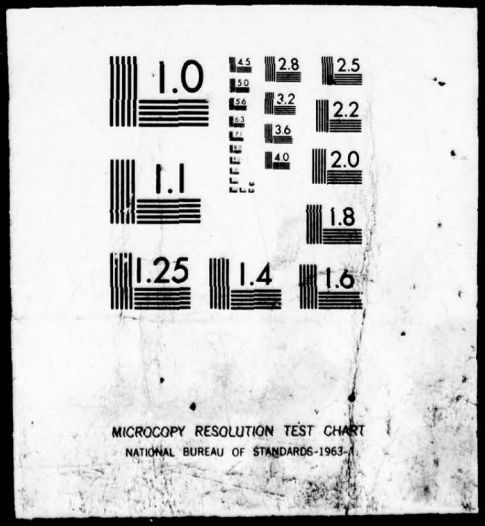
AD-A070 635 OFFICE OF NAVAL RESEARCH ARLINGTON VA F/G 15/5
NAVAL RESEARCH LOGISTICS QUARTERLY. VOLUME 26, NUMBER 2.(U)
JUN 79

UNCLASSIFIED NL

1 OF 3
AD A070635



LEVEL III

B.S.

AD A070635

NAVAL RESEARCH LOGISTICS QUARTERLY.

Volume 26, Number 2,

DDC RECEIVED JUN 29 1979 A

DDC FILE COPY.

JUNE 1979
VOL. 26, NO. 2

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

102p.

OFFICE OF NAVAL RESEARCH

NAVSO P-1278

265250

79 06 15 016

NAVAL RESEARCH LOGISTICS QUARTERLY

The Naval Research Logistics Quarterly is devoted to the dissemination of scientific information in logistics and will publish research and expository papers, including those in certain areas of mathematics, statistics, and economics, relevant to the over-all effort to improve the efficiency and effectiveness of logistics operations.

Information for Contributors is indicated on inside back cover.

The Naval Research Logistics Quarterly is published by the Office of Naval Research in the months of March, June, September, and December and can be purchased from the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402. Subscription Price: $11.15 a year in the U.S. and Canada, $13.95 elsewhere. Cost of individual issues may be obtained from the Superintendent of Documents.

Issuance of this periodical approved in accordance with Department of the Navy Publications and Printing Regulations, P-35 (Revised 1-74).

A METHODOLOGY FOR STUDYING THE DYNAMICS OF EXTENDED LOGISTIC SYSTEMS*

Stephen C. Graves† and Julian Keilson

Graduate School of Management
The University of Rochester
Rochester, New York

ABSTRACT

An extended logistic system is a well-defined configuration of equipment, modules, inventories, and repair and replacement facilities modeling a complex, repairable system with on-going repair. The design of such systems has been based largely on the static tools of inventory theory and reliability theory, i.e., on steady-state distributions and on associated means and variances. Such static tools suppress the scale of real time and ignore system persistence time in up-states and persistence time in down-states.

A reasonably simple dynamic methodology is presented, focusing on system failure time as a more meaningful objective function for system-design trade-off studies. In the presence of good reliability, it is shown that different candidates for system failure time effectively merge to yield an unambiguous, single system failure time. Examples illustrating the importance of dynamic information for system design are given.

INTRODUCTION

The design of complex, repairable systems has been based largely on static tools of inventory theory and reliability theory, i.e., on steady-state distributions and on associated mean values and variances. Such static tools neglect real time, a crucial dimension for the description of system behavior. Specifically, the persistence times of both satisfactory and unsatisfactory system performance are unavailable in static studies. Information on these persistence times is vital to an understanding of the dynamic behavior of the system, and hence it is vital to the evaluation of system performance. A *dynamic* treatment of system behavior quantifies such persistence times, and permits the study of the influence of system-design trade-offs on their frequency and duration. The purpose of this paper is to propose a methodology for the analysis of this dynamic behavior and to give a preliminary indication of how this analysis would be incorporated into the design decisions for complex systems.

This study consists of five parts. In Section I an informal study of the state of the art in system design is presented, i.e., of the tools of inventory theory and reliability theory currently employed. Here we discuss the limitations of current inventory theory and reliability theory for

*The research reported on here was supported in large part by the Air Force Business Research Management Center at Wright-Patterson Air Force Base, Dayton, Ohio, whose aid is gratefully acknowledged.
†Currently at A. P. Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Massachusetts

the study and understanding of the dynamic behavior of complex systems. In Section II, an alternative approach is proposed which may improve system understanding and, hence, improve system design. This model is an outgrowth of ideas developed earlier [1,2] for the study of system reliability. In Section III, the model is illustrated in terms of a simple single-item system. Section IV demonstrates how these single-item inventory techniques combine readily for the study of multi-item systems. In Section V, a simple example is presented to illustrate the necessity of dynamic information for making system-design decisions.

This document is presented at a reasonably simple mathematical level in keeping with practical needs. The supporting papers underlying the methodology [1-3] are more mathematically complete.

I. STATE-OF-THE-ART ANALYSIS

A. Extended Logistic Systems

An extended logistic system is a well-defined configuration of complex equipment, supporting inventory levels of components and modules, supporting maintenance facilities, supporting transportation system between local and remote inventory and maintenance sites, and procedures governing the allocation and shipment of components from remote to local sites.

For systems of interest to this study, breakdown, repair, and replacement are intrinsic elements, and system availability depends critically on component reliability, redundance built into the equipment, availability of repair personnel, levels of supporting inventories, and the delays associated with repair, replacement, and transportation from remote sites.

Randomness is implicit in all these elements. Specifically, components and modules are governed by failure-time distributions. Each failure type has an associated repair or replacement time distribution, and pipeline transportation times are random. Consequently, inventory levels fluctuate and system availability is impaired correspondingly.

Examples of extended logistic systems would be a squadron of aircraft, a radar system, or a network of communication satellites. For each of these systems, the basic unit of interest (an aircraft, a radar unit, or a satellite) is a complex combination of components which are subject to failure. For each component there are supporting inventory and/or repair facilities, and specific replacement procedures for such failures. System performance is evaluated according to system availability and the logistic costs required to obtain that level of availability.

B. Present Theory for System Design and Procurement

System design and procurement has relied heavily on current reliability and inventory theory. In this section we will briefly characterize the available tools from reliability and inventory theory and indicate the shortcomings of these tools for understanding the dynamic behavior of complex systems. The intent here is not to provide a comprehensive review of the current literature, but to indicate the need for a dynamic theory.

Reliability Theory

Reliability theory addresses itself to the modeling of systems subject to breakdown. A given reliability model may be classified by the presence or absence of the following characteristics in the system being described:

(a) *Complexity:* A system is complex if there are many modules that may fail.

(b) *Redundance:* A system which is complex has redundance when one or more modules may fail without the system failing. The redundance may be built into the system equipment or may be present in the form of standby equipment with instant replacement. For a redundant system, the functioning or nonfunctioning of the system is specified by a system structure function telling for which subsets of working modules the system works.

(c) *Activity:* A system model may be called active when modules which break down undergo repair or replacement. Up periods and down periods for the system then alternate. A system model that is not active may be called passive; consequently, passive systems are concerned solely with the time until the first system failure.

(d) *Dynamics:* The system model may be called dynamic when it is concerned with the transient behavior of the system, e.g., with the distribution of up times and down times for the system, and the correlation between these. When only steady-state information is given, e.g., the steady-state probability that the system is up (system availability), the model may be called static.

For most of the literature of reliability theory, one or more of these model characteristics are absent. Often treated are models in which the failure-time distribution or mean failure time of the system is inferred from that of the components in the absence of repair, i.e., passive models. Other reliability models which are active deal with systems whose level of complexity or redundance is modest, with only a few modules present or with very simple series or parallel structure. Those few models which are complex, redundant, and active are static and somewhat narrow in their applicability. The level of complexity present in these is far removed from most complex systems of interest, e.g., radar stations, or aircraft squadrons.

Inventory Theory

Inventory model studies are intrinsically redundant and active (drawing continuous replacement) in the sense above. Model characteristics appropriate to inventory models are:

(a) *Multi−item*: Here the concern is with stockage of a set of modules or items, as against single-item studies where there may or may not be dependence across items.

(b) *Multi−echelon*: Such studies are concerned for example, with forward inventories, regional inventories, and central inventories and their interaction, i.e., the relation between the delays and shortages at the different echelons.

(c) *Dynamic*: Again models are dynamic when they are concerned with the distribution of persistence times for shortages, distributions of times to depletion, etc. Studies concerned only with steady-state (ergodic) stockage levels, their distributions, expectation, variance, etc., may be called static.

The studies in the OR literature are largely single-item studies, with one echelon. The multi-echelon studies are single-item and static. The multi-item case can be treated if the different items behave independently and if only steady-state distributions are wanted.

For complex equipment to function, all modules must be functioning. Even though the inventory levels of different modules are effectively independent or may be treated as independent, and the joint distribution of module levels is the product of the individual distributions of

levels, of crucial concern to system availability is the availability of *complete sets* of individual modules. The stockage levels of individual components and modules interact through this requirement, and the dynamics of system availability is much harder than the statics.

For system design and procurement, the models needed are multi-item, multi-echelon and dynamic. Such models have not been available. In particular, the dynamic behavior of multi-item, multi-echelon systems has not been dealt with in logistics studies, and understanding, qualitative or quantitative, of fluctuation in system availability is poor. The only available tool for studying the dynamics of complex systems has been very detailed simulations, which can provide the system behavior for a specific system configuration, but which give little insight into an overall understanding of system behavior. The significance of this gap in system understanding will be discussed next.

C. The Study of System Trade-offs

The design of an extended logistic system consists logically of:

(a) Prediction of system performance and reliability characteristics attendant to a choice of system design parameters, i.e., component and/or module parameters; stockage levels at field and depot sites; personnel support levels; pipeline parameters;

(b) Determination of "feasible" system parameter choices described in (a) assuring minimally acceptable performance and reliability characteristics;

(c) Evaluation of net present value of total system cost obtained by adding the cost of initial procurement to inventory, repair, and pipeline operations cost (discounted) over the planned life of the system;

(d) Selection of that feasible system with minimal total system cost.

Such an extended logistic system-design procedure viewed as an optimization problem is enormously ambitious for complex systems, and full achievement is correspondingly unrealistic. Difficulties surrounding the problem of predicting system performance from its parameters and gaps in OR techniques for such prediction have been described in Section IB above. The optimization phase of the design — evaluation of total cost and selection of a minimal cost system — may only be possible crudely by examining cost changes associated with system-design trade-offs.

Such trade-off examinations are clearly impossible except within the framework of an overall system study, in which phases (a), (b), (c), and (d) are pursued as parts of a well-defined organic whole. The absence of any unified theory of extended logistic systems hampers such trade-off examinations by system designers on any meaningful, systematic basis. System design as now practiced is correspondingly makeshift and appears to be more of an engineering art than an engineering discipline.

Trade-offs of interest are component quality vs redundancy inside equipment vs stockage levels, field stockage levels vs central stockage levels, quality in one type of system component or module vs that in another type, and inventory levels of one item vs those of another. There is need to refine the intuitive notion of a "balanced" system (i.e., a system in which a change in investment represented in any two system levels meeting system needs raises system costs) and to develop simple rules permitting quick recognition of system imbalance.

D. The Importance of Understanding Dynamic System Behavior

Consider a single-item, single-echelon inventory, that of aircraft engines, say. The inventory level $n(t)$ fluctuates in time. A sample history of such a fluctuating level might have the appearance shown in Figure 1.

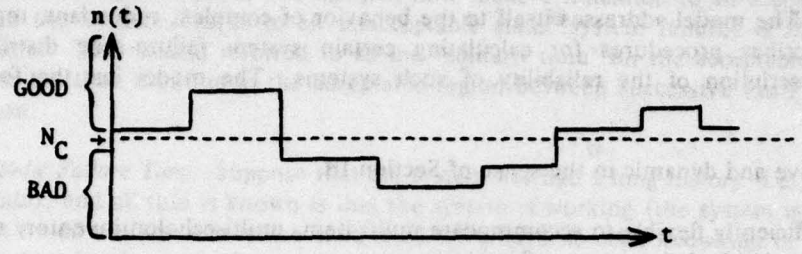


FIGURE 1.

A static description of the item availability might take any of the following forms:

(1) $\bar{n}$: the steady-state average number of items available.

(2) P_{GOOD}: the probability that the fluctuating level exceeds a specified critical level N_C regarded as being satisfactory. That is, for $n(t) \geqslant N_C$ the system is said to be in the "good" state and is functioning, while for $n(t) < N_C$ the system is in the "bad" state or has failed. P_{GOOD} is the "availability" of the item.

(3) p_r: the steady-state probability that r items are available at some time chosen at random.

The sequence of numbers p_r, $r = 0, 1, 2, \ldots$, is the steady-state probability distribution of the inventory level. From this distribution one may calculate all static information, such as the availability P_{GOOD}, the average $\bar{n}$, or the variance σ_n^2 of the distribution.

Of competing interest for such a fluctuating level $n(t)$ is the persistence time T_{GOOD} of levels above the critical level N_C and T_{BAD} for levels below N_C. One might wish to know the average of T_{GOOD}, say, or its variance, or the probability that T_{BAD} exceeds three days. Ideally, one would like the distribution of the good and bad times, but expectations and variances might suffice. Such dynamic information, however, is totally unavailable from the static information contained in the steady-state distribution P_r, and certainly not from any static descriptive such as $\bar{n}$ or P_{GOOD} obtained from it. Indeed, the steady-state probabilities entirely suppress any time-scale parameter of interest. The availability P_{GOOD} is a duty cycle parameter describing the ratio of good times to bad times, but it says nothing about whether there is fluctuation from GOOD to BAD once a day or once a year.

A fluctuation phenomenon of interest for system availability is *jitter*, the rapid fluctuation from good to bad levels when one is near the critical level. The character of such level jitter is again unavailable from the static information.

For multi-item, multi-echelon systems, the limitations of static answers are magnified. The higher dimension of the multi-item, multi-echelon process makes the system behavior more intricate and harder to visualize. The failure time of the system is the central random

variable describing the behavior of interest, and this cannot be obtained from steady-state information. (The concept of system failure time for complex, redundant, repairable systems is ambiguous. Three candidates of operational interest are described in Section II.)

E. A Possible Theoretical Framework for System Design

A theoretical model has been described in [1]. The methodology supporting the model is contained in [2]. The model addresses itself to the behavior of complex, redundant, repairable systems and describes procedures for calculating certain system failure-time distributions natural to the description of the reliability of such systems. The model has the following features:

(a) It is active and dynamic in the sense of Section IB.

(b) It is sufficiently flexible to accommodate multi-item, multi-echelon inventory systems present in the extended logistic systems of interest.

(c) The model describes the distribution of persistence times in acceptable (good) and unacceptable (bad) system states. Two related system failure times of operational interest, the "ergodic exit time" and "quasi-stationary exit time" are described and related to the persistence time. The three failure times describe, respectively, the duration of working times, the time remaining to failure when the system is working routinely, and the time remaining for veteran systems known to have been healthy for a very long time.

(d) Expected system failure times are available explicitly in terms of underlying parameters in the presence of item independence or time-reversibility, permitting calculation of steady-state probabilities of the system states.

II. SYSTEM FAILURE TIMES FOR COMPLEX, REDUNDANT, REPAIRABLE SYSTEMS

A. Definition of System Failure Times

Consider a complex, redundant, repairable (active) system as described in Section IA. It is assumed that the laws governing the system (e.g., failure rates, repair rates) do not change in time. Suppose the system can be either in an acceptable state (good state) or an unacceptable state (bad state). For example, in a single-item inventory* system with $N(t)$ being the number of failed items waiting to be repaired or replaced at time t, the system may be in an acceptable state if $N(t) \leqslant N_C$ = a specified critical level, while if $N(t) > N_C$ the system is in an unacceptable state. In such systems, we are interested in the behavior of the persistence time for the system in the acceptable region and in the unacceptable region. To study this behavior, we define *system failure* to occur when the system moves from an acceptable state to an unacceptable state. System failure time is the persistence time of the system in the acceptable region, i.e., the time until the sytem fails by going to an unacceptable state. For complex fluctuating systems, we need to find the system failure-time distribution to capture the dynamic behavior of the system.

For general systems, a sensible choice of (definition of) failure time is not obvious. For this study we will consider the following four specific system failure times whose simplicity of structure is in keeping with their natural intuitive simplicity:

*In a field setting the inventory level might be the availability level, i.e., the number of parts in service plus working spares.

Failure Time from the Perfect State: Assume that the system is in its best possible state: all components in the system are new and working, and the component inventory is fully stocked. The time until the system first reaches an unacceptable state (system failure) is called the "failure time from the perfect state."

Post-Recovery Failure Time: Suppose the system has just recovered, i.e., the system has been in an unacceptable state, and has just now made a transition to an acceptable state. The time until the system returns to an unacceptable state (system failure) is the "post-recovery failure time." This is also referred to as the "sojourn time" on the acceptable region, since it represents the time duration in the acceptable region between successive visits to the unacceptable region.

Ergodic Failure Time: Suppose that the system has had a long history (i.e., the system is in steady-state), and all that is known is that the system is working (the system is in an acceptable state, but we do not know in which state it is). We have *no other knowledge* of the performance of the system in the past. The time until the system stops working (system failure) is the "ergodic failure time."

Quasi-Stationary Failure Time: It is known that the system is currently working *and* that for as long as anyone can remember the system has been working. Again, we do not know in what specific working state the system is, other than it is in an acceptable state. The time until the system reaches an unacceptable state (system failure) is the "quasi-stationary failure time." This differs from the ergodic failure time in that for the ergodic failure time the possibility of one or more recent system failures is not dismissed.

In a similar manner, four analogous system recovery times could be defined, where we define system recovery to occur when the system moves from an unacceptable state to an acceptable state.

B. Properties of System Failure Times

The properties of these system failure times and their relationships are presented and discussed in [1]. The discussion given in [1] is based on the theoretical development and methodology contained in [2]. As employed in these papers, the following notation will be used to represent the four system failure times:

T_P = failure time from the perfect state;
T_V = post-recovery failure time;
T_E = ergodic failure time;
T_Q = quasi-stationary failure time.

We will assume that the system can be modeled as a stationary, *time-reversible* Markov chain (see [2], §1.3, §2.4). A stationary chain is time reversible if, for all values of t_1, t_2, m, and n,

$$\Pr[N(t_1) = m,\ N(t_2) = n] = \Pr[N(t_1) = n,\ N(t_2) = m],$$

where $N(t)$ is the state of the system at time t. Time reversibility is a common property for many Markov systems and is discussed in greater detail in [2]. In particular, the system consisting of a collection of independent components with exponentially distributed failure times and exponentially distributed repair or replacement times is time reversible.

Distributions for System Failure Times

We are interested in the probability distributions for these failure times. It is shown in [2], §6.6, that the quasi-stationary time T_Q is distributed as a pure exponential. That is, we have

(1) $s_Q(\tau)$ = probability density function for T_Q

$= \gamma \exp(-\gamma\tau)$

and

(2) $\bar{F}_Q(\tau)$ = survival function for T_Q

$= \text{Prob}[T_Q > \tau] = \exp(-\gamma\tau)$,

where $\bar{T}_Q$ = expected value of $T_Q = 1/\gamma$.

The distributions for both the ergodic failure time T_E and the post-recovery exit time T_V are mixtures of pure exponential distributions (see [2], §6.9). We can write for the ergodic failure time

(3) $s_E(\tau)$ = probability density function for T_E

$= \sum_j p_{E_j} \gamma_j \exp(-\gamma_j \tau)$

and

(4) $\bar{F}_E(\tau)$ = survival function for T_E

$= \text{Prob}(T_E > \tau) = \sum_j p_{E_j} \exp(-\gamma_j \tau)$,

where $p_{E_j} \geq 0$ for all j, $\gamma_j > 0$ for all j,

$\sum_j p_{E_j} = 1$,

$\bar{T}_E$ = expected value of $T_E = \sum_j p_{E_j} (1/\gamma_j)$.

The set of values $\{\gamma_j\}$ is common to all four failure times, and γ, the failure rate for the quasi-stationary time, is equal to $\min_j \{\gamma_j\}$ (see (2)).

Similarly, we have for the post-recovery failure time

(5) $s_V(\tau) = \sum_j p_{V_j} \gamma_j \exp(-\gamma_j \tau)$

and

(6) $\bar{F}_V(\tau) = \sum_j p_{V_j} \exp(-\gamma_j \tau)$,

where $p_{V_j} \geq 0$ for all j, $\sum_j p_{V_j} = 1$,

$\bar{T}_V = \sum_j p_{V_j} (1/\gamma_j)$.

The terms $\{p_{E_j}\}$ and $\{p_{V_j}\}$ are the *mixing distributions* for the ergodic failure time and the post-recovery failure time, respectively. These two system failure times differ only by their mixing distributions, $\{p_{E_j}\}$ and $\{p_{V_j}\}$.

One useful property of a probability distribution that is a mixture of pure exponential distributions is that the survival function, $\bar{F}(\tau) = \text{Prob}(T > \tau)$, is log-convex, that is, $G(\tau) = \log [\bar{F}(\tau)]$ is a convex function Note that in the case where $\bar{F}(\tau) = \exp(-\gamma\tau)$, the survival function is both log-convex and log-concave. Thus the survival functions for both the post-recovery and the ergodic failure times are log-convex, while the survival function for the quasi-stationary failure time is both log-convex and log-concave.

