivals of cars and \( y_k \)'s are beginning ages of revamped batteries; and \( A_j \) is that part of the sample space where, under policy \( S_{i+1} \), battery \( i \) is issued to the \( j \)th car after the present one and \( s_j < \tau \). Thus \( A_j \) is determined completely by \( S_{i+1} \).

By convexity of \( p \), we have

\[
p(t_{i+1} + s_j) - p(t_i + s_j) \geq p(t_{i+1}) - p(t_i), \quad \text{for all } s_j \geq 0,
\]

so that

\[
R(i) - R(i+1) \geq p(t_i) - p(t_{i+1}) (1 - \sum_{j=1}^{\infty} \mathbb{P}(A_j)) \geq 0,
\]

by monotonicity of \( p \). Thus \( R(i) \geq R(i) \geq R(i+1) \), and the proof is complete.

**Variation 3:**

In this model we assume that \( n \) batteries of equal ages are on hand at the beginning of the problem, and that no replacements are possible. However, the batteries deteriorate at different rates. In particular, battery \( i \) has an associated function \( p_i \), such that if it is placed in a car at age \( x \), the car will run successfully with this battery for at least \( t \) units of time with probability \( p_i(x) \tilde{F}(t) \).
We assume that battery \( j \) deteriorates faster than battery \( i \), \( i+1 \leq j \leq n \), in the sense that

\[
(5.25) \quad p_i(x+y) - p_i(x) \geq p_j(x+y) - p_j(x) \quad \text{for all} \quad x, y \geq 0.
\]

If \( p_i \) and \( p_j \) are differentiable, this condition is equivalent to

\[
\frac{d}{dx} p_i(x) \geq \frac{d}{dx} p_j(x) \quad \text{for all} \quad x \geq 0.
\]

Again we attempt to find an issuing policy which maximizes the total expected lifetime the first \( n \) automobiles receive from these \( n \) batteries.

**Theorem 5.4:** If the functions \( p_1, \ldots, p_n \) satisfy (5.25) for all \( i < j \), then a FIFO policy is optimal. That is, it is best to issue the batteries in the order \( n, n-1, \ldots, 1 \).

**Proof:** The proof is by induction on \( n \). Suppose that when \( n-1 \) batteries are on hand and are all of the same age, a FIFO policy is optimal. Now suppose there are \( n \) batteries on hand, all of age \( x_0 \), a car arrives, battery \( i \) is issued, and an optimal policy (a FIFO policy) is used thereafter. The expected reward from this policy is

\[
p_i(x_0) + \left( \int_0^\infty p_n(x_0+x_1) \, dx_1 \right) + \left( \int_0^\infty p_{n-1}(x_0+x_1+x_2) \, dx_2 \right) + \cdots + \left( \int_0^\infty p_1(x_0+x_1+\cdots+x_{n-1}) \, G(dx_{n-1}) \cdots G(dx_1) \right) N.
\]
We show that this is less than or equal to the expected reward from issuing battery \( i+1 \) first and then using a FIFO policy. We have

\[
\left\{ p_i(x_0)\mu + \left( \int_0^\infty p_n(x_0 + x_1) + \cdots + \int_0^\infty p_{i+1}(x_0 + x_1 + \cdots + x_{n-i}) \right) \\
+ \int_0^\infty p_{i-1}(x_0 + x_1 + \cdots + x_{n-i+1}) \\
+ \cdots + \int_0^\infty p_1(x_0 + x_1 + \cdots + x_{n-1}) \, G(dx_{n-1}) \cdots G(dx_1) \mu \right\} \\
- \left\{ p_{i+1}(x_0)\mu + \left( \int_0^\infty p_n(x_0 + x_1) + \cdots + \int_0^\infty p_i(x_0 + x_1 + \cdots + x_{n-i}) \right) \\
+ \int_0^\infty p_{i-1}(x_0 + x_1 + \cdots + x_{n-i+1}) \\
+ \cdots + \int_0^\infty p_1(x_0 + x_1 + \cdots + x_{n-1}) \, G(dx_{n-1}) \cdots G(dx_1) \mu \right\}
\]