The distribution for the failure time from the perfect state may also be written in terms of pure exponential distributions, but it is *not* a mixture of exponential distributions. That is, we have

$$s_P(\tau) = \sum_j p_{p_j} \gamma_j \exp(-\gamma_j \tau) \tag{7}$$

and

$$\bar{F}_p(\tau) = \sum_j p_{p_j} \exp(-\gamma_j \tau), \tag{8}$$

where $\sum_j p_{p_j} = 1$,

$$\bar{T}_p = \sum_j p_{p_j}(1/\gamma_j).$$

This is very similar to the expressions for the ergodic failure time and the post-recovery failure time. The critical distinction is that here p_{p_j} is not nonnegative for all j; hence, $\{p_{p_j}\}$ is not a mixing distribution. Thus the survival function for the failure time from the perfect state is not log-convex. Rather, it can be shown that the survival function is log-concave: $G(\tau) = \log [\bar{F}(\tau)]$ is a concave function.

In the context of reliability theory, we can add a further interpretation for notions of log-convexity and log-concavity of the survival functions (see [2], §5.9). A distribution with a strictly log-convex survival function (e.g., T_E and T_V) has a *decreasing failure rate.* The failure rate or likelihood of failure in any instant in time decreases with the age of the system; the longer the system operates without failure, the more reliable the system becomes.* Conversely, a distribution with a strictly log-concave survival function (e.g., T_p) has an increasing failure rate. Here the system becomes less reliable as the time since the last system failure increases.

The four system failure times are related by the following inequalities (see [2], §6.9):

$$\bar{F}_p(\tau) \geq \bar{F}_Q(\tau) \geq \bar{F}_E(\tau) \geq \bar{F}_V(\tau). \tag{9}$$

Hence, in terms of system survival (absence of failure), it is better to start in the perfect state than in the quasi-stationary state ($\bar{F}_p(\tau) \geq \bar{F}_Q(\tau)$), better to start in the quasi-stationary state than in the ergodic state ($\bar{F}_Q(\tau) \geq \bar{F}_E(\tau)$), and better to start in the ergodic state than in the post-recovery state ($\bar{F}_E(\tau) \geq \bar{F}_V(\tau)$). A direct result from this relationship is the ordering of the expected values of the failure times:

$$\bar{T}_p \geq \bar{T}_Q \geq \bar{T}_E \geq \bar{T}_V. \tag{10}$$

*This outwardly paradoxical behavior must be understood in its context. The failure times are for systems observed only at $t = 0$ and the subsequent failure time. An observed random system is continually reconditioned by observation, and prediction is constantly modified.

C. Limiting Behavior of System Failure Times

The limiting behavior of the system failure times as system reliability increases is very informative. It is shown in [2], §8.4, that the distributions for both the ergodic failure time and the failure time from the perfect state are exponential in the limit as system reliability is increased. Examples of highly reliable systems are systems for which the failure rates for the components are small relative to the components' repair or replacement rates or for which the inventory of space components is large relative to the need for replacement components. In practice, most complex systems are designed to be highly reliable, due to the heavy penalties for system failure. Furthermore, the convergence to exponentiality of the distributions $\bar{F}_E(\tau)$ and $\bar{F}_p(\tau)$ is quite fast for increasing system reliability, as will be seen in the next section.

The limiting behavior of the post-recovery failure time is complicated by the presence of *jitter*. Jitter occurs when the system, having just recovered from being in the unacceptable region, has a tendency to vacillate between the acceptable and unacceptable region before embarking upon an extended sojourn to the acceptable region. The amount of jitter in a system will vary with the relative magnitudes of the failure and repair rates at the boundary between the acceptable and unacceptable regions. If at this boundary the repair rates dominate the failure rates, then the jitter factor will be small. For highly reliable systems, we would expect this to be the case and, hence, the amount of jitter to be small. In the absence of jitter, the distribution of the post-recovery failure time will become exponential as the system reliability increases.

It is important to know how close to exponential these failure times are. A measure for the closeness to exponentiality (in the class of mixtures of exponentials) for the ergodic failure time and for the post-recovery failure time is $(\sigma^2/\bar{T}^2) - 1$ (see [2], §8.7, and [4]), where σ^2 is the variance of the failure time, and $\bar{T}$ is the expected value of the failure time. For both the ergodic failure time and the post-recovery failure time, this measure is nonnegative; the closer the measure is to zero, the closer the distribution is to being an exponential distribution.

We have defined four system failure times that are of interest in the design and analysis of complex, redundant repairable systems. We have stated various properties for the distributions of these failure times, and have shown the interrelationship among these distributions. Of particular interest is the fact that the distributions of the system failure times for highly reliable systems are exponential or nearly exponential. The significance of this observation lies in the fact that the exponential distribution is completely characterized by one parameter, the mean value. Hence, for reliable systems, given only the means for the system failure times, we may very accurately approximate the distributions for these failure times. Furthermore, the mean values for the system failure times, particularly for the post-recovery failure time, are obtainable (see [2], §6.7, §6.8). For particular systems, such as a system modeled by a birth-death process, these mean values can be computed from analytical expressions or, at worst, can be found by means of simple computer programs. Thus, the information provided by these system failure times can be easily incorporated into the design and analysis of reliable systems.

III. ILLUSTRATION OF SYSTEM FAILURE TIME BEHAVIOR

In this section we consider systems which can be modeled as birth-death processes. We indicate how the system failure-time distributions may be computed, and illustrate these distributions for a set of specific models.

A. Birth-Death Processes

Consider a single-item inventory system which may be modeled as a birth-death process. The system at time t is characterized by $N(t)$ = the number of items that have failed and are being repaired or are waiting to be replaced. The state variable ranges from 0 failed items (the perfect state) to K failed items, where K equals the maximum possible number of available items. For any state $N(t) = n$, system transitions are characterized by λ_n, equal to the transition rate from n failed items to $n + 1$ failed items, and by μ_n, equal to the transition rate from n failed items to $n - 1$ failed items. That is, when n items are not working ($K - n$ items are available), λ_n is the *failure rate* and μ_n is the *repair rate*. By definition, $\mu_0 = 0$ and $\lambda_K = 0$. The system and its transitions may be visualized as in Figure 2, where the boxes represent the states of the system and the dotted arrows are the potential transitions. For these systems, the acceptable region is specified by N_C, equal to the critical level; the system is in an acceptable state if $N(t) \leqslant N_C$ and is in an unacceptable state if $N(t) > N_C$.

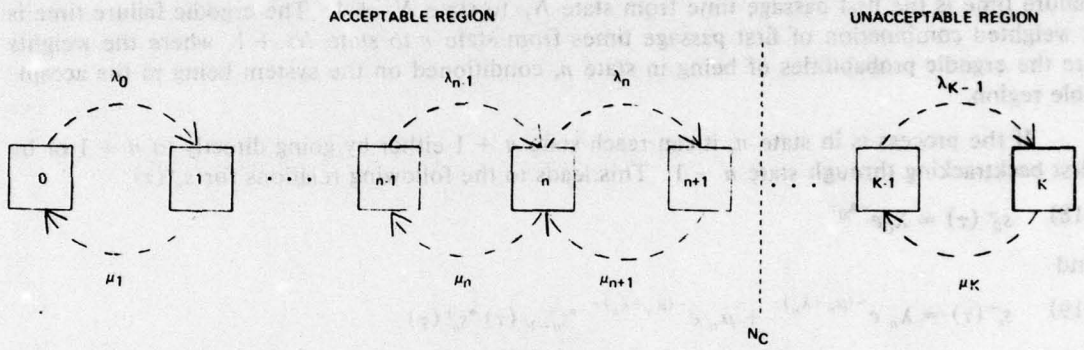


FIGURE 2.

In the traditional static models, attention is focused on the steady-state properties for the system. For the birth-death process, the steady-state probabilities are easily found as follows (see [2], §3.3):

(11) e_n = steady-state probability of system being in state n

$$= \Pr[N(t = \infty) = n]$$

$$= \pi_n \Big/ \sum_{m=0}^{K} \pi_m,$$

where

(12) $\pi_0 = 1$

and

(13) $\pi_n = \pi_{n-1}\, \lambda_{n-1} / \mu_n \quad \text{for} \quad n = 1, 2, \ldots, K.$

B. Computation of System Failure-Time Distribution*

Let $s_n^+(\tau)$ be the first-passage-time density from state n to state $n + 1$ for the process $N(t)$. The densities for the system failure times (except for s_0) can be stated in terms of $s_n^+(\tau)$ as follows:

*The methodology in this section was introduced in [3] and extends earlier calculations of Ross and Huisjes [5].

(14) $s_p(\tau) = s_0^+ {}^* s_1^+ {}^* \cdots {}^* s_{N_C}^+ = \hat{s}_0(\tau)$,

(15) $s_V(\tau) = s_{N_C}^+(\tau)$,

and

(16) $s_E(\tau) = \sum_{n=0}^{N_C} e_n \hat{s}_n(\tau) \Big/ \sum_{n=0}^{N_C} e_n$,

where

(17) $\hat{s}_n(\tau)$ = first passage time density from state n to state $N_C + 1$ for $n \leqslant N_C$

$= s_n^+ {}^* s_{n+1}^+ {}^* \cdots {}^* s_{N_C}^+(\tau)$

and * indicates convolution.

In words, the failure time from the perfect state is the first passage time from the perfect state (0) to state $N_C + 1$ (the first state reached in the unacceptable region). The post-recovery failure time is the first passage time from state N_C to state $N_C + 1$. The ergodic failure time is a weighted combination of first passage times from state n to state $N_C + 1$, where the weights are the ergodic probabilities of being in state n, conditioned on the system being in the acceptable region.

If the process is in state n, it can reach state $n + 1$ either by going directly to $n + 1$ or by first backtracking through state $n - 1$. This leads to the following relations for $s_n^+(\tau)$:

(18) $s_0^+(\tau) = \lambda_0 e^{-\lambda_0 \tau}$

and

(19) $s_n^+(\tau) = \lambda_n e^{-(\mu_n + \lambda_n)\tau} + \mu_n e^{-(\mu_n + \lambda_n)\tau} {}^* s_{n-1}^+(\tau) {}^* s_n^+(\tau)$,

for $n > 0$.

Taking Laplace transforms of (18) and (19), we obtain

(20) $\sigma_0^+(s) = L[s_0^+(\tau)] = \dfrac{\lambda_0}{\lambda_0 + s}$

and

(21) $\sigma_n^+(s) = L[s_n^+(\tau)] = \dfrac{\lambda_n}{s + \lambda_n + \mu_n - \mu_n \sigma_{n-1}^+(s)}$.

Differentiation of (20) and (21) gives

(22) $\dfrac{d}{ds}\sigma_0^+(s) = \dfrac{-\lambda_0}{(\lambda_0 + s)^2}$

and

(23) $\dfrac{d}{ds}\sigma_n^+(s) = \dfrac{-\lambda_n}{[s + \lambda_n + \mu_n - \mu_n \sigma_{n-1}^+(s)]^2}\left[1 - \mu_n \dfrac{d}{ds}\sigma_{n-1}^+(s)\right]$.

By induction over n, we get

(24) $\dfrac{d}{ds}\sigma_n^+(s) < 0, \quad n = 0, 1, 2, \ldots,$

for all s, except at singularities. Furthermore, we have, also by induction

(25) $\lim_{s \to \pm\infty} \sigma_n^+(s) = 0, \quad n = 0, 1, 2, \ldots,$

and

(26) $\sigma_n^+(0) = 1, \quad n = 0, 1, 2, \ldots .$

It is easy to show by induction that $\sigma_n^+(s)$ is a rational function of s and can be written as

(27) $$\sigma_n^+(s) = \frac{Q_n(s)}{P_{n+1}(s)},$$

where $Q_n(s)$ = polynomial of degree n

and

$P_{n+1}(s)$ = polynomial of degree $n + 1$.

Substituting (27) into (21) and (20) gives the following recursive equations:

(28) $Q_n(s) = \lambda_n P_n(s), \qquad n = 1, 2, \ldots ,$

(29) $P_{n+1}(s) = (s + \lambda_n + \mu_n)P_n(s) - \mu_n Q_{n-1}(s), \quad n = 1, 2, \ldots ,$

(30) $Q_0(s) = \lambda_0,$

and

(31) $P_1(s) = \lambda_0 + s.$

Now, substituting (28) into (29) and accounting for the initial conditions, we can rewrite (27) as

(32) $$\sigma_n^+(s) = \frac{\lambda_n P_n(s)}{P_{n+1}(s)}, \qquad n = 0, 1, 2, \ldots ,$$

where

(33) $P_{n+1}(s) = (s + \lambda_n + \mu_n)P_n(s) - \mu_n \lambda_{n-1} P_{n-1}(s), \quad n = 1, 2, 3, \ldots ,$

(34) $P_0(s) = 1,$

(35) $P_1(s) = \lambda_0 + s.$

We will now show the following theorem:

THEOREM: For $n = 0, 1, 2, \ldots , \sigma_n^+(s)$ has n simple zeros and $n + 1$ simple poles, and between any two adjacent poles there lies exactly one simple zero. In addition, the zeros of $\sigma_n^+(s)$ correspond to the poles of $\sigma_{n-1}^+(s)$.

PROOF: Consider $\sigma_0^+(s) = \dfrac{\lambda_0}{\lambda_0 + s}$; $\sigma_0^+(s)$ has a simple pole at $s = -\lambda_0$. Then $\sigma_1^+(s)$ has a simple zero at $s = -\lambda_0$; now, from properties (24), (25), and (26), $\sigma_1^+(s)$ must have a pole in $(-\infty, -\lambda_0)$ and a pole in $(-\lambda_0, 0)$. Furthermore, these poles are simple, since the denominator of $\sigma_1^+(s)$ is a polynomial of degree two.

Suppose the theorem holds for $j = 0, 1, \ldots , n - 1$; we need to show that it holds for $\sigma_n^+(s)$. Clearly, $\sigma_n^+(s)$ has n simple zeros corresponding to the n simple poles of $\sigma_{n-1}^+(s)$. By properties (24), (25), and (26), $\sigma_n^+(s)$ must have $n+1$ poles and each zero must be flanked by two poles. Finally, these poles are simple, since $\sigma_n^+(s)$ has a maximum of $n + 1$ poles.

A direct implication of this theorem is that we have

(36) $$P_n(s) = \prod_{i=1}^{n} (s + \theta_{ni}), \qquad n = 1, 2, \ldots ,$$

where

(37) $\theta_{11} = \lambda_0$

(38) $0 < \theta_{n,i} < \theta_{n,i+1}$, for $n = 1, 2, \ldots, i = 1, 2, \ldots, n-1$,

(39) $\theta_{n,i} < \theta_{n-1,i} < \theta_{n,i+1}$, for $n = 1, 2, \ldots, i = 1, 2, \ldots, n-1$.

Now we can use (32) and (36) to reexpress (14) as

$$(40)\quad L[s_P(\tau)] = \frac{\lambda_0\lambda_1 \cdots \lambda_{N_C}}{P_{N_C+1}(s)} = \frac{\lambda_0\lambda_1 \cdots \lambda_{N_C}}{\prod_{i=1}^{N_C+1} (s + \theta_{N_C+1,i})} = \prod_{i=1}^{N_C+1} \frac{\beta_{Pi}}{s + \theta_{N_C+1,i}},$$

where β_{Pi} is the residue of $L[s_P(\tau)]$ at the pole corresponding to $s = -\theta_{N_C+1,i}$. Inverting (40), we obtain

$$(41)\quad s_P(\tau) = \sum_{i=1}^{N_C+1} \beta_{Pi} \exp(-\tau\theta_{N_C+1,i}).$$

Similarly, for the post-recovery time, we have from (15)

$$(42)\quad L[s_V(\tau)] = \frac{\lambda_{N_C} P_{N_C}(s)}{P_{N_C+1}(s)} = \frac{\prod_{i=1}^{N_C} (s + \theta_{N_C,i})}{\prod_{i=1}^{N_C+1} (s + \theta_{N_C+1,i})} = \prod_{i=1}^{N_C+1} \frac{\beta_{Vi}}{s + \theta_{N_C+1,i}}.$$

Hence

$$(43)\quad s_V(\tau) = \sum_{i=1}^{N_C+1} \beta_{Vi} \exp(-\tau\theta_{N_C+1,i}),$$

where β_{Vi} is the residue at the pole $s = -\theta_{N_C+1,i}$. For the ergodic failure time, we have from (16)

$$(44)\quad L[s_E(\tau)] = \frac{\sum_{n=0}^{N_C} e_n \lambda_n \lambda_{n+1} \cdots \lambda_{N_C} P_n(s)}{P_{N_C+1}(s) \sum_{n=0}^{N_C} e_n}$$

$$= \sum_{i=1}^{N_C+1} \frac{\beta_{Ei}}{s + \theta_{N_C+1,i}}.$$

Hence

$$s_E(\tau) = \sum_{i=1}^{N_C+1} \beta_{Ei} \exp(-\tau\theta_{N_C+1,i}), \tag{45}$$

where β_{Ei} is the residue at the pole $s = -\theta_{N_C+1,i}$.

Therefore, to compute these three failure time distributions we need to find the zeros of $P_{N_C+1}(s)$ and the three sets of residues from the Laplace transforms. The quantity $P_n(s)$ can be computed recursively from relations (33), (34), and (35). At each stage of the recursion, the n zeros of $P_n(s)$ can be found by direct search. Relations (37), (38), and (39) can be used to speed up this search; that is, adjacent zeros for $P_{n-1}(s)$ are used as brackets for the search of the zeros of $P_n(s)$. To find the set of residues, we need the Laplace transform of the failure-time density; given the polynomials $P_n(s)$, these transforms are easily found using (40), (42), or (44). The residue β_{*i} is found by evaluating $(s + \theta_{N_C+1,i})\, L[s_*(\tau)]$ at $s = -\theta_{N_C+1,i}$.

The quasi-stationary failure time density is now just the pure exponential,

$$s_Q(\tau) = \theta_{N_C+1,1} \exp(-\tau\theta_{N_C+1,1}). \tag{46}$$

C. Graphs Demonstrating Failure Time Behavior

We consider four possible systems. For each system we assume that $K = 10$ and $N_C = 5$. That is, there are a maximum of ten items available, and at least five of these items are required to be in working order at any time. If six or more items have failed (less than five working), the system is in an unacceptable state. The four systems are specified by their repair and failure rates as follows (λ and μ are constant parameters):

A. $\lambda_n = \lambda,\ \mu_n = \mu;$

B. $\lambda_n = (K - n)\lambda,\ \mu_n = \mu;$

C. $\lambda_n = \lambda,\ \mu_n = n\mu;$

D. $\lambda_n = (K - n)\lambda,\ \mu_n = n\mu.$

System A assumes that the repair rate (μ_n) is independent of the number of failed items, and the failure rate (λ_n) is independent of the number of working items. For instance, the repair facility may be able to work on only one item at a time, or there may only be one repairman. Item failures may depend on item usage, but total usage is held constant over the system; hence, the more items working, the less usage, and thus the less exposure to failure there is for each of the working items. This system is analogous to an M/M/1 queueing model: the system has one server (repairman), and the arrivals to the system (failures) occur at a constant rate independent of the queue length (number of working items).

System B differs from System A in that the failure rate (λ_n) now depends on the number of working items ($K - n$). For instance, item failure again depends on item usage, but now individual item usage is constant for all working items and does not depend on the number of working items. Thus, total usage and hence the failure rate increase with the number of working items.

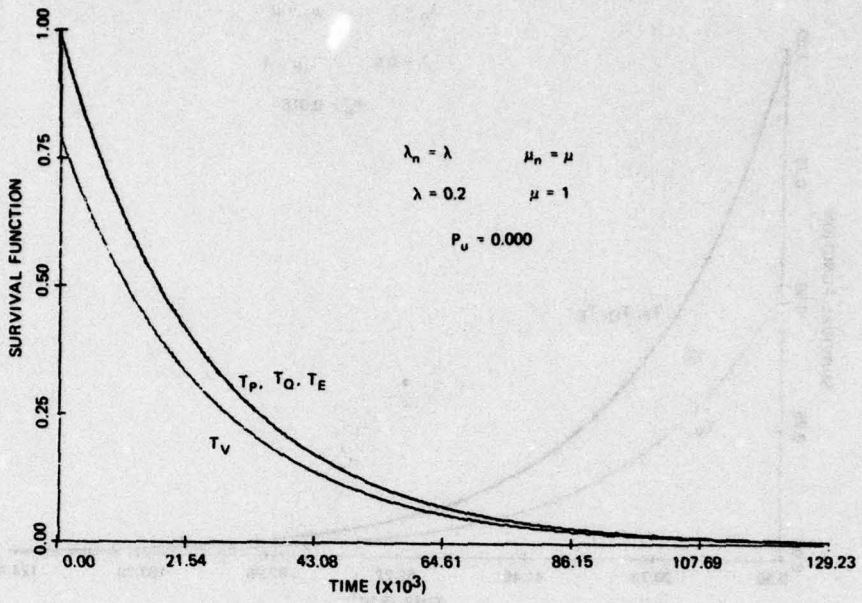
System C differs from System A in that the repair rate (μ_n) now depends on the number of failed items (n). The repair facility has sufficient capacity or repairmen to be able to work concurrently on all failed units. This system is analogous to the M/M/∞ queueing system; the system has an infinite number of servers (infinite repair capacity), but arrivals to the system (failures) are governed by a constant arrival rate.

System D assumes that the repair rate (μ_n) depends on the number of failed items and the failure rate (λ_n) depends on the number of working items.

We have computed and graphed the system failure time distributions for the four systems for $\mu = 1$, $\lambda = 0.2$ and for $\mu = 1$, $\lambda = 0.5$. For each system and each set of parameter values (λ, μ), we have graphed (1) the survival functions for the system failure times, and (2) the log of the survival functions for the system failure times. In addition, for each system and for each set of parameter values, P_u will represent the steady-state probability that the system is in an unacceptable state ($N(t) > N_C = 5$). This probability will be an indication of the reliability of the system.

From Figures 3 through 10, the properties of the system failure times are quite evident. The survival functions and their logs are clearly ordered. The quasi-stationary failure time is a pure exponential; its log survival function is linear. Furthermore, from the graphs of the log survival functions, the log-convexity of the ergodic failure time and post-recovery time are quite evident, as is the log-concavity of the failure time from the perfect state. It is also clear from the log survival functions that all four of the system failure times are quite exponential in nature. Each of their log survival functions quickly becomes linear (i.e., the survival function is exponential) and parallel. Thus, after an initial "burn-in" stage, each of the four system failure times behaves as if it were a pure exponential — that is, the rate of system failure is constant over time. In addition, we see that for highly reliable systems (e.g., system A, system C, or system D), as denoted by P_u, the system failure times are essentially exponential. Here, the ergodic failure time and the failure time from the perfect state coincide with the quasi-stationary failure time, while the post-recovery failure time differs from these only by an initial displacement of probability. This initial displacement of probability is what we defined to be jitter; that is, it represents the tendency of the system to vacillate on the boundary between the acceptable and unacceptable regions. For system B, for the particular parameter values (λ, μ) that we have considered, the system is not reliable; here the four system failure times are quite distinct. However, as previously mentioned, the system failure times for system B are still quite exponential in nature, as seen by the log survival functions, which become linear and parallel.

In summary, this evidence clearly supports the possible utility of these system failure times in decision-making. For reliable systems, the system failure times are essentially exponential, and consequently they can be completely characterized by the mean values of the distributions. Thus, the information provided by these system failure times may be easily incorporated into any analysis concerned with the inherent trade-offs in a complex system (e.g., increasing the inventory of an item vs decreasing the item's failure rate vs increasing the capacity of the repair facility). Furthermore, for more realistic real world models, failure and repair time distributions need not be and, in general, will not be exponentially distributed. However, the presence of exponentiality in the *system failure time distributions*, induced by reliability, is a *robust* property, i.e., the property is not sensitive to the underlying failure and repair time distributions. Hence, the conclusions from these examples are extendable to more realistic systems.



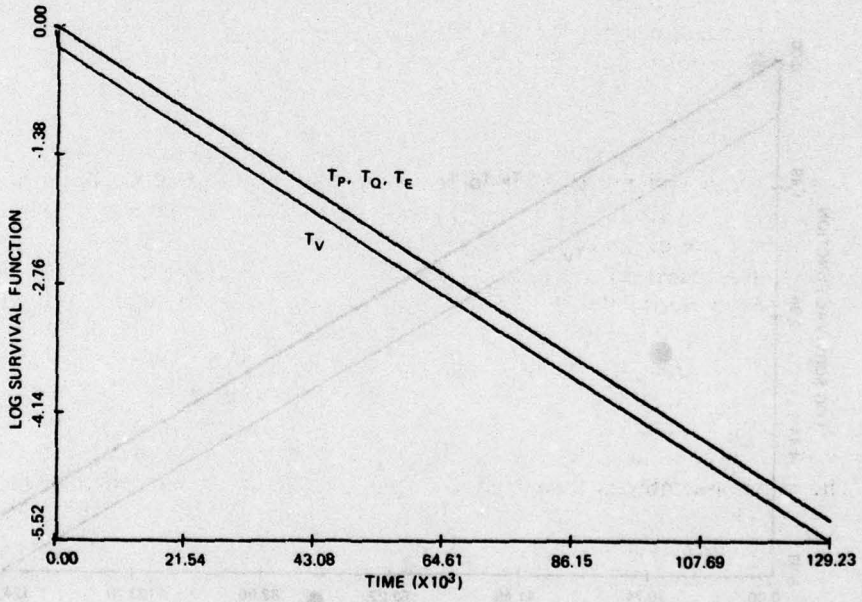


FIGURE 3. System A

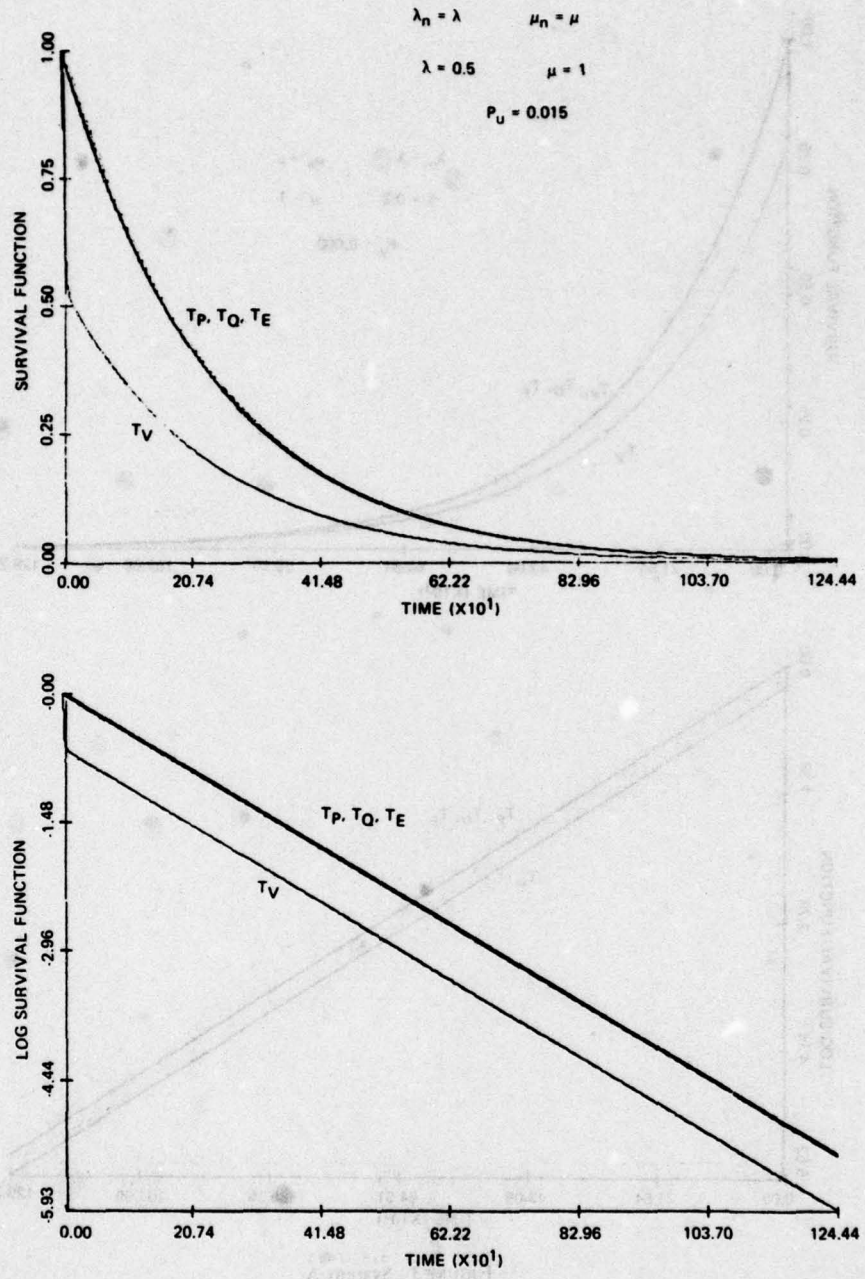


FIGURE 4. System A

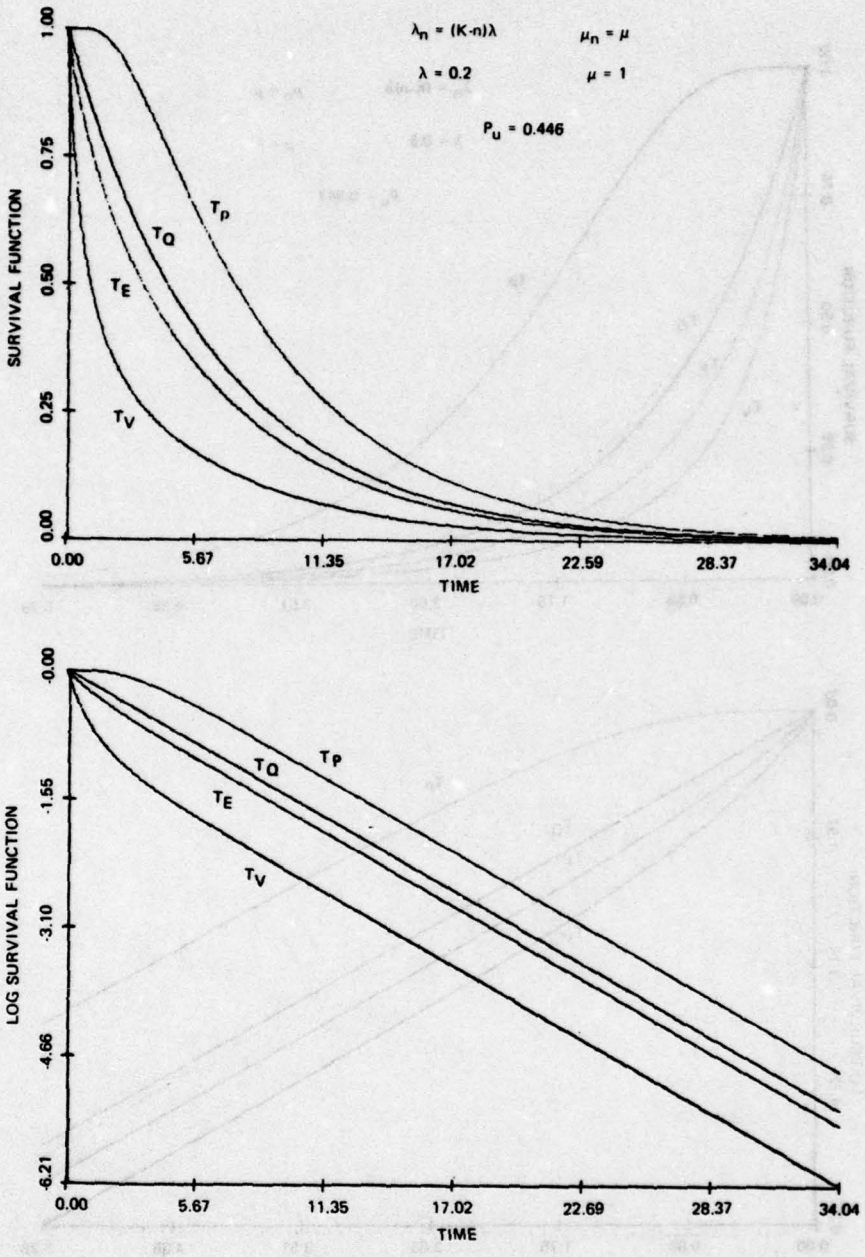
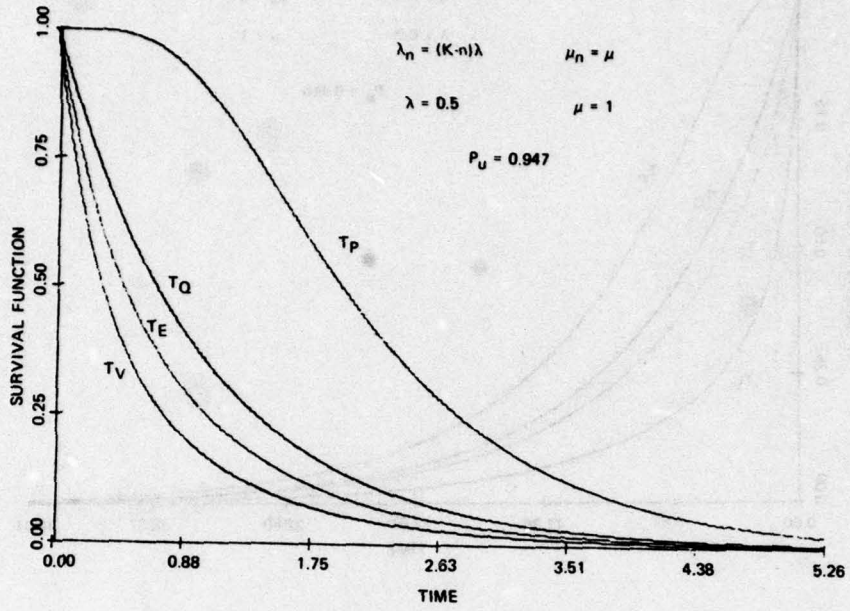