\[
= \left\{ \int_0^\infty \cdots \int_0^\infty \left( p_{i+1}(x_0 + x_1 + \cdots + x_{n-i}) - p_{i+1}(x_0) \right) \, G(dx_{n-i}) \cdots G(dx_1) \right\} \mu \\
- \left\{ \int_0^\infty \cdots \int_0^\infty \left( p_i(x_0 + x_1 + \cdots + x_{n-1}) - p_1(x_0) \right) \, G(dx_{n-i}) \cdots G(dx_1) \right\} \mu \leq 0
\]

from (5.25).

From this we conclude that it is best to issue battery \( n \) first and then follow a FIFO policy. Thus it is optimal to issue in the order \( n, n-1, \ldots, 1 \).
Again, notice that nothing is needed about monotonicity of the functions \( p_1, \ldots, p_n \). Also we do not need to know ordering relations between \( p_i(x) \) and \( p_j(x) \) for any \( x \). However, if \( p_1(0) = \cdots = p_n(0) \), then it will necessarily follow that \( p_i(x) \geq p_j(x) \) for all \( x \geq 0 \) when \( i < j \).
REFERENCES


In this paper men with fixed values \( p_1 \leq \cdots \leq p_n \) are to be assigned to jobs. The jobs arrive sequentially and assume values \( x_1, x_2, \cdots \), which are random variables. When a job arrives, its value is observed, a man is assigned to it, and a reward \( r(p,x) \), usually \( px \), is received. Various models are treated: (1) exactly \( n \) jobs are arriving, but the \( x \)'s may not be i.i.d. and not all of the men can do all of the jobs; (2) the number of jobs and the times they arrive are random, and not all of the jobs have to be used; (3) men can finish jobs and then become reavailable; (4) random replacements for assigned men are immediately available. Under the reward structure \( r(p,x) = px \), simple critical number policies, independent of the \( p \)'s, are found which maximize the expected total reward or reward per unit time. Other interpretations of the models, besides the men-jobs interpretation, are given when applicable.
### INSTRUCTIONS

1. **ORIGINATING ACTIVITY:** Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization (corporate author) issuing the report.

2. **REPORT SECURITY CLASSIFICATION:** Enter the overall security classification of the report. Indicate whether “Restricted Data” is included. Marking is to be in accordance with appropriate security regulations.

3. **GROUP:** Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

4. **REPORT TITLE:** Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

5. **DEScriptive NOTES:** If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

6. **AUTHOR(S):** Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

7. **REPORT DATE:** Enter the date of the report as day, month, year, or month, year. If more than one date appears on the report, use date of publication.

8. **TOTAL NUMBER OF PAGES:** The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

9. **NUMBER OF REFERENCES:** Enter the total number of references cited in the report.

10. **CONTRACT OR GRANT NUMBER:** If applicable, enter the applicable number of the contract or grant under which the report was written.

11. **PROJECT NUMBER:** Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

12. **ORIGINATOR’S REPORT NUMBER(S):** Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

13. **OTHER REPORT NUMBER(S):** If the report has been assigned any other report numbers (either by the originator or by the sponsor), also enter this number(s).

14. **AVAILABILITY/LIMITATION NOTICES:** Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

   1. “Qualified requesters may obtain copies of this report from DDC.”
   2. “For official announcement and dissemination of this report by DDC is not authorized.”
   3. “U.S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through ____________________________.”
   4. “U.S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through ____________________________.”
   5. “All distribution of this report is controlled. Qualified DDC users shall request through ____________________________.”

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

15. **SUPPLEMENTARY NOTES:** Use for additional explanatory notes.

16. **SPONSORING MILITARY ACTIVITY:** Enter the name of the departmental project office or laboratory sponsoring (paying for) the research and development. Include address.

17. **ABSTRACT:** Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as (T), (S), (C), or (U).

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 223 words.

18. **KEY WORDS:** Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, roles, and weights is optional.