FIGURE 5. System B



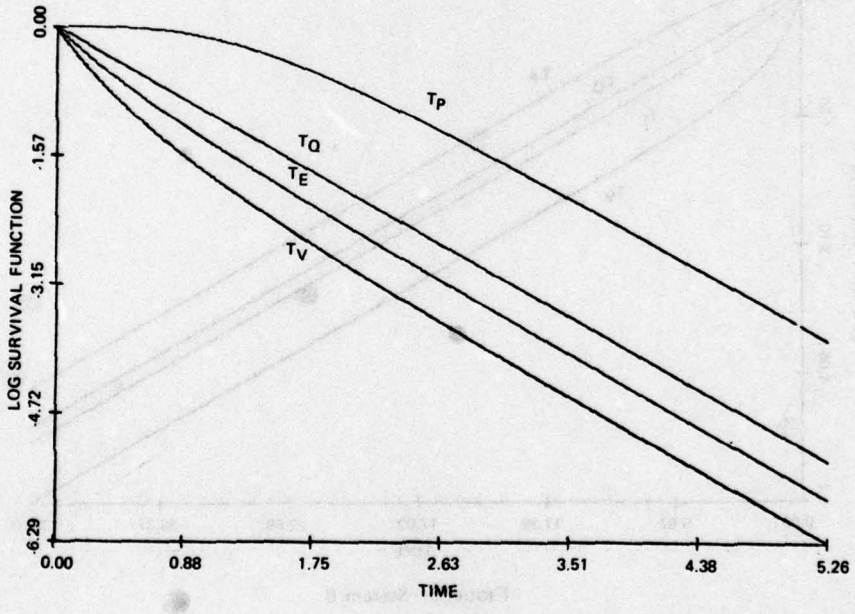


FIGURE 6. System B

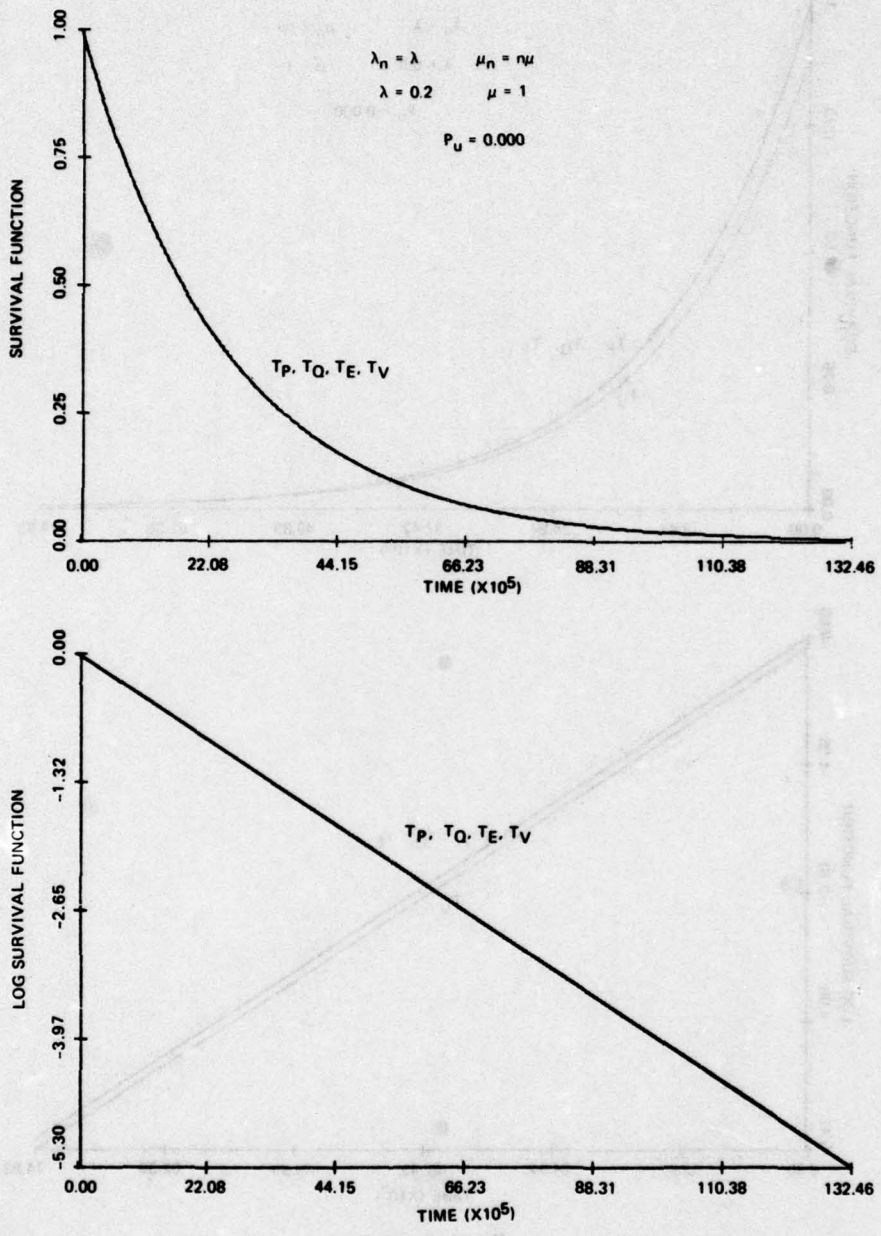


FIGURE 7. System C

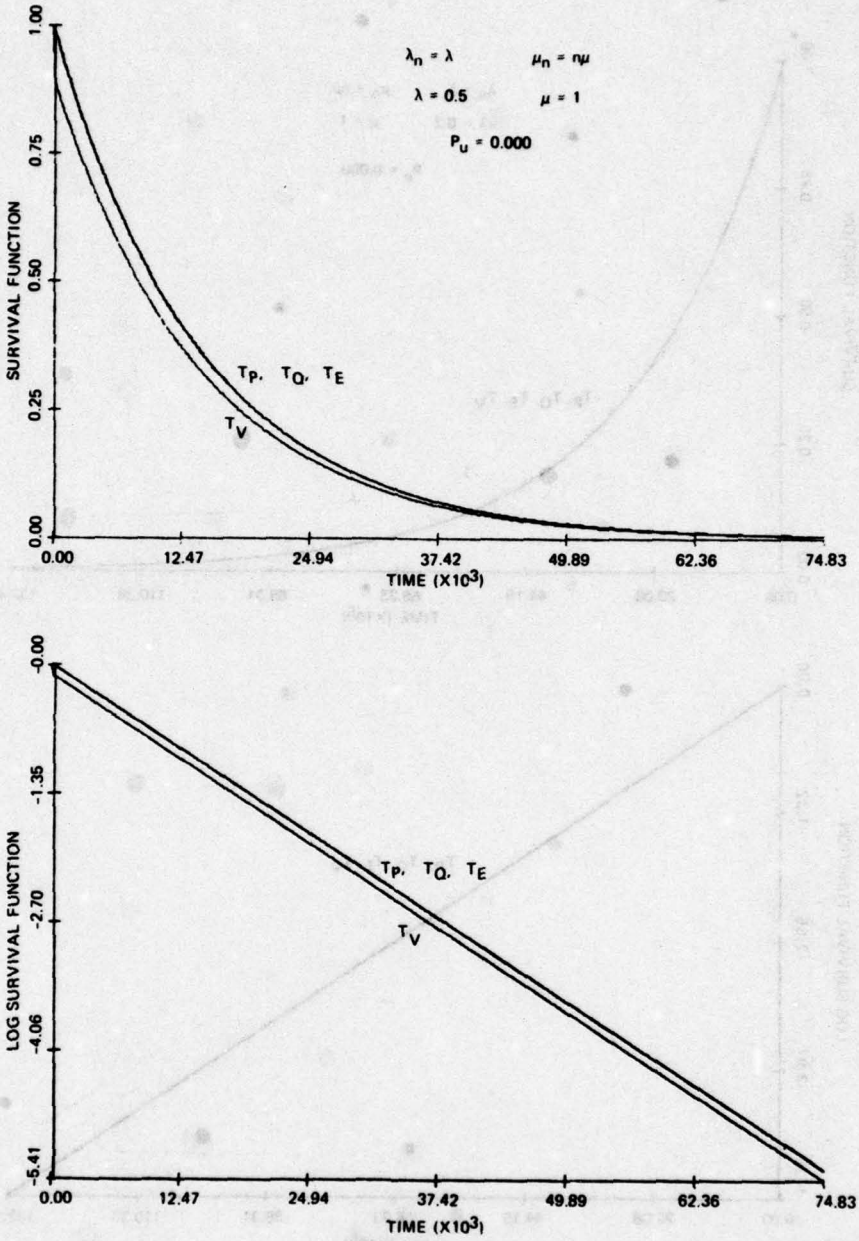


FIGURE 8. System C

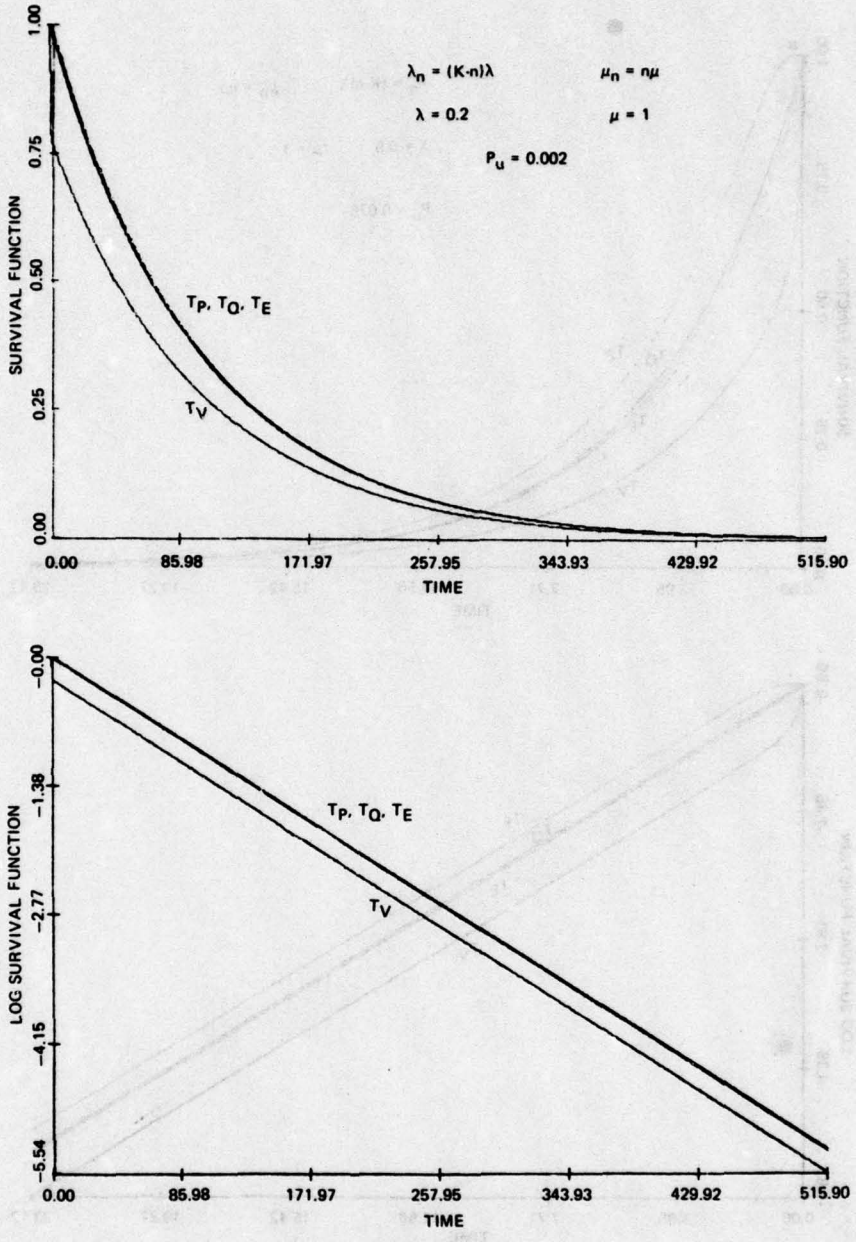


FIGURE 9. System D

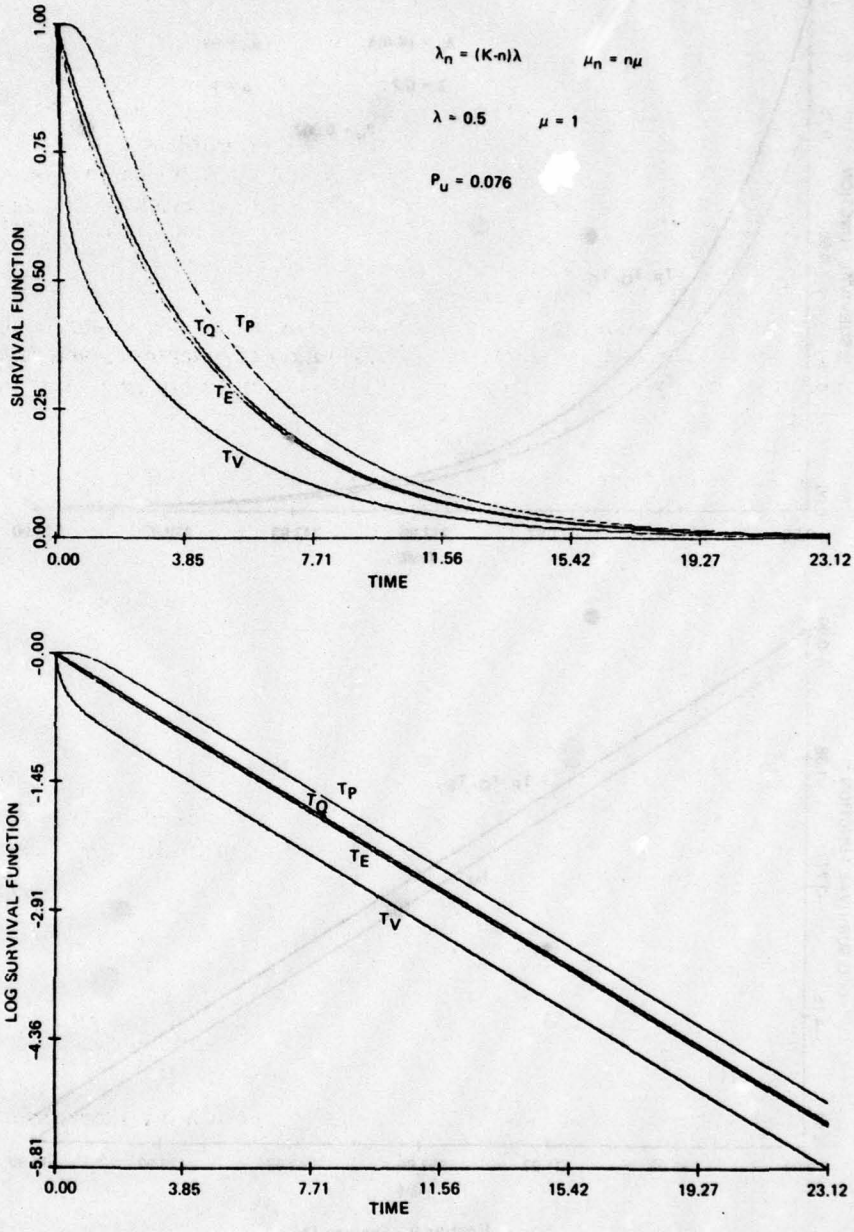


FIGURE 10. System D

IV. EXTENSION TO MULTI-ITEM INVENTORIES

In this section we will indicate how the analysis presented in the previous section for a single-item inventory system may be extended to multi-item inventory systems.

A. System Failure Times

Suppose now that, instead of a single-item inventory, one has R items with independent inventories for each. One might have, for example, a squadron of aircraft, with certain key costly items, such as engines or radars, considered as modules of special concern. For a minimally acceptable number of aircraft N_C to be operating, it is essential that there be $2N_C$ available engines or N_C radars, for instance. At any given time t, component j will have $N_j(t)$ available (in use plus functioning spares), and these availability levels for each item will fluctuate with failures and repairs (or replacements). To extend the ideas of Section III to this multi-item context, one must discuss the failure time of the system, i.e., the random time that elapses between some known initial satisfactory configuration of inventory levels (system state) and a system failure brought about by one of the key item inventory levels falling below the level needed for N_C operating aircraft.

If the failure time for the inventory level of type 1, 2, ..., R is $T_1, T_2, \ldots, T_R$, respectively, then the system failure time is

$$T = \min(T_1, T_2, \ldots, T_R).$$

The failure time distribution for the system will be described by its survival function $\bar{F}_T(\tau)$, and one clearly has, from the independence of the item inventories,

(47) $$P[T > \tau] = P[T_1 > \tau]\, P[T_2 > \tau] \ldots P[T_R > \tau],$$

i.e.,

(48) $$\bar{F}_T(\tau) = \bar{F}_{T_1}(\tau)\, \bar{F}_{T_2}(\tau) \ldots \bar{F}_{T_R}(\tau).$$

Just as one can speak (as in Section II) of four basic failure times of interest for each item inventory, one can now speak of four corresponding failure times for the system, characterized by the following initial conditions:

System Failure Time from the Perfect State

Here, at $t = 0$, all inventories are perfect so that

(49) $$\bar{F}_P(\tau) = \bar{F}_{P1}(\tau)\, \bar{F}_{P2}(\tau) \ldots \bar{F}_{PR}(\tau),$$

where $\bar{F}_{Pr}(\tau)$ is the survival function for the failure time of the r th inventory from the perfect state.

Ergodic Failure Time for the System

At $t = 0$, each inventory is in its steady (ergodic) state, conditional on the inventory level being acceptable, i.e., the initial state for each inventory is precisely that for the ergodic failure time of that inventory. It follows again, from independence, that

(50) $$\bar{F}_E(\tau) = \bar{F}_{E1}(\tau)\, \bar{F}_{E2}(\tau) \ldots \bar{F}_{ER}(\tau).$$

Quasi-Stationary Failure Time for the System

At $t = 0$, each inventory is in its steady state, conditioned on the inventory state being acceptable and the inventory having been acceptable for as long as anyone can remember. Independence then implies that

$$\bar{F}_Q(\tau) = \bar{F}_{Q1}(\tau)\,\bar{F}_{Q2}(\tau)\,\ldots\,\bar{F}_{QR}(\tau). \tag{51}$$

Post-recovery Failure Time

Here, it is known that the system has had a long history and has just had a recovery, i.e., one of the inventory subsystems has just recovered, and all other item subsystems are working. Then the system survival function $\bar{F}_V(\tau)$ may be seen to be

$$\begin{aligned}\bar{F}_V(\tau) = {} & \theta_1 \bar{F}_{V1}(\tau)\,\bar{F}_{E2}(\tau)\,\ldots\,\bar{F}_{ER}(\tau) \\ & + \theta_2 \bar{F}_{E1}(\tau)\,\bar{F}_{V2}(\tau)\,\ldots\,\bar{F}_{ER}(\tau) \\ & + \ldots\ldots \\ & + \theta_R \bar{F}_{E1}(\tau)\,\bar{F}_{E2}(\tau)\,\ldots\,\bar{F}_{VR}(\tau),\end{aligned} \tag{52}$$

where $\bar{F}_{Vr}(\tau)$ is the post-recovery survival function for the type-r inventory subsystem alone. The parameter θ_r is the relative long-run frequency of system failures due to subsystem failures of type r *so that* $\theta_r \geq 0$, $\sum_1^R \theta_r = 1$.

For birth-death models used to describe the individual item inventories, the long-run failure frequency of the type-r inventory is (see Section III for these birth-death models):

$$h_r = e^{(r)}_{n_r^*}\,\lambda^{(r)}_{n_r^*} \tag{53}$$

and

$$\theta_r = \frac{h_r}{\sum_1^R h_r}, \tag{54}$$

where n_r^* is the critical level for component type r, $e_m^{(r)}$ is the steady-state (ergodic) probability that m items are failed, and $\lambda_m^{(r)}$ is the repair rate for level m, all as described in Section III.

B. Properties of the System Failure Times and their Interrelation

It will be seen from equations (49), (50), (51), and (52) that the structure of the different system failure time distributions and their interrelation is almost identical to that for the single-item inventory subsystems described in Section II. We see specifically that

(a) $\bar{F}_P(\tau)$ is log-concave;

$\bar{F}_Q(\tau)$ is log-linear;

$\bar{F}_E(\tau)$ is log-convex;

$\bar{F}_V(\tau)$ is log-convex .

(b) $\bar{F}_P(\tau) \geq \bar{F}_Q(\tau) \geq \bar{F}_E(\tau) \geq \bar{F}_V(\tau)$; i.e., the system failure time survival functions are ordered precisely as before. Consequently,

(c) $\bar{T}_P \geq \bar{T}_Q \geq \bar{T}_E \geq \bar{T}_V$;

that is, the mean failure times have the same order.

(d) When the inventory subsystems are highly reliable, $\bar{F}_{Pr}(\tau)$, $\bar{F}_{Qr}(\tau)$, and $\bar{F}_{Er}(\tau)$ virtually coincide for each item r, and all are exponential. Consequently, the *system* failure time survival functions $\bar{F}_P(\tau)$, $\bar{F}_Q(\tau)$, and $\bar{F}_E(\tau)$ *will coincide and will be exponential.*

(e) Because of the exponentiality, the system may then be described by a single parameter, its failure rate λ, with

$$\lambda = \sum_{r=1}^{R} \lambda_r$$

and

$$\bar{T} = \left[\sum_{r=1}^{R} \bar{T}_r^{-1}\right]^{-1}$$

where λ_r is the failure rate of inventory system r, and $\bar{T}_r$ is the mean failure time of subsystem r. (Here, $\bar{T}_r = \bar{T}_{Qr} \approx \bar{T}_{Er} \approx \bar{T}_{Pr}$.)

(f) The system post-recovery failure time survival function will then lie below the other survival functions to the extent to which there is *jitter* in the individual item inventories. The jitter factor for each subsystem may be defined to be

$$J_r = \lim_{\tau \to \infty} \frac{\bar{F}_{Qr}(\tau)}{\bar{F}_{Vr}(\tau)}.$$

It then follows at once, from the multiplicative structure of (49), (50), (51), and (52), that the system jitter factor

$$J = \lim_{\tau \to \infty} \frac{\bar{F}_Q(\tau)}{\bar{F}_V(\tau)} = J_1 J_2 \ldots J_R.$$

V. AN EXAMPLE ILLUSTRATING THE IMPORTANCE OF DYNAMIC INFORMATION

In this section an example is presented showing the limitations of static analysis and the corresponding need for an understanding of the dynamic system behavior. The example is simplistic, but it does clearly state the case for dynamic over static analysis.

Consider a single-item system that can be modeled as a birth-death process. Using the notation of Section IIIC, suppose that λ_n = failure rate = $(K - n)\lambda$, and μ_n = repair rate = μ. Assume that $K = 10$ and $N_C = 3$; that is, there are ten identical items in the system, and at least seven of them must be working for the system to be in an acceptable state. Assume for the current system that $\lambda = 0.1$ and $\mu = 1$.

The static analysis of this system yields an ergodic probability of being in the unacceptable state, P_u, of 0.223. The dynamic analysis will provide distribution information on the four failure times previously defined. For this example, we will restrict our attention to just the post-recovery failure time or sojourn time on the acceptable region. The expected sojourn time

on the acceptable region is 7.18 time units, while the expected sojourn time on the unacceptable region is 2.06 time units. This implies that the system will alternate between being in the satisfactory state and being in the unsatisfactory state; on average it will spend 7.18 time units in the satisfactory state followed by 2.06 time units in the unsatisfactory state.

Now suppose there are two alternative and exclusive options for improving the system's performance. The first option would increase the repair rate from $\mu = 1$ to $\mu = 2$. The second option would decrease the failure rate by cutting λ from 0.1 to 0.05. In comparing these two options, the static analysis would find that each option results in $P_u = 0.024$. Hence, based on the static analysis, we are indifferent between the two options, assuming that they are comparable in terms of implementation and cost.

However, the two options result in quite distinct system behavior, as can only be seen from the dynamic analysis. In terms of the sojourn times, the first option which increases the repair rate will have, on average, a sojourn time of 28.81 time units in the acceptable state, followed by 0.70 time units in the unacceptable state. This is to be contrasted with the second option, for which the respective expected sojourn times are 57.62 time units and 1.39 time units. In each instance, these sojourn times can be considered to be exponentially distributed. Hence, while the two options yield the same availability levels, the second option results in average sojourn times which are twice as long as for the first option. For instance, if the system has just failed, then the average time until recovery to the acceptable state is twice as long for option two than for option one; that is, the system will not be acceptable for 1.39 time units as opposed to 0.70 time units.

The essence of this result carries over for perhaps more meaningful models. For instance, for each of the birth-death systems considered in Section IIIC, an example with behavior and results similar to the one just given can be constructed. Indeed, for more complex models the need for the dynamic analysis is generally greater due to the inherent complexity of the system's behavior.

The intent of this example has been to illustrate the differences in information from static analysis and dynamic analysis. The system considered is admittedly very simplistic, but the conclusions from the example are equally applicable to more complex and realistic systems. Whereas the static analysis is unable to distinguish between two potential system configurations, the dynamic analysis shows that there is a distinct difference. The example shows that the static analysis is limited, and must be supplemented, if not replaced, by dynamic analysis if one is to understand the full implications of various design trade-offs. Of course, in any actual application the cost differences between the various alternatives must also be considered. For instance, if there is a fixed cost budget, design alternatives are to be chosen so as to maximize the performance of the system subject to the budget constraint. In order to reflect completely the system performance for a specific configuration, the dynamic analysis is necessary.

REFERENCES

[1] Keilson, J., "Systems of Independent Markov Components and their Transient Behavior," *Reliability and Fault Tree Analysis*, SIAM, Philadelphia (1975) pp. 351-364.

[2] Keilson, J., *Markov Chain Models — Rarity and Exponentiality*, Monograph, CSS 74-01, September 1974, to be published by Springer-Verlag, New York, Spring 1979.

[3] Keilson, J., "Log-Concavity and Log-Convexity in Passage Time Densities of Diffusion and Birth-Death Processes," *Journal of Applied Probability 8*, pp. 391-398 (1971).

[4] Keilson, J., and F. W. Steutel, "Mixtures of Distributions, Moment Inequalities and Measures of Exponentiality and Normality," *Annals of Probability 2*, No. 1, pp. 112-130 (1974).

[5] Ross, H., and W. Huisjes, "Log-Concavity of Passage-Time Densities in Birth-Death Processes: Review and Illustrations," Report CSS 71-10, Center for System Science, The University of Rochester, Rochester, New York (December, 1971).

A THREE-ECHELON, MULTI-ITEM MODEL FOR RECOVERABLE ITEMS*

J. A. Muckstadt

School of Operations Research and Industrial Engineering
College of Engineering
Cornell University
Ithaca, New York

ABSTRACT

The main objective of this paper is to develop a mathematical model for a particular type of three-echelon inventory system. The proposed model is being used by the Air Force to evaluate inventory investment requirements for alternative logistic structures. The system we will model consists of a group of locations, called bases, and a central depot. The items of concern in our analysis are called recoverable items, that is, items that can be repaired when they fail. Furthermore, each item has a modular or hierarchical design. Briefly, the model is used to determine the stock levels at each location for each item so as to achieve optimum inventory-system performance for a given level of investment. An algorithm for the computation of stock levels for each item and location is developed and illustrated. Some of the ways the model can be used are illustrated with Air Force data.

INTRODUCTION

The main objective of this paper is to develop a mathematical model for a particular type of three-echelon inventory system. The model is being used by the Air Force to evaluate inventory investment requirements for alternative logistics structures. The system we will study consists of a group of locations, called bases, and a central depot. The items of concern in our analysis are called recoverable items, that is, items that can be repaired when they fail. Briefly, the model will be used to determine what the stock levels should be at each location for each item so as to achieve optimum inventory-system performance for a given level of investment.

We assume that each item in the system has a hierarchical or modular design. By a hierarchically designed recoverable item we mean one that has components which are also recoverable items. In the Air Force context, when an aircraft fails, a recoverable assembly is often found to be faulty. It is removed from the aircraft and replaced by a serviceable assembly of the same type. If this failed assembly has a hierarchical design, it may be taken to a shop on the base where a faulty recoverable component may be identified. To return the failed assembly to a serviceable condition requires the removing and replacing of the defective recoverable component. For example, many avionics-system components on newer aircraft, such as the

*This research was supported in part by the Office of Naval Research under contract N00014-75-C-1172 Task NR042-335

radar-target digital processor on the Air Force's new F-15 fighter, have this type of hierarchical design. This assembly has many recoverable components which are, for the most part, integrated-circuit boards. Recognition and description of the hierarchical relationship among the recoverable items is a major element of the model we will develop. As we will see, the relationship between the stocking of inventories of assemblies and recoverable components at one echelon and the performance at that and lower echelons is demonstrated through the equations for the average resupply time for each base. These equations are the backbone of the model.

The three-echelon system, as we have stated, consists of a group of bases and a depot. Each base is capable of performing only certain types of maintenance. Some bases, which form the second echelon, have extensive repair centers, called maintenance centers, collocated with them. The remaining bases, called operating bases, perform only a minimal amount of maintenance. These bases comprise the third and lowest echelon. By definition, an operating base does not have a collocated maintenance center. Lastly, the depot, the first echelon, has the capability to perform all types of repair.

Each customer demand occurs at a base and is always a requisition for a serviceable assembly. Furthermore, we assume that each customer demand is triggered by the failure of an assembly. The failed assembly is then repaired either at a maintenance center or at the depot. The location at which repair takes place depends only on the nature of the failure. Maintenance centers repair assemblies; they use diagnostic equipment located in shops to isolate the assemblies' defective components. In some instances, these components may also be repaired at the maintenance center, although in most real applications they are normally repaired at the depot. Since repair of assemblies is performed only at maintenance centers and at the depot, components need to be stocked only at the depot and at bases having a collocated maintenance center. Assemblies, on the other hand, can be stocked at all locations.

Customer demands for serviceable assemblies are always satisfied by an organization at each base called base supply. If a serviceable spare assembly is immediately available from base supply, the customer demand is satisfied immediately. On the other hand, if no serviceable stock is on hand, the assembly is placed in a backorder status and the satisfaction of the customer demand is delayed. As we have described, the failed assembly is then either repaired at the base or sent to a higher echelon to be repaired.

Correspondingly, resupply of base supply can occur in one of two ways. If a failed assembly is repaired at a maintenance center, resupply occurs from the maintenance center; if the assembly is repaired at the depot, then resupply will occur from the depot. In either case, the organization that resupplies the base supply activity does so by exchanging a serviceable part for a failed part on a one-for-one basis. The resupply time—that is, the time it takes to replace an assembly demanded from base supply with a serviceable one—depends on the source of resupply. For example, at a base having a collocated maintenance center, the average resupply time for base supply for an assembly equals the average repair time when the assembly is repaired at the maintenance center. The average repair time for the assembly clearly depends on the availability of the components needed to accomplish the repair. If adequate component stocks are on hand, repair will be completed with minimal delay. On the other hand, if repair of the assembly takes place at the depot, the average resupply time for base supply equals the average depot-to-base shipping time plus the expected waiting time before a serviceable assembly is available for shipment to the base. This expected waiting time depends on the depot stock level for the assembly.

Additionally, in this three-echelon system we will assume the structure of the supply and maintenance system for an assembly family—an assembly and its subordinate components—can be represented by a tree, as displayed in Figure 1. Specifically, we assume that the set of bases can be partitioned into a collection of mutually exclusive and collectively exhaustive sets. Each set, which has exactly one maintenance center and a collection of operating bases logistically supported by the maintenance center, is called a Consolidated Support Family. Each operating base is assumed to receive all maintenance-center-level resupply from the maintenance center in its Consolidated Support Family.

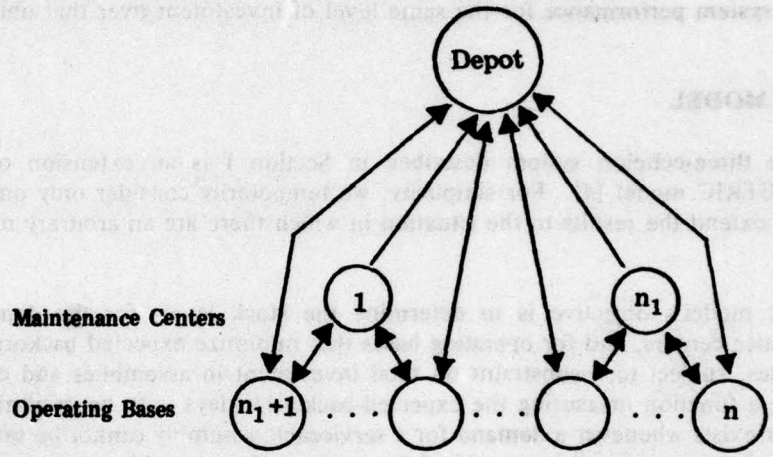


FIGURE 1. The supply and maintenance system for an assembly family

Bases at which a maintenance center is located may or may not have customers requesting serviceable assemblies. If the base has such customers, all requests for spare assemblies are made to the base maintenance center. The maintenance center performs all resupply for all customers at that base and is also a resupply point for all the operating bases in the same Consolidated Support Family. Some assemblies that fail at a base having a maintenance center are repaired at the base, but others may be sent to the depot for repair. We assume, for the sake of simplicity, that any failed assembly sent to a location for repair by a lower-echelon base is not sent on and is repaired there.

In the next section we develop a mathematical model for the system we have described for a single hierarchically designed item. The model recognizes the hierarchical relationship among the recoverable items, and it explicitly accounts for the relationship between the stocking of inventories of assemblies and components at one echelon and the performance at that echelon, as well as at other echelons. An analytic solution is obtained for the three-echelon stockage problem under the steady-state demand assumption. Under this assumption, the solution depends only on the mean resupply times rather than on the resupply time distributions. Mathematical results are stated in terms of the Poisson assumption, but they can be readily extended to cover the compound Poisson case.

The third section contains a description of an algorithm for computing stock levels for the assembly and its components. A method for computing stock levels for systems consisting of a large number of assemblies is presented and illustrated in Section IV.

The model presented in this paper was developed originally to assist Air Force planners in their study of alternatives to the Air Force's current two-echelon (depot-base) logistics-system structure. In the current system, all bases have collocated maintenance centers. The model is being used to assess the differences in stockage requirements between the Air Force's current structure and the three-echelon structure described previously. Some of the ways the model can be used are illustrated in Section V with Air Force data. The illustrations provide some interesting insights into how inventory requirements change when a three-echelon system is operated rather than a two-echelon system.

The final section contains a brief summary and an example that indicates that using relatively sophisticated inventory models, such as the one described in this paper, will significantly improve system performance for the same level of investment over that obtained using simple models.

II. THE MODEL

The three-echelon system described in Section I is an extension of the two-echelon MOD-METRIC model [4]. For simplicity, we temporarily consider only one assembly family. Later we extend the results to the situation in which there are an arbitrary number of assembly families.

The model's objective is to determine the stock levels for the depot, for bases with maintenance centers, and for operating bases that minimize expected backorders for assemblies at all bases, subject to a constraint on total investment in assemblies and components. More precisely, a function measuring the expected-backorder-days is to be minimized. An assembly backorder exists whenever a demand for a serviceable assembly cannot be satisfied by base supply at the base at which the assembly failure occurred. Assembly-resupply delays at the depot or at a maintenance center are measured in the model only insofar as they influence backorders for assemblies at bases; component shortages are also measured indirectly. Observe that a backorder for an assembly at a base indicates that a customer demand is unsatisfied. Since components are only used to repair assemblies, a component backorder only delays repair of the assembly; it does not directly cause a customer's demand to be unsatisfied immediately. Consequently, the impact of assembly and component backorders on customer satisfaction is quite different. We describe the exact nature of the assembly/component interaction in detail later in this section.

Before presenting the mathematical model for this decision problem, we first state the underlying assumptions and then develop the average resupply-time equations for assemblies for all bases. As will be shown, these equations are the backbone of the model. They represent the manner in which assembly and component stock levels interact, and they explicitly state how resupply capability for each echelon depends on the stock levels for all higher echelons. Having established these equations, we next determine the probability distributions describing the number of units in resupply for each location. Using these probabilities, we can then calculate the expected number of assembly backorders outstanding at any time at each base; that is, we can state the model's objective function.

Basic Assumptions

The basic assumptions* underlying the model, in addition to those mentioned earlier, include:

*A complete discussion of these assumptions and their implications is given in Ref. 7.

1. Demand for assemblies at each base is a stationary Poisson process.
2. There is no lateral resupply among bases.*
3. All failed parts are repaired.
4. The probability of a failure of one assembly is independent of failures occurring for other assemblies.
5. Repair times are statistically independent.
6. There is no waiting or batching of items before starting the repair of any item.
7. The echelon at which repair is performed depends only on the complexity of the repair.
8. Each assembly failure repaired at a maintenance center is caused by a failure of at most a single component.

Average-Resupply-Time Equation for Assemblies

After defining some necessary notation, we first derive the average-resupply-time equation for assemblies for each maintenance center and describe in detail the exact nature of the assembly/component interaction. Next, we develop the average-resupply-time equations for assemblies for both the bases having collocated maintenance centers and the operating bases subordinate to the maintenance centers.

Let $N = \{1, \ldots, n_1\}$ denote the set of locations having maintenance centers, and let $N(k)$ denote the set of locations resupplied by $k \in N$; let $M = \{n_1+1, \ldots, n\}$ be the set of operating bases. An index j will refer to an operating base, an index k to a base having a collocated maintenance center, and index 0 will refer to the depot.

Let

$\lambda'_k \triangleq$ expected daily customer demand for assemblies at the base collocated with maintenance center k, $k \in N$;

$\lambda'_j \triangleq$ expected daily customer demand for assemblies at operating base j, $j \in M$;

$w_{vt} \triangleq$ probability that a failed assembly occurring at location v is both repaired and resupplied by location t;

$\lambda_k \triangleq$ expected daily resupply requests for assemblies levied on maintenance center k, $k \in N$.

The expected number of requests for resupply for assemblies levied on maintenance center k equals the expected number of daily assembly failures at the base collocated with maintenance center k plus the expected number of daily resupply requests for assemblies generated by lower-echelon bases supported by maintenance center k. Thus,

$$\lambda_k = \lambda'_k + \sum_{j \in N(k)} w_{jk}\lambda'_j.$$

*This assumption is consistent with Air Force policy for computing recoverable item stock levels and was made for this reason.

Furthermore, let

r'_k ≜ probability that an assembly failure at the base collocated with maintenance center k is repaired at maintenance center k;

r_k ≜ probability that an assembly arrival to maintenance center k is repaired there (see the next paragraph for a discussion of r_k);

B_v ≜ the expected assembly repair-cycle time at location v, measured in days, including repair-time delay for unavailable components, $v = 1, \ldots, n$;

A_{vt} ≜ the expected order-and-ship time between t and v for assemblies measured in days where $t = 0, \ldots, n$, and $v = 1, \ldots, n$;

D ≜ the expected depot repair-cycle time for assemblies measured in days;

s_t ≜ the stock level for the assembly at location t.*

By assumption, all failed assemblies shipped from an operating base to maintenance center k are actually repaired at maintenance center k; however, some assemblies that fail at the base collocated with maintenance center k are sent to the depot for repair. Then the expected number of failed assemblies arriving at maintenance center k each day that are repaired there equals the total number of maintenance-center-k expected daily resupply requests minus the expected number of assembly failures per day occurring at the base collocated with maintenance center k that require depot-level repair. Thus, the probability that an assembly arriving at maintenance center k will actually be repaired there is

$$r_k = \frac{\lambda_k - \lambda'_k(1 - r'_k)}{\lambda_k} = 1 - \frac{\lambda'_k}{\lambda_k}(1 - r'_k) .$$

We are now ready to establish the equation for the average resupply time for assemblies for maintenance center k, which we denote by T_k. The expected assembly resupply time at maintenance center k equals the probability r_k that the assembly will be repaired at maintenance center k times the average maintenance-center-k assembly-repair time B_k plus the probability the assembly will be repaired at the depot $(1 - r_k)$ times the average depot-to-maintenance center k resupply time.

The average depot-to-maintenance center k resupply time equals the average assembly order-and-ship time A_{k0} plus the expected number of days before a serviceable assembly is available at the depot for shipment to the base. This depot delay time can be found from the following formula: expected delay days per demand equals the expected number of assemblies being delayed at any point in time—the expected number of depot backorders—divided by the expected daily depot demand rate for assemblies. Let

$$\lambda = \sum_{t=1}^{n} w_{t0}\lambda'_t ,$$

the expected number of daily demands for assembly resupply placed on the depot. It follows from assumptions 1 and 4 that $p(x|\lambda D)$—the probability that x assemblies are in resupply at the depot, given that the expected demand over the depot resupply cycle is λD—has a Poisson distribution with mean λD. Thus, the expected number of delay days experienced by each assembly resupplied by the depot can be expressed as

*The stock level is defined to be the on-hand plus on-order inventory minus backorders.

$$H(s_0) \triangleq \frac{\text{expected depot backorders given the depot assembly stock level } s_0}{\text{expected daily depot demand for assemblies}}$$

or

$$H(s_0) \triangleq \frac{1}{\lambda} \sum_{x > s_0} (x - s_0) p(x|\lambda D)$$

Combining these observations, we see that the average assembly-resupply time at maintenance center k can be expressed as $T_k = r_k B_k + (1 - r_k)\,[A_{k0} + H(s_0)]$. This equation indicates how the depot stock affects the average resupply time for assemblies at maintenance center k.

But T_k also depends on the component stock levels. The average assembly repair-cycle time B_k at maintenance center k is the sum of two terms. The first reflects the portion of the repair-cycle time related to the operation of the maintenance and transportation systems. In particular, this term represents administrative delay time plus queuing time, plus fault isolation time, plus component remove-and-replace time. It also includes transportation time, if the assembly is sent to maintenance center k for repair by an operating base. Denote this portion of the average assembly repair-cycle time by R_k.

The second term reflects expected delay in completing an assembly's repair due to the shortage of serviceable components. If a particular component is the cause of the assembly's failure and no serviceable component of that type is on hand, then the assembly repair time is lengthened. Consequently, B_k depends on component stock levels both at maintenance center k and at the depot. Let $G_k(s_{1k}, \ldots, s_{mk}; s_{10}, \ldots, s_{m0}) \triangleq G_k$ represent the average delay days per demand in maintenance-center-k assembly repair, given component i stock level s_{ik} at base k and component i depot stock level s_{i0}, where m represents the number of different components in the assembly. We will now develop an explicit expression for G_k.

Recall we assume that if an assembly is repaired at a maintenance center, at most one component needs to be replaced. Then the expected delay in assembly repair time at maintenance center k, given the failure of component i, is the expected number of components of type i at maintenance center k on which delay is being incurred at any point in time—the expected backorders for component i at maintenance center k—divided by the expected component i daily removal or demand rate at maintenance center k. Denote this conditional expected delay by g_{ik}; that is,

$$g_{ik} \triangleq \frac{\text{expected backorders for component } i \text{ at maintenance center } k \text{ at any point in time}}{\text{expected daily removal rate for component } i \text{ at } k},$$

or

$$g_{ik} \triangleq \frac{1}{\lambda_{ik}} \sum_{x > s_{ik}} (x - s_{ik}) p(x|\lambda_{ik} T_{ik}) ,$$

where

$\lambda_{ik} \triangleq$ average number of daily removals of component i at maintenance center k;
$r_{ik} \triangleq$ probability that if component i fails at location k, the component will be repaired at maintenance center k;

B_{ik} ≜ average component i repair time at maintenance center k;
A^i_{k0} ≜ average component i order-and -ship time from the depot to maintenance center k;
D_i ≜ average depot repair-cycle time for component i;
T_{ik} ≜ average resupply time for component i at maintenance center k.

The average-resupply-time equation T_{ik} for component i at maintenance center k equals the probability r_{ik} that the component is repaired at maintenance center k times the average maintenance-center-k repair time B_{ik} for component i, plus the probability the component will be repaired at the depot $(1 - r_{ik})$ times the sum of the depot-to-maintenance center order-and-ship time A^i_{k0} and the expected number of days before a serviceable component is available at the depot for shipment to the maintenance center. We denote this latter delay by $H_{ik}(s_{i0})$, where

$$H_{ik}(s_{i0}) \triangleq \frac{\text{expected number of unsatisfied depot demands for component } i \text{ at any point in time given the depot stock level for component } i \text{ is } s_{i0}}{\text{expected daily depot demand rate for component } i},$$

or

$$H_{ik}(s_{i0}) \triangleq \frac{1}{\theta_i} \sum_{x > s_{i0}} (x - s_{i0}) p(x|\theta_i D_i),$$

and where

$$\theta_i = \sum_{k=1}^{n_1} \lambda_{ik} (1 - r_{ik}),$$

and

$$p(x|\theta_i D_i) = \text{the probability of } x \text{ units of component } i \text{ in depot resupply at any point in time}$$

$$= e^{-\theta_i D_i} \frac{(\theta_i D_i)^x}{x!}.$$

Thus,

$$T_{ik} = r_{ik} B_{ik} + (1 - r_{ik})[A^i_{k0} + H_{ik}(s_{i0})].$$

The probability that an assembly failure repaired at maintenance center k is caused by component i is $\lambda_{ik}/r_k\lambda_k$. Then the expected delay time in repair of an assembly at maintenance center k due to the unavailability of component stock is found by multiplying the conditional delays g_{ik} by $\lambda_{ik}/r_k\lambda_k$ and summing over component types. Thus,

$$G_k = \sum_{i=1}^{m} \frac{\lambda_{ik}}{r_k\lambda_k} g_{ik}.$$

We have now seen that B_k, the average repair time for an assembly at maintenance center k, can be represented as the sum of two terms, R_k and G_k. We therefore have shown that the average resupply time for an assembly at maintenance center k can be represented as

$$T_k = r_k B_k + (1 - r_k)[A_{k0} + H(s_0)]$$
$$= r_k(R_k + G_k) + (1 - r_k)[A_{k0} + H(s_0)].$$

This equation indicates how the depot stock level for the assembly and the maintenance center k and depot component stock levels affect the assembly resupply time at maintenance center k.

We now develop the average-resupply-time equation for base k, $k \in N$. Since base k is physically collocated with the maintenance center, no assembly stock will be allocated exclusively to it. Immediate resupply is assumed to be always available (zero lead time) for customers at the base from the maintenance center if serviceable stock is on hand. From a system's viewpoint there is no advantage to allocating exclusive stock to the base, since all assemblies assigned there, by assumption, would be unavailable for redistribution. This would degrade expected system performance. Since all resupply for customer demands at base k comes from maintenance center k, the average resupply time for an assembly for base k, call it T'_k, equals the expected number of delay days before a serviceable assembly becomes available at maintenance center k. Therefore,

$$T'_k = \sum_{x > s_k} \frac{(x - s_k)p(x|\lambda_k T_k)}{\lambda_k}, \quad k \in N.$$

The average-resupply-time equation for operating base j, call it T_j, can be found by the same method we used to determine T_k. Let us temporarily assume that location j receives resupply from maintenance center k. Then T_j equals r_j, the probability that the failed assembly is repaired at base j, times the average base repair time for the assembly B_j plus the probability w_{jk} that the assembly is repaired at maintenance center k, times the sum of maintenance center k to operating base j order-and-ship time A_{jk} and the expected delay in shipment due to the unavailability of a serviceable assembly at maintenance center k (call this quantity $H_k(s_k)$), plus the probability that the assembly is shipped to the depot for repair w_{j0} times the sum of the depot-to-base j order-and-ship time A_{j0} and the expected delay before a serviceable assembly is available at the depot for shipment to the base $H(s_0)$. In general, let $g(j) \in N$ denote the maintenance center for which $w_{jk} > 0$. Then we may express the average resupply time for operating base j as

$$T_j = r_j B_j + w_{j,g(j)}\,[A_{j,g(j)} + H_{g(j)}(s_{g(j)})] + w_{j0}[A_{j0} + H(s_0)].$$

The average number of days a resupply request for an assembly levied on maintenance center k is delayed before a serviceable assembly becomes available for shipment, given the stock level s_k, was denoted by $H_k(s_k)$. This function is

$$H_k(s_k) \triangleq \frac{\text{expected maintenance center } k \text{ assembly backorders at any point in time given the maintenance center } k \text{ stock level of } s_k}{\text{expected daily assembly demand at maintenance center } k}$$

or

$$H_k(s_k) \triangleq \frac{1}{\lambda_k} \sum_{x > s_k} (x - s_k)p(x|\lambda_k T_k),$$

where $p(x|\lambda_k T_k)$ is the probability of x assemblies in the maintenance-center-k resupply system. In the expression, $p(x|\lambda_k T_k)$ is approximated by a Poisson distribution whose mean is $\lambda_k T_k$.

Mathematical Statement of the Model

The goal of the model is to find the assembly and component stock levels for each location that minimize the system's average number of backorders for customer demands for assemblies outstanding at any point in time, subject to a restriction on inventory investment. For each operating base, that is, for each $j \in M$, we express the average number of outstanding backorders at any time for customer demands for assemblies as

$$\sum_{x>s_j} (x - s_j)p(x|\lambda'_j T_j) ,$$

where s_j represents the stock level for the assembly at base j. Recall that no stock is explicitly allocated to the base at a location having a maintenance center. All stock at base k is under the administrative control of the maintenance center. Thus, the average number of customer backorders for assemblies at base k at any time is

$$\sum_{x\geq 0} xp(x|\lambda'_k T'_k) = \lambda'_k T'_k .$$

Therefore, the objective function for the model is

$$\sum_{j\in M}\left\{\sum_{x>s_j} (x - s_j)p(x|\lambda'_j T_j)\right\} + \sum_{k\in N} \lambda'_k T'_k .$$

Note that the backorder expression for each base depends on its average resupply time.

The inventory-investment constraint in the model states that the system investment in assemblies and components cannot exceed some maximum value. If

c = the unit cost of an assembly,
c_i = the unit cost of component i,
s_{it} = stock level for component i at location t, and
C = the available budget,

the mathematical representation of the investment constraint is

$$c \sum_{t=0}^{n} s_t + \sum_{i=1}^{m} c_i \sum_{t=0}^{n_1} s_{it} \leqslant C .$$

Combining the above, we write the mathematical statement of the model as follows:

(P) $$\min \sum_{j\in M} \left\{ \sum_{x>s_j} (x - s_j)p(x|\lambda'_j T_j)\right\} + \sum_{k\in N} \lambda'_k T'_k$$

subject to $$c \sum_{t=0}^{n} s_t + \sum_{i=1}^{m} c_i \sum_{t=0}^{n_1} s_{it} \leqslant C ,$$

where s_t and s_{it} are nonnegative integers,

$t = 0, \ldots, n$, and $i = 1, \ldots, m$.

We will call this Problem P.

III. AN ALGORITHM FOR DETERMINING STOCK LEVELS

The model's objective function represents the total system backorders existing at any point in time for customer demand for the assembly. As stated in the previous section, the expected backorder expression for the assembly for each operating base, that is, for each $j \in M$, is

$$\sum_{x>s_j} (x - s_j)p(x|\lambda'_j T_j) ,$$

which depends on T_j. But T_j is a function of both depot and maintenance-center assembly and component stock levels. Similarly, the expected backorder expression for customer demands for assemblies for each $k \in N$ depends on depot assembly and component stock levels as well as on maintenance-center stock levels for components. Consequently, Problem P is not a separable programming problem. Furthermore, the objective function need not be convex.

The strategy we employ to solve Problem P circumvents these difficulties. Specifically, we will solve a finite sequence of subproblems, each corresponding to a fixed investment in assemblies. For a fixed total budget C, it is possible to purchase either 0,1,..., or Q assemblies, where Q is the greatest integer less than or equal to C/c. The proposed algorithm requires evaluating the solution—at least implicitly— to $Q + 1$ subproblems, one for each possible investment in assemblies. Each subproblem can be stated as follows:

$$\min_{s_{it}} \left\{ \left\{ \min_{s_t} \sum_{j \in M} \sum_{x>s_j} (x - s_j)p(x|\lambda'_j T_j) + \sum_{k \in N} \lambda'_k T'_k : \right. \right.$$

s_{it} is fixed for all i and t (thereby establishing the component delay in T_k), and s_t and s_{it} are nonnegative integers;

$$\sum_{t=0}^{n} s_t = \bar{N} ,$$

where $\bar{N}$ represents the number of assemblies available for distribution, and $\bar{N}$ is the greatest integer less than or equal to

$$\left. \left. \frac{C - \sum_{i=1}^{m} c_i \sum_{t=0}^{n_i} s_{it}}{c} \right\} \right\}$$

Consequently, each subproblem can be two partitioned into two parts, one corresponding to components and the other to assemblies. The first part establishes the manner in which a limited budget $(C - c\bar{N})$ is allocated among the m components. Once a specific allocation of components to the depot and maintenance centers has been determined, the expected delay in assembly repair time due to components is known. This in turn affects the resupply time and ultimately the expected backorders for assemblies at each base. The optimal allocation of the $\bar{N}$ assemblies among the bases and depot— which corresponds to the second portion of the above problem— is obtained from the expected delay in assembly repair time at each maintenance center.

Suppose $U \triangleq C - c\bar{N}$ dollars are available for investment in components. How should it be allocated among the m components? Clearly, we should make the investment so that the total of expected customer backorders for assemblies is reduced by the greatest amount. If all n_1 Consolidated Support Families are identical, it is not hard to show that this corresponds to

an allocation in which the stock levels are selected so that total weighted expected delay in assembly repair due to components is minimized, where the weights reflect the expected number of daily assembly failures repaired at a maintenance center. Although it is only an approximation in cases where the Consolidated Support Families are not identical, we will use this objective to determine the allocation of the available U dollars among the components for each subproblem. A considerable amount of experimentation was done by the Air Force Logistics Command using this type of approximation in the MOD-METRIC model [4]. The approximation produced the optimal allocation in all cases. Thus, the component stock levels in the $Q + 1$ subproblems are obtained by solving the following problem, called Problem P1:

$$(P1) \quad \min \sum_{k \in N} r_k \lambda_k G_k = \min \sum_{k \in N} \sum_{i} \sum_{x > s_{ik}} (x - s_{ik}) p(x | \lambda_{ik} T_{ik})$$

$$\text{subject to} \sum_{i=1}^{m} \left(c_i s_{i0} + \sum_{k=1}^{n_1} c_i s_{ik} \right) \leqslant U ,$$

where s_{it} is a nonnegative integer.

Observe that minimizing the total weighted expected delay due to components is equivalent to minimizing total component backorders. The solution to this two-echelon component problem can be easily obtained using the method described in either Ref. 3 or Ref. 5.

When the component stock levels have been established, we must then determine the optimal method for allocating the $\bar{N}$ assemblies among the depot and bases. In particular, we must solve the following problem, called Problem P2:

$$(P2) \quad \min \sum_{j \in M} \sum_{x > s_j} (x - s_j) p(x | \lambda'_j T_j) + \sum_{k \in N} \lambda'_k T'_k$$

$$\text{subject to} \sum_{i=0} s_i = \bar{N} ,$$

where s_i is a nonnegative integer.

Due to the interaction of stock levels among echelons, Problem P2 is neither convex nor separable. We therefore employ a simple partitioning procedure to obtain its solution. The algorithm for solving this three-echelon problem is based on the system's nested-tree structure, as displayed earlier in Figure 1. The algorithm works up the tree by solving a sequence of independent two-echelon subproblems, one set of problems for each Consolidated Support Family; the solutions to these problems are then combined in an appropriate way to solve Problem P2. We now discuss the algorithm for solving Problem $P2$ in detail.

Suppose the depot stock level is fixed at s_0, and assume that a total of $\bar{N}_k$ assemblies are available for allocation to all bases in Consolidated Support Family k. Then the optimal allocation of the $\bar{N}_k$ assemblies among the bases can be found by solving the following problem, called Problem P3:

$$(P3) \quad B_k(\bar{N}_k; s_0) \triangleq \min \lambda'_k T'_k + \sum_{j \in N(k)} \sum_{x > s_j} (x - s_j) p(x | \lambda'_j T_j)$$

subject to s_0 fixed,

$$\sum_{j \in N(k)} s_j + s_k = \bar{N}_k , \text{ and}$$

s_j a nonnegative integer, $j \in N(k)$, $s_k \in R_k$,

where R_k represents a set whose elements are the candidate values for s_k. This problem may not be convex and is not separable. To obtain its solution we solve the subproblems

$$h(s_k, s_0) \triangleq \min \sum_{j \in N(k)} \sum_{x > s_j} (x - s_j) p(x|\lambda'_j T_j)$$

subject to s_0 and s_k fixed,

$$\sum_{j \in N(k)} s_j = \bar{N}_k - s_k \text{ , and}$$

s_j a nonnegative integer,

via marginal analysis (valid because of the convexity of the objective function). Then the solution to Problem P3 is found by solving

$$\min_{s_k \in R_K} h(s_k, s_0) + \lambda'_k T'_k .$$

Since the optimal value of $\bar{N}_k$ is unknown, Problem P3 is solved for all values of $\bar{N}_k \in \bar{R}_k$, where $\bar{R}_k$ represents the set of possible total family k stock levels.

To solve Problem P2 we use the solutions obtained for each Consolidated Support Family. More specifically, to solve P2 we solve Problem P4:

(P4)
$$B(\bar{N}) \triangleq \min \sum_{k=1}^{n_1} B_k(\bar{N}_k; s_0)$$

$$\text{subject to } \sum_{k=1}^{n_1} \bar{N}_k + s_0 = \bar{N} ,$$

$$\bar{N}_k \in \bar{R}_k, \; s_0 \in R_0 ,$$

where R_0 represents the set of candidate depot assembly stock levels. A dynamic programming algorithm is used to compute the optimal solution.

The amount of effort required to solve Problems P3 and P4 depends on the cardinality of the sets R_0, R_k, and $\bar{R}_k$. Fortunately, the number of stock levels that need to be explicitly considered for any location or Consolidated Support Family is generally not large. This is chiefly due to the nature of the functions $H(s_0)$ and $H_k(s_k)$, which rise very sharply for stock levels below the mean demand, and approach 0 rapidly for stock levels above the mean.* Experiments [4] on similar problems indicate that the cardinality of the R_0 and R_k sets should rarely exceed 10.

To find $\bar{R}_k$, we may first compute the total expected daily removals for family k, call it $\bar{\lambda}_k$. An estimate of the average family k resupply time $\bar{T}_k$ is found by weighting the expected resupply times for each location in the family by the proportion of family k daily demand occurring at that location and then summing these quantities over locations. An estimate of the depot and base k optimal stock levels obtained, for example, using the method described in Ref. 5, is employed to estimate the value of T_j and T_k used in the averaging. Using these values, we solve Problem P5:

*An illustration of this fact is given in Ref. 4, p. 479.

(P5)
$$\min \sum_{k=1}^{n_1} \sum_{x>\bar{s}_k} (x - \bar{s}_k) p(x|\bar{\lambda}_k \bar{T}_k)$$

$$\text{subject to } \sum_{k=1}^{n_1} \bar{s}_k = \bar{N} - \bar{s}_0, \text{ and}$$

$$\bar{s}_k \text{ is a nonnegative integer,}$$

where $\bar{s}_0$ is the estimate of the optimal depot stock level. Marginal analysis is used to obtain the optimal solution, since the objective function is convex. $\bar{R}_k$ is constructed based on the estimate $\bar{s}_k$. The minimum element of $\bar{R}_k$ can be set at $\max\{a\bar{s}_k, \bar{s}_k - b\}$ and the largest value at $\min\{c\bar{s}_k, \bar{s}_k + d\}$. The values of a, b, c, and d can be selected as a function of the size of $\bar{s}_k$. For larger values of $\bar{s}_k$, the range should be larger. Limited computational experience on a similar problem using this technique has shown that a maximum cardinality of 15 for $\bar{R}_k$ is adequate [6]. However, the best method for determining R_0, R_k, and $\bar{R}_k$ remains an open question.

Combining the above observations, we can state a basic algorithm for determining item stock levels:

INITIALIZATION STEP: Establish upper- and lower-bound constraints on assembly investment. Let u and l represent these upper and lower limits and let z' represent the best known objective function value. Set $z' = \infty$ and $U = C - l$. Assume C is an integer multiple of c.

STEP 1: Solve Problem P6:

(P6)
$$\min \sum_k r_k \lambda_k G_k$$

$$\text{subject to } \sum_{i=1}^{m} \left(c_i s_{i0} + \sum_{k=1}^{n_1} c_i s_{ik} \right) \leq U,$$

$$s_{ik} \text{ is a nonnegative integer.}$$

STEP 2: Solve Problem P7:

(P7)
$$\min z = \sum_{k=1}^{n_1} \lambda'_k T'_k + \sum_{j \in M} \left\{ \sum_{x>s_j} (x - s_j) p(x|\lambda'_j T_j) \right\}$$

$$\text{subject to } \sum_{i=0}^{n} c s_i = C - U,$$

$$T_j \text{ and } T'_k \text{ are calculated using the stock levels computed in Step 1, and } s_i \text{ is a nonnegative integer.}$$

STEP 3: If $z \geqslant z'$, go to Step 4; otherwise, set $z' = z$ and retain the corresponding stock levels as the incumbent stock levels. Go to Step 4.

STEP 4: Decrement U by c. If $C - U > u$, stop; otherwise, return to Step 1.

The algorithm outlined above suggests a rather tedious method for establishing the optimal investment level in assemblies and components. We will now see that the number of assembly budgets that need to be explicitly examined is generally quite small. First observe that when values of u and l are selected one should consider the marginal impact of investment in components on expected backorders for customer demands for assemblies. The marginal impact is negligible when investment in components is large; on the other hand, assembly resupply times are increased substantially, thereby increasing total assembly system backorders, when the investment in components is relatively low. Roughly stated, we would like to allocate the available budget C in such a way that the marginal reduction in backorders for customer demands for assemblies per dollar invested in components equals the marginal reduction per dollar invested in the assembly. The values of u and l should reflect this goal.

It is easy to obtain an estimate of the optimal total component investment. Suppose we estimate component and assembly depot stock levels, perhaps using the method described in Ref. 5. The total cost of this investment can be determined and subtracted from the available budget C. We next assume that all of the n_1 Consolidated Support Families have the same demand rates for assemblies and the same number of operating bases and that all demand for assemblies takes place at the corresponding base collocated with a maintenance center. Then a crude estimate of the optimal investment in components for each $k \in N$ corresponds to the investment level for which the partial derivative of maintenance center k's average resupply time with respect to dollar investment in components at the maintenance center k equals $1/c$. We can then easily estimate the optimal total system investment in components by multiplying the Consolidated Support Family estimate by n_1 and adding to this value the estimated required depot component investment.

Once u and l have been established, we simplify the search for the optimal partitioning of the budget by exploiting the apparent strict quasi-convexity of the total expected customer backorders for assemblies as a function of investment in assemblies. Using the Fibonacci search algorithm, we see that it is necessary to examine only a very small number of assembly investment levels explicity. For example, if $Q = 600$, only 13 problems need to be solved explicitly. Each problem requires the solution of two subproblems. The first subproblem has the form of Problem P6 in Step 1 of the algorithm, and the second subproblem has the form of Problem P7 in Step 2. The value of U, of course, corresponds to a specific total investment in assemblies.

Figure 2 displays the results of applying the proposed algorithm to one assembly/component family for the Air Force's F-15 aircraft. The graph relates total expected customer backorders for assemblies to the proportion of the total system budget invested in assemblies. As indicated on the graph, approximately two-thirds of the total budget should be allocated to the assembly. Investing either a greater or lesser proportion of the total budget in the assembly increases total backorders. A substantial misallocation of the available budget can seriously degrade system performance. For example, investing one-half rather than two-thirds of the budget in the assembly causes expected customer backorders for assemblies to double.

IV. MULTIPLE-ASSEMBLY PROBLEMS

We have developed a model of a three-echelon inventory system for one assembly type and its subordinate components. If this model could not be easily extended to multiple-assembly problems, it would be of little practical use. We now demonstrate how it can be extended.

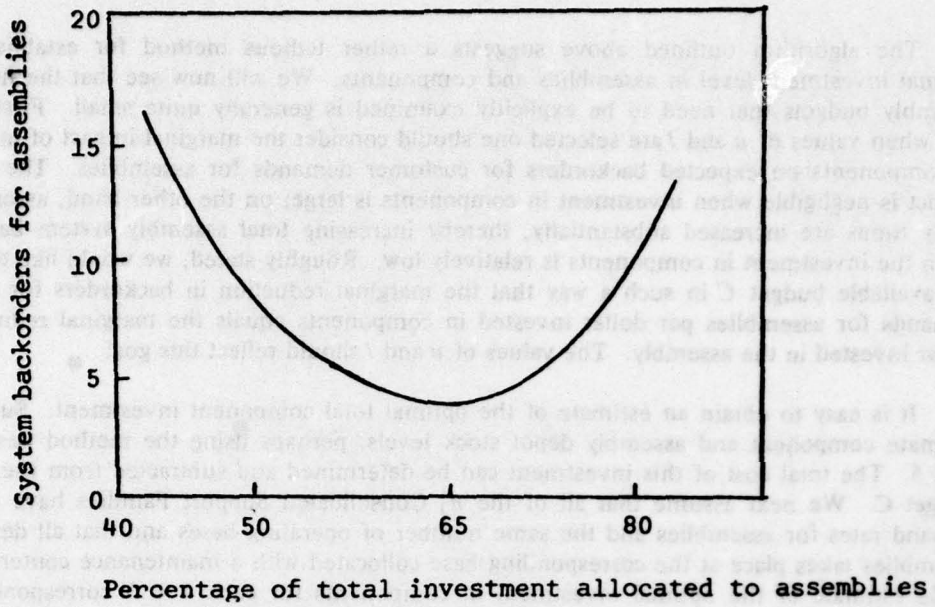


FIGURE 2. System backorders for assemblies as a function of the percentage of total budget allocated to assemblies

In practice, Problem P is solved for a finite number of budgets $C_1^i, C_2^i, \ldots, C_{q_i}^i$ for each assembly family i. The number of budgets explicitly examined q_i depends, in practice, on the expected assembly failure rate. Using the data obtained when solving these q_i problems, it is possible to plot performance vs investment. Figure 3 illustrates this trade-off data for one F-15 assembly family. A piecewise linear function can then be constructed to approximate the entire performance-vs-investment trade-off curve for this assembly family, as shown in Figure 4. If this curve is not convex, then replace it by its greatest convex minorant.

After the convex performance/investment trade-off curves are developed for each assembly family, they are combined to produce a curve relating total customer backorders for all types of assemblies as a function of investment in all assembly and component types. This curve is constructed by applying a simple marginal analysis algorithm. The first point on the system-performance curve corresponds to the total expected customer backorders for an assembly type i when the minimal amount C_1^i is invested in assembly family i.

Let $B_i(C_j^i)$ represent the total expected customer backorders for assembly i, given the investment in assembly family i is C_j^i. Then the first point on the system-performance curve is $\sum_i B_i(C_1^i)$, corresponding to an investment of $\sum_i C_1^i$. Next, compute

$$\Delta_1^i \triangleq \frac{B_i(C_1^i) - B_i(C_2^i)}{C_2^i - C_1^i}$$

for each assembly family. Then Δ_1^i measures the marginal reduction in system customer backorders for assemblies per dollar invested in assembly family i. Suppose $\Delta_1^k = \max \Delta_1^i$. Then the second point on the curve is $\sum_i B_i(C_1^i) - (B_k(C_1^k) - B_k(C_2^k))$, corresponding to an investment of $\sum_i C_1^i + C_2^k - C_1^k$. Now compute

$$\Delta_2^k = \frac{B_k(C_2^k) - B_k(C_3^k)}{C_3^k - C_2^k}$$

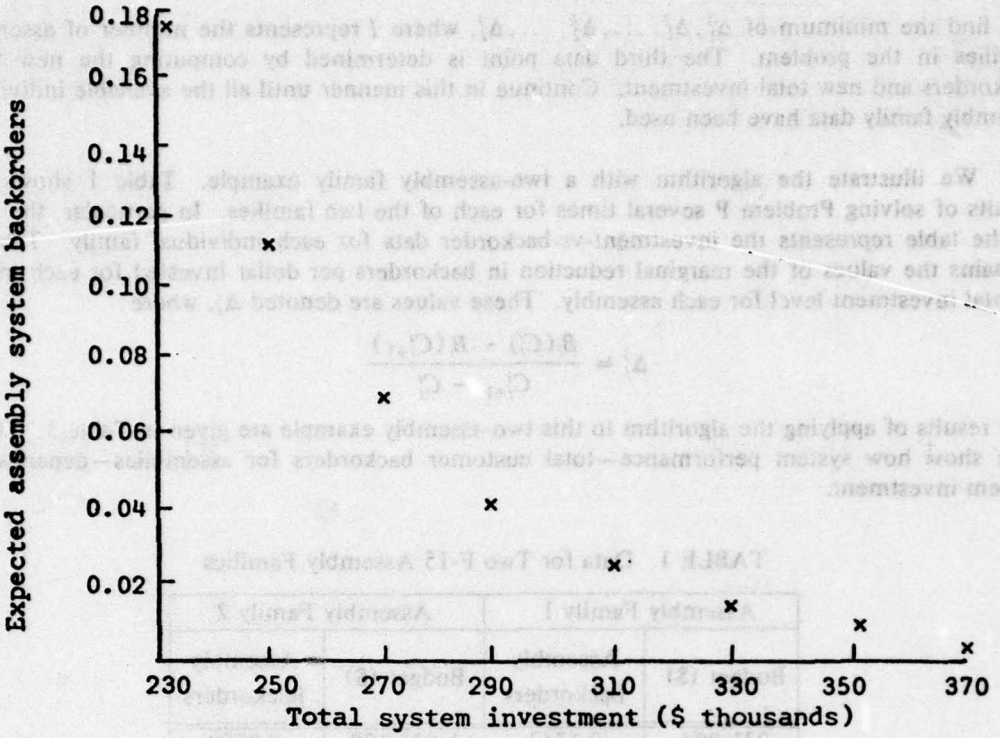


FIGURE 3. System backorders for assemblies as a function of various levels of total investment

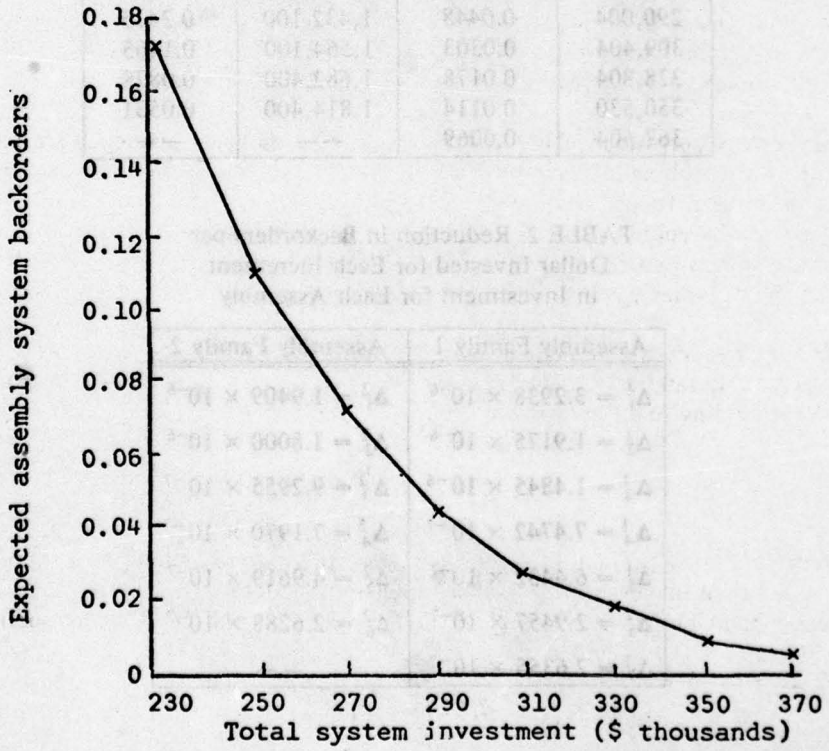


FIGURE 4. Piecewise-linear approximation of performance-vs-investment relationship

and find the minimum of $\Delta_1^1, \Delta_1^2, \ldots, \Delta_2^k, \ldots, \Delta_I^J$, where I represents the number of assembly families in the problem. The third data point is determined by computing the new total backorders and new total investment. Continue in this manner until all the available individual assembly family data have been used.

We illustrate the algorithm with a two-assembly family example. Table 1 shows the results of solving Problem P several times for each of the two families. In particular, the data in the table represents the investment-vs-backorder data for each individual family. Table 2 contains the values of the marginal reduction in backorders per dollar invested for each incremental investment level for each assembly. These values are denoted Δ_j^i, where

$$\Delta_j^i = \frac{B_i(C_j^i) - B_i(C_{j+1}^i)}{C_{j+1}^i - C_j^i}.$$

The results of applying the algorithm to this two-assembly example are given in Table 3. These data show how system performance—total customer backorders for assemblies—depends on system investment.

TABLE 1. Data for Two F-15 Assembly Families

Assembly Family 1		Assembly Family 2	
Budget ($)	Assembly Backorders	Budget ($)	Assembly Backorders
231,804	0.1747	1,036,100	0.8580
251,204	0.1108	1,168,100	0.6018
270,604	0.0736	1,300,100	0.3642
290,004	0.0448	1,432,100	0.2415
309,404	0.0303	1,564,100	0.1465
328,804	0.0178	1,682,400	0.0878
350,530	0.0114	1,814,400	0.0531
367,604	0.0069	--	--

TABLE 2. Reduction in Backorders per Dollar Invested for Each Increment in Investment for Each Assembly

Assembly Family 1	Assembly Family 2
$\Delta_1^1 = 3.2938 \times 10^{-6}$	$\Delta_1^2 = 1.9409 \times 10^{-6}$
$\Delta_2^1 = 1.9175 \times 10^{-6}$	$\Delta_2^2 = 1.8000 \times 10^{-6}$
$\Delta_3^1 = 1.4845 \times 10^{-6}$	$\Delta_3^2 = 9.2955 \times 10^{-7}$
$\Delta_4^1 = 7.4742 \times 10^{-7}$	$\Delta_4^2 = 7.1970 \times 10^{-7}$
$\Delta_5^1 = 6.4432 \times 10^{-7}$	$\Delta_5^2 = 4.9619 \times 10^{-7}$
$\Delta_6^1 = 2.9457 \times 10^{-7}$	$\Delta_6^2 = 2.6288 \times 10^{-7}$
$\Delta_7^1 = 2.6355 \times 10^{-7}$	--

TABLE 3. Backorder and Investment Data for the Combined System (Two Assembly Familes)

Investment ($)	Backorders
1,267,904	1.0327
1,287,304	0.9688
1,419,304	0.7126
1,438,704	0.6754
1,570,704	0.4378
1,590,104	0.4090
1,722,104	0.2863
1,741,504	0.2718
1,873,504	0.1768
1,892,904	0.1643
2,011,204	0.1056
2,032,930	0.0992
2,050,004	0.0947
2,182,004	0.0600

These data can then be used to determine what the individual assembly-family investment levels should be so that a target system budget or performance goal is achieved. For example, suppose planners decide that approximately $1.9 million is available for investment in these two assembly families. The closest tabulated value corresponds to a total investment of $1,892,904. This point in turn corresponds to an investment of $328,804 in the first assembly family and an investment of $1,564,100 in the second assembly family.

The tabulated values can be used in a second way as well. Suppose the planners decide they want no more than 0.1 expected customer backorders attributed to these two assemblies at any point in time. Then a budget of $2,032,930 must be made available for these assembly families, with $350,530 and $1,682,400 budgeted for the first and second families, respectively.

V. AN EXAMPLE OF ANALYSIS USING THE MODEL

As we have mentioned, the main reason for developing the model described in this paper was to assess the impact on inventory investment in recoverable spares of changing the Air Force's current logistics-system structure. To illustrate the use of the model in this regard, we postulated an operating environment involving three bases. In this example, aircraft correspond to customers. We assume there are two squadrons of F-15 aircraft stationed in the first base; one squadron of F-15 aircraft is stationed in each of the other two bases. Flying activities per aircraft are assumed to be the same at each of the three bases. A maintenance center is assumed to exist at each base. The average repair-cycle time is assumed to be 4 days, order and shipping time from the depot to each of the bases is 12 days, and the depot repair-cycle time is 52 days for assemblies.

We selected 18 high-demand assemblies related to the F-15 avionics system as our data base. There are 224 components associated with these 18 assemblies. For these assemblies, 80% to 95% of the malfunctioning items can be repaired at the maintenance center at the base. These avionics-type items were selected for analysis because it would make sense in the real world to consolidate their maintenance at some central location.

We made five sets of computer runs, and for each run we generated a trade-off curve between inventory investment and performance expressed in terms of the expected number of backorders on assemblies at the three bases:

CASE 1 is the base case in which the structure of the logistics system remains the same as the current one. In other words, maintenance is performed at maintenance centers at each location, and other system parameters are the same as those described above.

In CASE 2 it is assumed that repairs will be performed at the largest base, namely Base 1. Thus Base 1 has a maintenance center and Bases 2 and 3 do not. It is assumed to take an average of 4 days to ship defective assemblies to Base 1 from Bases 2 and 3 and to ship serviceables back from Base 1 to Bases 2 and 3. Shipping time from Base 1 to the depot and depot repair times remain the same as in the base case.

In CASE 3 all parameters remain unchanged from Case 2, except that it is assumed that the proportion of repairs that cannot be accomplished at the maintenance center for every assembly has been reduced by 50%. The reduction of 50% is hypothetical and is not based on any engineering study. Under this structure, however, the Air Force has in some actual tests reduced the proportion of items that have to be returned to the depot for repair by roughly this amount.

In CASE 4 we based our calculation on the same proportion of failures being repaired at the depot as in the base case and also used the same system parameters, except the shipping time to Base 1 from Bases 2 and 3 has been reduced from 4 to 2 days. This was done to check the effect of the responsiveness of the transportation system on this alternative type of structure.

Finally, in CASE 5 we assume that the shipping time can be set at 2 days, and it takes only 2 days to ship to the maintenance center. Furthermore, we assume that the proportion of repairs that cannot be accomplished at the maintenance center has been reduced, as in Case 3.

The results are shown in Figure 5. Each curve summarizes the analysis corresponding to each of the five cases described above. Each curve portrays the impact on the performance of the support system as a function of investment in inventory of spares and conditioned on system parameters, as described above. The performance is stated in terms of the number of assembly backorders throughout the system. For example, if we take the base-case trade-off curve, we see that for an investment of $45 million, there will be 9 backorders on the average.

A comparison of Case 2 with the base case shows that introducing a change in the logistics structure of the type described would imply that additional spares requirements of nearly $5 million would be needed to maintain the same level of performance. However, Case 3 results show that if the maintenance capability at the intermediate level could be enhanced to the extent that a greater proportion of defective assemblies could be fixed at the maintenance center, instead of having to be shipped to the depot, additional spares requirements would be minimal. Even without the assumed improvement in the maintenance productivity, if it takes only 2 days to ship serviceable and broken assemblies from operating bases to the maintenance center, then the alternative structure does not require any additional spares investment, as depicted in the comparison between Case 1 and Case 4. It was mentioned earlier that, in the alternative structure, an additional requirement for assemblies may be offset by a reduction in component stockage. When Case 4 was compared to Case 1 at a performance level of 25 assembly backorders, it was found that the composition of stockage had changed as follows: For Case 1, inventory investment for components was $13.8 million, and for the assemblies, $27.5 million.

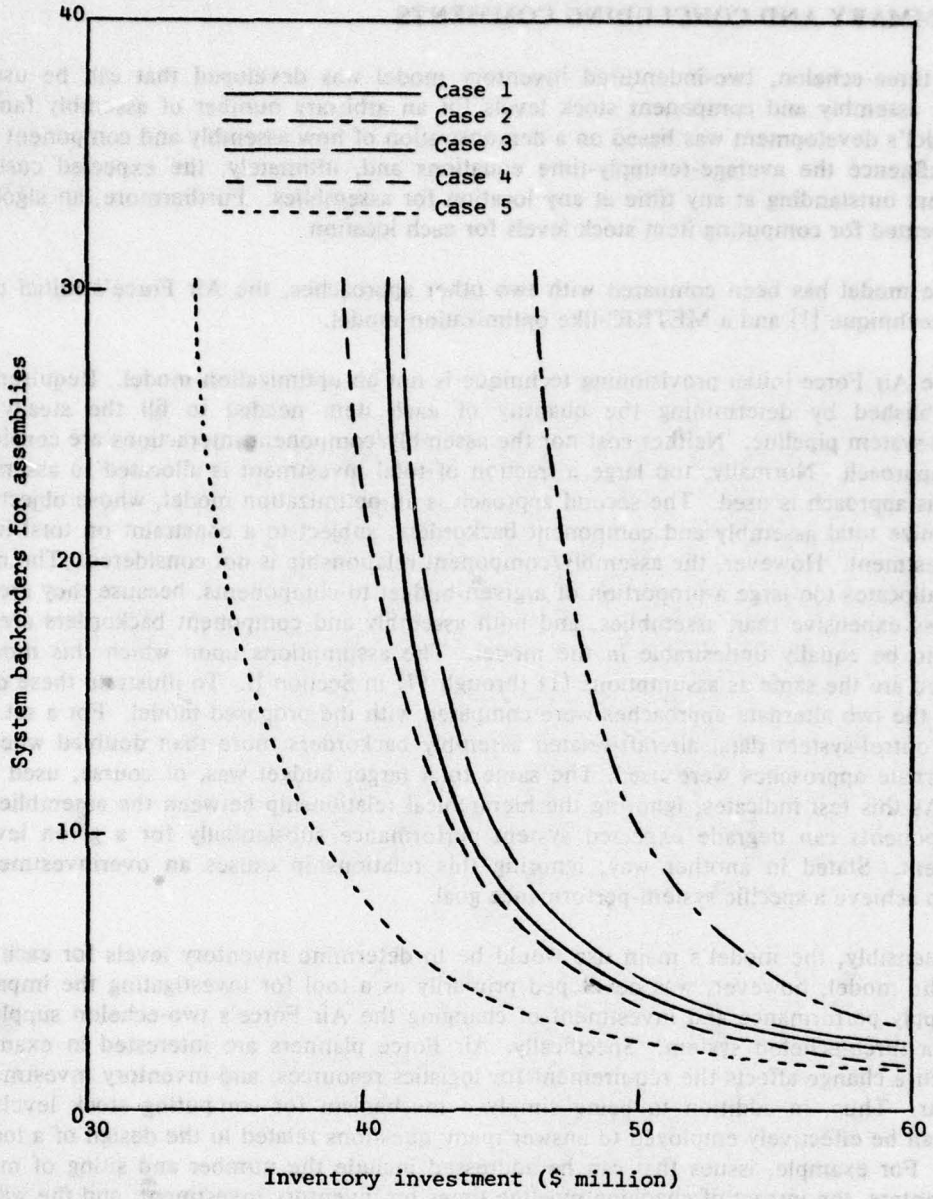


FIGURE 5. Analysis of structures: maintenance centralization vs decentralization

For Case 4, they were $11.3 million and $28.4 million, respectively. Thus, under the hypothesized operating conditions, a saving in component stockage investment more than offset a need for more assemblies.

Finally, Case 5 suggests that if the alternative three-echelon structure can improve maintenance productivity as well as rely on a highly responsive transportation system, economic gains in the area of spares requirements are possible.

VI. SUMMARY AND CONCLUDING COMMENTS

A three-echelon, two-indentured inventory model was developed that can be used to establish assembly and component stock levels for an arbitrary number of assembly families. The model's development was based on a demonstration of how assembly and component stock levels influence the average-resupply-time equations and, ultimately, the expected customer backorders outstanding at any time at any location for assemblies. Furthermore, an algorithm was presented for computing item stock levels for each location.

The model has been compared with two other approaches, the Air Force's initial provisioning technique [1] and a METRIC-like optimization model.

The Air Force initial provisioning technique is not an optimization model. Requirements are established by determining the quantity of each item needed to fill the steady-state resupply-system pipeline. Neither cost nor the assembly/component interactions are considered in this approach. Normally, too large a fraction of total investment is allocated to assemblies when this approach is used. The second approach is an optimization model, whose objective is to minimize total assembly and component backorders, subject to a constraint on total inventory investment. However, the assembly/component relationship is not considered. The model usually allocates too large a proportion of a given budget to components, because they are generally less expensive than assemblies, and both assembly and component backorders are considered to be equally undesirable in the model. The assumptions upon which this model is developed are the same as assumptions (1) through (7) in Section II. To illustrate these observations, the two alternate approaches were compared with the proposed model. For a set of F-15 fire-control-system data, aircraft-related assembly backorders more than doubled when the two alternate approaches were used. The same total target budget was, of course, used in all cases. As this test indicates, ignoring the hierarchical relationship between the assemblies and its components can degrade expected system performance substantially for a given level of investment. Stated in another way, ignoring this relationship causes an overinvestment in spares to achieve a specific system-performance goal.

Ostensibly, the model's main use would be to determine inventory levels for each location. The model, however, was developed primarily as a tool for investigating the impact on both supply performance and investment of changing the Air Force's two-echelon supply system to a three-echelon system. Specifically, Air Force planners are interested in examining how such a change affects the requirement for logistics resources, and inventory investment in particular. Thus, in addition to being simply a mechanism for computing stock levels, the model can be effectively employed to answer many questions related to the design of a logistics system. For example, issues that can be addressed include the number and siting of maintenance centers, the impact of changing pipeline times on inventory investment, and the way that repair capability—measured in the model in terms of the probability that an item is repaired at a particular location—alters the investment in inventory.

REFERENCES

[1] Air Force Logistics Command Regulation 57-27, "Determination of Requirements of Initially Provisioned Items" (January 1969).

[2] Feeney, George J., and Craig C. Sherbrooke, "The (S-1,S) Inventory Policy Under Compound Poisson Demand," Management Science *12*, 391-411 (1966).

[3] Fox, Bennett, and M. Landi, "Searching for the Multiplier in One-Constraint Optimization Problems," Operations Research *18*, 253-262 (1970).

[4] Muckstadt, John A., "A Model for a Multi-Item, Multi-Echelon, Multi-Indenture Inventory System," Management Science *20*, 472-481 (1973).

[5] Muckstadt, John A., "Some Approximations in Multi-Item, Multi-Echelon Inventory Systems for Recoverable Items," Naval Research Logistics Quarterly *25*, 377-394 (September 1978).

[6] Muckstadt, John A., "NAVMET: A Four-Echelon Model for Determining the Optimal Quantity and Distribution of Navy Spare Aircraft Engines," Report R-7511, Naval Weapons Engineering Support Activity, Washington Navy Yard, Washington, D.C. (October 1976).

[7] Sherbrooke, Craig C., "METRIC: A Multi-Echelon Technique for Recoverable Item Control," Operations Research *16*,, 122-141 (1968).

BAYESIAN ESTIMATION AND OPTIMAL DESIGNS IN PARTIALLY ACCELERATED LIFE TESTING*

Morris H. DeGroot

Carnegie-Mellon University
Pittsburgh, Pennsylvania

Prem K. Goel

Purdue University
Lafayette, Indiana

ABSTRACT

A method of life testing is proposed which combines both ordinary and accelerated life-testing procedures. It is assumed that an item can be tested either in a standard environment or under stress. The amount of stress is fixed in advance and is the same for all items to be tested. However, the time x at which an item on test is taken out of the standard environment and put under stress can be chosen by the experimenter subject to a given cost structure. When an item is put under stress its lifetime is changed by the factor α. Let the random variable T denote the lifetime of an item in the standard environment, and let Y denote its lifetime under the partially accelerated test procedure just described. Then $Y = T$ if $T \leqslant x$, and $Y = x + \alpha (T - x)$ if $T > x$. It is assumed that T has an exponential distribution with parameter θ. The estimation of θ and α and the optimal design of a partially accelerated life test are studied in the framework of Bayesian decision theory.

1. INTRODUCTION

In many problems of life testing, the experimenter realizes that the test process may require an unacceptably long time period for its completion if the test is simply carried out under specified standard stress conditions. In such problems, the experimenter is generally able to run the life test under stresses that are higher than the specified standard in order to accelerate the process and shorten the time to its completion. Furthermore, he can either start the life test under these higher stresses and continue the test under these conditions to completion, or he can start the test under the standard conditions and only apply the higher stresses if the test is not completed by some specified time.

This type of problem does not seem to have been treated in the literature on life testing or accelerated life testing, where it is usually assumed that the experimenter can control the levels of higher stress to be used in the test. However, it is also assumed that the entire test

*This research was supported in part by the National Science Foundation under Grant SOC 77-07548.

must be carried out at this fixed higher level. Some of the standard work in this area will now be described briefly.

Epstein [4] has presented life-testing problems in which a number of items are put on test and the testing process is carried out for some period of time and then terminated in accordance with some specified stopping rule. The lifetimes of different items are assumed to be independent, each is assumed to have some specified distribution function (d.f.) $F(t, \theta)$, and inferences about the parameter θ are to be made on the basis of the outcomes of the testing process. Problems of optimal design connected with this life-testing process would involve the questions of how many items to put on test, whether or not to replace items when they fail, and how to specify an optimal stopping rule.

Chernoff [2] and Bessler et al. [1] introduced and studied the concept of accelerated life tests. In these tests, the parameter θ, which appears in the d.f. $F(t, \theta)$, is regarded as a specified function $\theta = \psi(s, \alpha)$ of an environmental stress s, to which an item on test can be subjected, and an unknown parameter α. They consider problems of estimation of α and of optimal design of the testing process in both sequential and nonsequential contexts. The distribution of life-times is assumed to be exponential and the function ψ is usually taken to be linear.

Some recent references on accelerated life testing are Meeker and Nelson [7], which considers Weibull and extreme-value distributions, and Nelson and Kielpinski [8], which considers censored tests for normal and lognormal distributions. They also assume that the function ψ is linear.

As previously mentioned, in our work we assume that the experimenter can control the time at which a test item is switched from the standard stress conditions to higher stresses. In many problems, such as those of accelerated life testing, it will also be possible for the experimenter to choose various levels of higher stresses. For simplicity, in this paper we shall restrict ourselves to problems in which the higher level of stress is fixed in advance and is the same for all items to be tested.

Since this framework combines both ordinary and accelerated life-testing procedures, we will call it *partially accelerated life testing.*

We shall denote the lifetime of an item tested under the standard conditions by the random variable T, and we shall let $F(t, \theta)$ denote the d.f. of T. Here, the value of the parameter θ is unknown and is to be estimated. Suppose that if the item has not failed by some specified time x, then it is switched to the higher level of stress and the test is continued until the item fails. We assume that the effect of this switch is to multiply the remaining lifetime of the item by some unknown factor $\alpha > 0$.

In general, α will be a function of the higher stress levels that are chosen. However, since we are assuming that only one higher stress level is used, α can be regarded as a constant. Furthermore, since the effect of switching to the higher stress level will typically be to shorten the life of the test item, usually α will be less than 1.

To describe the model for this partially accelerated life test, we shall let Y denote the total lifetime of a test item. Thus, Y is defined by the relation

$$Y = \begin{cases} T & \text{for } T \leqslant x, \\ x + \alpha(T - x) & \text{for } T > x. \end{cases} \tag{1.1}$$

Since switching to the higher stress level can be regarded as tampering with the ordinary life test, Y is called a *tampered random variable*, x is called the *tampering point*, and α is called the *tampering coefficient*. This model and an application were introduced by Goel [5].

We shall assume that an experimenter starts with a sample of n items and subjects them to test in the standard environment. If item i has not failed by some prespecified time x_i, then it is put under the higher stress and the test is continued. If T_i would be the lifetime of item i in the standard environment, then the total lifetime Y_i of item i under this partially accelerated life test is given by (1.1). It would be possible to consider problems in which the tampering point x_i for item i is chosen sequentially, after the experimenter has observed whether or not some of the other items have previously failed, but we shall not do so in this paper.

The statistical problems involved in using the model (1.1) are (i) the estimation of θ and α for given values of the tampering points $x_1, \ldots, x_n$ and (ii) the choice of an optimal design for this estimation, i.e., the selection of the best tampering points.

The estimation of θ and α based on tampered random variables $Y_1, \ldots, Y_n$ corresponding to the tampering points $x_1, \ldots, x_n$, which may or may not be distinct, has been discussed by Goel [6]. In that paper, the consistency and asymptotic normality of the maximum likelihood estimators was demonstrated for different types of distributions under various conditions on the tampering points.

We shall consider the estimation problem in the framework of Bayesian decision theory and determine the optimal design for various loss and cost structures. The discussion in this paper is restricted to cases in which the random variable T has an exponential distribution with parameter θ. However, the results on optimal design will be valid for a somewhat broader class of distributions.

2. TERMINOLOGY AND NOTATION

The following terminology and notation will be used throughout the paper. A sample of n observations $Y_1, \ldots, Y_n$ is obtained on the random variable Y corresponding to preassigned tampering points $x_1, \ldots, x_n$. If the observed value y_i of Y_i is less than the corresponding tampering point x_i, then Y_i is called an *untampered observation*. Otherwise, Y_i is called a *tampered observation*. Thus, an untampered observation comes from a test item that failed under the standard conditions, and a tampered observation comes from a test item that failed after it had been switched to the higher-stress level.

The number of tampered observations among $Y_1, \ldots, Y_n$ is denoted by the random variable M. Also, we shall let A denote the set of indices $i \in \{1, \ldots, n\}$ for which Y_i is a tampered observation and let $\bar{A}$ denote the complementary set of indices corresponding to untampered observations. Thus, A contains M elements, $\bar{A}$ contains $n - M$ elements, $Y_i > x_i$ for $i \in A$, and $Y_j \leqslant x_j$ for $j \in \bar{A}$.

Let $\pi_i(\theta)$ denote the probability that the i^{th} test item will be tampered. Then

$$\pi_i(\theta) = \Pr(T_i > x_i \mid \theta). \tag{2.1}$$

It is convenient to introduce independent random variables $\zeta_1, \ldots, \zeta_n$ such that $\Pr(\zeta_i = 1) = \pi_i(\theta)$ and $\Pr(\zeta_i = 0) = 1 - \pi_i(\theta)$, for $i = 1, \ldots, n$. Then the conditional distribution of M

given θ is identical to the conditional distribution of $\sum_1^n \zeta_i$ given θ. Hence, for any prior distribution of θ, the prior (predictive) distribution of M is identical to the marginal distribution of $\sum_1^n \zeta_i$.

We shall now present a summary of the results obtained in this paper. In Section 3 it is assumed that T has an exponential distribution with the following p.d.f.:

$$f(t \mid \theta) = \begin{cases} \theta \exp(-t\theta) & \text{for } t > 0, \\ 0 & \text{for } t \leqslant 0. \end{cases} \tag{2.2}$$

It is assumed that the joint prior distribution of θ and α belongs to an appropriate conjugate family, and the Bayes estimators of θ and α are then obtained for a particular class of loss functions. The corresponding Bayes risks are given in a form suitable for use in the optimal design problem.

In Section 4, we consider the optimal design problem for a general class of risk functions. First, a random cost structure is assumed, whereby one pays a fixed cost for each tampered observation and another fixed cost for each untampered observation. It is proved that for this cost structure the optimal design uses only two different tampering points, namely, $x = 0$ and $x = \infty$. In other words, some observations are immediately tampered, and the rest are not tampered at all. For a wide class of other cost functions, the optimal design is shown to be of the same structure. The paper is written in such a way that a reader who is interested mainly in the optimal design problem may skip Section 3 and proceed directly to Section 4.

In Section 5, the results of Section 4 are applied to the specific estimation problems discussed in Section 3. For each problem the optimal solution is obtained for a class of cost functions which admit a two-point optimal design. Finally, a cost function, for which the optimal design is not concentrated on two points, is presented.

3. BAYES ESTIMATION FOR EXPONENTIAL DISTRIBUTIONS

In this section the p.d.f. of the random variable T is assumed to be of the form (2.2). We first assume that the parameter θ is known, say $\theta = \theta_0$, and we want to estimate the unknown parameter α. We will then use these results for the case in which θ is unknown. It is convenient to work with the parameter $\beta = 1/\alpha$. In most problems of partially accelerated life testing β will be greater than 1. However, in order not to restrict the applicability of this model we shall consider prior distributions for β that assign positive density to all positive values of β. If the experimenter is almost certain that $\beta > 1$, then he can choose a prior distribution of the form (3.1) below that assigns a suitably small probability to the interval $0 < \beta < 1$.

We shall assume that the prior distribution of β is a gamma distribution with parameters r and $s\theta_0$, the p.d.f. of which is

$$g(\beta) = \begin{cases} [(s\theta_0)^r/\Gamma(r)]\,\beta^{r-1} \exp(-s\theta_0\beta) & \text{for } \beta > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{3.1}$$

Since these distributions form a conjugate family in this problem, it follows that, given the values of $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$, the posterior distribution of β is again a gamma distribution with parameters r_1 and $s_1\theta_0$ (see DeGroot [3], p. 166), where

$$r_1 = m + r \text{ and } s_1 = s + \sum_{i \in A} (Y_i - x_i). \tag{3.2}$$

If there are no tampered observations in the sample, then we obtain no information about the value of α and the posterior distribution of α is the same as the prior.

Since β is a scale parameter, it is reasonable to consider loss functions for its estimation that are invariant under changes in the units of measurement of lifetimes. The following two loss functions have this property:

$$L_1(\hat{\beta}, \beta) = \left(\frac{\hat{\beta}}{\beta} - 1\right)^2 \text{ and } L_2(\hat{\beta}, \beta) = \left(\frac{\beta}{\hat{\beta}} - 1\right)^2. \tag{3.3}$$

Each of these loss functions measures the relative squared error and combines the standard squared-error loss with the invariance requirement. The loss function L_1 measures the squared error relative to the actual value of β, and L_2 measures the squared error relative to the magnitude of the estimate $\hat{\beta}$.

In fact, however, we shall obtain our results for a wider class of loss functions containing L_1 and L_2. Specifically, we shall assume that the loss function is of the form

$$L(\hat{\beta}, \beta) = \hat{\beta}^k \beta^l (\hat{\beta} - \beta)^2, \tag{3.4}$$

where $-2 \leqslant k \leqslant 0$ and $-\infty < l < \infty$. It should be noted that only values of k in the interval $-2 \leqslant k \leqslant 0$ are of interest, because the Bayes risk is infinite for $k < -2$ and is 0 for $k > 0$. Furthermore, if $L(\hat{\beta}, \beta)$ satisfies (3.4) and we let $\hat{\alpha} = 1/\hat{\beta}$, then $L(\hat{\beta}, \beta) = \hat{\alpha}^{k_1} \alpha^{l_1} (\hat{\alpha} - \alpha)^2$, where $k_1 = -k-2$ and $l_1 = -l-2$. Again, k_1 will lie in the interval $-2 \leqslant k_1 \leqslant 0$. Hence, estimation of β is equivalent to estimation of α for appropriate values k and l. The loss function L reduces to L_1 when $k = 0$ and $l = -2$, and to L_2 when $k = -2$ and $l = 0$.

THEOREM 1: For $r + l > 0$, the Bayes estimator $\hat{\beta}$ with respect to the loss function L is given by

$$\hat{\beta} = \gamma_k(\eta_1) \frac{\delta_1}{s_1 \theta_0}, \tag{3.5}$$

where r_1 and s_1 are given by (3.2), $\eta_1 = \delta_1 = r_1 + l$, and for $\eta > 0$ the function γ_k is defined by

$$\gamma_k(\eta) = \begin{cases} \dfrac{1}{k+2}\left\{k + 1 + \left[1 - \dfrac{k(k+2)}{\eta}\right]^{\frac{1}{2}}\right\} & \text{for } -2 < k \leqslant 0, \\ 1 + \dfrac{1}{\eta} & \text{for } k = -2. \end{cases} \tag{3.6}$$

The Bayes risk ρ_1, for given tampering points $x_1, \ldots, x_n$, is

$$\rho_1 = \frac{\Gamma(r + k + l + 2)}{\Gamma(r)\,(s\theta_0)^{k+l+2}} E_M\left[\frac{\Gamma(r_1 + l + 1)}{\Gamma(r_1 + l + k + 2)} \rho(\delta_1, \eta_1)\right]. \tag{3.7}$$

The expectation in (3.7) is taken with respect to the conditional distribution of M given $\theta = \theta_0$, as defined in Section 2, with $\pi_i(\theta_0) = \exp(-\theta_0 x_i)$, and $\rho(\delta, \eta)$ is defined as follows for $\delta > 0$ and $\eta > 0$:

$$\rho(\delta, \eta) = \delta^{k+1}[\gamma_k(\eta)]^k \left\{\frac{1}{\eta} + [\gamma_k(\eta) - 1]^2\right\} \tag{3.8}$$

PROOF: It follows from (3.4) that the posterior risk for any estimator $\tilde{\beta}$ is given by

$$R(\tilde{\beta}) = \tilde{\beta}^{k+2} E(\beta^l) - 2\tilde{\beta}^{k+1} E(\beta^{l+1}) + \tilde{\beta}^k E(\beta^{l+2}),$$

where the posterior moments of β are given by $E(\beta^l) = \Gamma(r_1+l)/[\Gamma(r_1)(s_1\theta_0)^l]$ for $r_1 + l > 0$. For $-2 \leqslant k \leqslant 0$, the solution $\hat{\beta}$ of the equation $\partial R(\tilde{\beta})/\partial\tilde{\beta} = 0$ that minimizes $R(\tilde{\beta})$ is given by (3.5) and (3.6). Hence, $\hat{\beta}$ is the Bayes estimator. It follows that the posterior Bayes risk can be written as

$$R(\hat{\beta}) = \frac{\Gamma(r_1+l+1)}{\Gamma(r_1)(s_1\theta_0)^{k+l+2}} \rho(\delta_1, \eta_1). \tag{3.9}$$

The Bayes risk ρ_1 is the expectation of $R(\hat{\beta})$ with respect to the joint marginal distribution of $Y_1, \ldots, Y_n$. However, the distribution of β, given $x_1, \ldots, x_n$ and the set A, is the same as its prior distribution. It can be shown, therefore, that the p.d.f. of the posterior parameter s_1, given $x_1, \ldots, x_n$ and the set A, is

$$f(s_1) = \frac{\Gamma(m+r)}{\Gamma(m)\Gamma(r)} (s_1 - s)^{m-1} \frac{s^r}{s_1^{m+r}} \quad \text{for } s_1 \geqslant s,$$

and $f(s_1) = 0$ otherwise. Hence,

$$E[s_1^{-(k+l+2)} \mid M = m] = \frac{\Gamma(r+k+l+2)\,\Gamma(m+r)}{\Gamma(r)\,\Gamma(m+r+l+k+2)}. \tag{3.10}$$

It follows from (3.9) and (3.10) that the Bayes risk ρ_1 is given by (3.7).

It is noteworthy that θ_0 does not appear in the conditional p.d.f. $f(s_1)$.

When θ is unknown, a conjugate family of joint prior distribution for β and θ can be specified as follows:

Given $\theta = \theta_0$, the conditional prior distribution of β is a gamma distribution with parameters r and $s\theta_0$, and the prior distribution of θ is a gamma distribution with parameters r_0 and s_0.

It follows that, given the observations $y_1, \ldots, y_n$ and the values of $x_1, \ldots, x_n$ and the set A, the joint posterior distribution of β and θ can be specified as follows:

Given $\theta = \theta_0$, the conditional posterior distribution of θ is a gamma distribution with parameters r_1 and $\theta_0 s_1$, where r_1 and s_1 are given by (3.2), and the posterior distribution of θ is a gamma distribution with parameters r_2 and s_2, where r_2 and s_2 are defined by

$$r_2 = r_0 + n - m \text{ and } s_2 = s_0 + \sum_{i \in A} x_i + \sum_{j \in \bar{A}} Y_j. \tag{3.11}$$

If all the observations are untampered, then $m = 0$, $s_1 = s$, and $s_2 = s_0 + \sum_{i=1}^{n} Y_i$. If all the observations are tampered, then $m = n$, $s_1 = s + \sum_{i=1}^{n} (Y_i - x_i)$, and $s_2 = s_0 + \sum_{i=1}^{n} x_i$.

It should be noted that this posterior distribution does not depend on the values of the tampering points corresponding to the untampered observations. Hence, it does not depend on the method by which these points were chosen.

We will consider the following estimation problems in the remainder of this section: (i) the estimation of θ only, (ii) the estimation of β only, (iii) the estimation of both β and θ.

Estimation of θ: It is assumed that the loss function due to estimation error in θ is of the form $L(\hat{\theta}, \theta) = \hat{\theta}^k \theta^l (\hat{\theta} - \theta)^2$, with $-2 \leqslant k \leqslant 0$. Since the posterior distribution of θ is similar to the posterior distribution of β when $\theta = \theta_0$, the Bayes estimator $\hat{\theta}$ can be derived by methods used in Theorem 1. Furthermore, after some algebraic manipulations, the posterior risk of $\hat{\theta}$ can be written as

$$R(\hat{\theta}) = \frac{\Gamma(r_2 + l + 1)}{\Gamma(r_2)\, s_2^{k+l+2}}\, \rho(\delta_2, \eta_2), \tag{3.12}$$

where r_2 and s_2 are given by (3.11), $\rho(\delta, \eta)$ is given by (3.8), and $\delta_2 = \eta_2 = r_2 + l$. If $k + l + 2 \neq 0$, it seems impossible to integrate (3.12) with respect to the joint marginal distribution of $y_1, \ldots, y_n$. However, if $k+l+2 = 0$, the posterior risk $R(\hat{\theta})$ is a function of M only and the expectation of (3.12) reduces to a form similar to (3.7). It should be noted that the loss functions L_1 and L_2 in (3.3) satisfy the condition $k + l + 2 = 0$. The following theorem results from this discussion:

THEOREM 2: If $r_0 + l > 2$, then the Bayes estimator $\hat{\theta}$ with respect to the loss function L is given by

$$\hat{\theta} = \gamma_k(\eta_2)\, [\delta_2 / s_2]. \tag{3.13}$$

If $k + l + 2 = 0$, then the Bayes risk is given by

$$\rho_2 = E_M \left[\frac{\Gamma(r_2 + l + 1)}{\Gamma(r_2)}\, \rho(\delta_2, \eta_2) \right]. \tag{3.14}$$

The distribution of M in (3.14) is as defined in Section 2, with $\pi_i(\theta) = \exp(-\theta x_i)$, $i = 1, \ldots, n$.

Estimation of β: The next result gives the Bayes estimator $\hat{\beta}$ and the corresponding Bayes risk.

THEOREM 3: For $r + l > 0$ and $r_0 > l + 2$, the Bayes estimator $\hat{\beta}$ with respect to the loss function L is given by

$$\hat{\beta} = \gamma_k(\eta_3)\, \delta_3 s_2 / s_1, \tag{3.15}$$

where

$$\delta_3 = \frac{r_1 + l}{r_2 - l - 1} \text{ and } \eta_3 = \frac{(r_1 + l)(r_2 - l - 2)}{r_1 + r_2 - 1} \tag{3.16}$$

and all other variables are as previously defined. If $k + l + 2 = 0$, then the Bayes risk is given by

$$\rho_3 = E_M \left[\frac{\Gamma(r_1 + l + 1)\Gamma(r_2 - l - 1)}{\Gamma(r_1)\,\Gamma(r_2)} \rho(\delta_3, \eta_3) \right], \tag{3.17}$$

where the distribution of M is as given in Theorem 2.

PROOF: Since the lth posterior moment of β is given by

$$E(\beta^l) = \frac{\Gamma(r_1 + l)\,\Gamma(r_2 - l)}{\Gamma(r_1)\,\Gamma(r_2)} \left(\frac{s_2}{s_1} \right)^l, \tag{3.18}$$

it follows, from a discussion similar to that given for Theorem 1, that $\hat{\beta}$ is given by (3.15). Furthermore, the posterior Bayes risk can be written as a product of a function of M and $(s_2/s_1)^{k+l+2}$. If $k + l + 2 = 0$, expression for ρ_3 follows from a discussion similar to that given for Theorem 2.

Estimation of β and θ: We shall assume that the loss from estimating both β and θ is a linear combination of the losses resulting from the estimation of each of the parameters separately. Thus, the loss function is given by

$$L(\hat{\theta}, \hat{\beta}; \theta, \beta) = \lambda_1 \hat{\beta}^{k_1} \beta^{l_1} (\hat{\beta} - \beta)^2 + \lambda_2 \hat{\theta}^{k_2} \theta^{l_2} (\hat{\theta} - \theta)^2, \tag{3.19}$$

where λ_1 and λ_2 are positive constants.

The Bayes estimators $\hat{\theta}$ and $\hat{\beta}$ are given in Theorems 2 and 3 with the appropriate choices of the values of k and l. Furthermore, if $k_i + l_i + 2 = 0$ for $i = 1, 2$, then the Bayes risk is given by the corresponding linear combination of ρ_2 and ρ_3.

The foregoing analysis can be performed for other loss functions and different distributions of the random variable T. In particular, when T follows a uniform distribution, the analogous results are given in [5].

In this section all the expressions for the Bayes risks are given in the form $E[h(M)]$, where the function h is explicitly known. It should be noted that it is difficult to find the expectation and the risk as an explicit function of $x_1, \ldots, x_n$. In Section 4 we shall consider the problem of choosing optimal designs based only on the knowledge of the function h. The results obtained will, therefore, be applicable to other distributions of the random variable T as long as the function h is known.

4. OPTIMAL DESIGNS FOR ESTIMATION

Suppose now that the experimenter has to pay a cost for each item tested. In general, this cost will depend on the tampering point x and on whether or not the observation is actually tampered. Under these conditions, the experimenter desires to choose an optimal design for

the estimation of the unknown parameters, i.e., to choose n tampering points $x_1, \ldots, x_n$ such that the total risk (the sum of the Bayes risk due to the estimation error and the cost of choosing the tampering points) is a minimum. In general, it will be difficult to get a closed form solution to this minimization problem unless a simple expression for the Bayes risk is available. In accordance with the results in Section 3, we shall now assume that this risk can be written as

$$\rho = E[h(M)], \tag{4.1}$$

where the function h is explicitly known and M is the number of tampered observations. We shall show that the optimal design problem can then be solved for a wide class of cost functions.

The results to be presented in this section are not restricted to the case in which T has an exponential distribution but are applicable for any specified family of continuous distributions $F(t, \theta)$ supported on the nonnegative part of the real line.

Suppose, first, that the cost of any observation depends only on whether or not it is tampered and not in any other way on the value of the tampering point x. We shall assume that the cost of an observation is $\nu_1 \geqslant 0$ if it is untampered and $\nu_2 > \nu_1$ if it is tampered. This cost structure seems to be valid in many practical problems. Furthermore, it will also be used later in this section as a device to solve optimal design problems with other cost functions. Therefore, the total cost of the observations is equal to $n\nu_1 + (\nu_2 - \nu_1)M$, and we must minimize the total expected risk

$$R_0 = \rho + (\nu_2 - \nu_1)E(M) + n\nu_1 = E[h(M) + (\nu_2 - \nu_1)M] + n\nu_1. \tag{4.2}$$

Except for a constant, the risk R_0 in (4.2) is the expected value of the random variable $h(M) + (\nu_2 - \nu_1)M$. Therefore, among all possible distributions of M, it is minimized when the distribution of M assigns probability 1 to the integer m_0 satisfying

$$h(m_0) + (\nu_2 - \nu_1)m_0 = \min_{0, 1, \ldots, n} [h(i) + (\nu_2 - \nu_1)i]. \tag{4.3}$$

This degenerate distribution of M is achieved by if we choose m_0 tampering points at $x = 0$, so that these observations are tampered immediately, and the remaining $n - m_0$ tampering points at $x = \infty$, so that these observations are never tampered. Thus, under the optimal design the experimenter never leaves to chance whether or not an observation will be tampered.

The cost structure we have just considered is random in the sense that the cost of an observation is not fixed in advance but depends on whether or not the observation turns out to be tampered. We shall now assume that the cost $c(x)$ of each observation is fixed in advance and depends only on the tampering point x. Now for the optimal design we need to choose the tampering points $x_1, \ldots, x_n$ to minimize

$$R(x_1, \ldots, x_n) = E[h(M)] + \sum_{i=1}^{n} c(x_i). \tag{4.4}$$

For any given tampering point x, let

$$p(x) = E[\Pr(T > x \mid \theta)], \tag{4.5}$$

where the expectation is evaluated with respect to the given prior distribution of θ. In other words, $p(x)$ is the prior probability that an observation will be tampered when the tampering point x is used. It follows that $E(M) = \sum_{i=1}^{n} p_i$, where $p_i = p(x_i)$ for $i = 1, \ldots, n$.

Under the assumptions on the distribution of T made at the beginning of this section, $p(x)$ is a strictly decreasing function of x for $x \geq 0$. Therefore, the cost function $c(x)$ can be written as a function of p, rather than of x, which we denote by $c^*(p)$. In other words, $c^*(p)$ is the cost of choosing a tampering point for which the probability is p that the observation will be tampered.

In particular, if the cost function c^* is of the form

$$c_0^*(p) = \nu_1 + (\nu_2 - \nu_1)p, \tag{4.6}$$

then the risk defined by (4.4) is equal to the risk R_0 defined by (4.2). Therefore, it follows from the discussion for the random cost structure that the optimal design is to choose m_0 tampering points at $x = 0$ and the remaining $(n - m_0)$ tampering points at $x = \infty$, where m_0 is defined by (4.3). In fact, as we shall now show, there is a wide class of functions with this type of solution.

THEOREM 4: Suppose that the cost function $c^*(p)$ satisfies the condition

$$c^*(p) \geq pc^*(1) + (1 - p)c^*(0) \quad \text{for } 0 < p < 1. \tag{4.7}$$

Then the total risk is minimized by the solution obtained for the cost function c_0^* satisfying (4.6), with $c_0^*(0) = \nu_1$ and $c_0^*(1) = \nu_2$.

PROOF: Since the cost function satisfies (4.7), it follows that the risk defined in (4.4) satisfies the relation

$$R(x_1, \ldots, x_n) \geq E[h(M)] + \sum_{i=1}^{n} c_0^*(p_i) \geq R_0^*, \tag{4.8}$$

where R_0^* is the risk corresponding to the optimal design for the cost function c_0^*. However, $R(x_1, \ldots, x_n) = R_0^*$ when $x_i = 0$ (and $p_i = 1$) for m_0 values of i, and $x_i = \infty$ (and $p_i = 0$) for the remaining $n - m_0$ values of i. Hence, this solution is also the optimal design for the cost function satisfying (4.7).

COROLLARY 1: If the cost function $c^*(p)$ is a concave function of p on $[0, 1]$, then only the values $x = 0$ and $x = \infty$ need be used in an optimal design.

The results presented thus far in this section indicate that, for a wide class of cost functions, the optimal design does not involve a *partially* accelerated test on any item. For some of the items, the test is carried out entirely under the standard stress conditions; and for the remaining items, it is carried out entirely under the higher stress.

The techniques used in proving the above results can also be helpful in problems in which the optimal design does involve partially accelerated life tests. The following results indicate the kind of simplification that can be obtained in characterizing the optimal design.

THEOREM 5: If $h(m)$ is monotone on the integers $0, 1, \ldots, n$ and the cost function $c(x)$ is constant on some interval $a \leq x \leq b$, then there exists an optimal design that does not use any tampering point in the interior of that interval.

COROLLARY 2: Suppose that $h(m)$ is nonincreasing on the integers $0, 1, \ldots, n$, and there exist values $0 = x_0^* < x_1^* < \ldots < x_k^* < x_{k+1}^* = \infty$ and $\nu_0 > \nu_1 > \ldots > \nu_k > \nu_{k+1} \geq 0$ such that $c(x) = \nu_i$ for $x_i^* \leq x < x_{i+1}^*$. Then only the values $0, x_1^*, \ldots, x_k^*$ need be used in an optimal design.

An analogous result can be given when $h(m)$ is nondecreasing and $c(x)$ is a left-continuous, increasing step function.

5. APPLICATIONS

In this section we shall apply the results of Section 4 to the estimation problems considered in Section 3. Throughout this section we shall assume that the loss functions have the relative squared-error form (3.3).

Estimation of β when θ is known: It follows from Theorem 1 that the function h is of the form

$$h(m) = b/(\Delta + m), \tag{5.1}$$

where b and Δ are positive constants that depend on the prior distribution of β and on which of the two loss functions in (3.3) is used. Here, the parameter θ of the exponential distribution is known, and only the tampering parameter β is unknown. Suppose that the cost function $c(x)$ is a nondecreasing function of x. Then the obvious solution to the optimal design problem is to choose $x_i = 0$ for all i and to tamper all observations immediately. These observations will simultaneously be the cheapest ones available and will yield the maximum information about β. In general, for any cost function $c(x)$, if $a_1 < a_2$ and $c(a_1) \leq c(a_2)$, then the point a_2 need not be considered as a possible tampering point in the optimal design.

EXAMPLE 1: Suppose that the cost function $c(x)$ satisfies the following property:

$$c(0) = \nu_2,\ c(\infty) = \nu_1 < \nu_2, \text{ and}$$
$$c(x) \geq \nu_1 + (\nu_2 - \nu_1) \exp(-\theta_0 x) \text{ for } 0 < x < \infty, \tag{5.2}$$

where θ_0 is the known value of θ. Here $c(\infty)$ is the cost of not tampering an observation at all. It follows from Theorem 4 that there exists an optimal design with m_0 tampering points at $x = 0$ and the remaining $n - m_0$ tampering point at $x = \infty$. It can be proved that the optimal value of m_0 is given by

$$m_0 = k \text{ for } t_{k+1} \leq \nu_2 - \nu_1 < t_k \ (k = 0, 1, \ldots, n), \tag{5.3}$$

where $t_0 = \infty$, $t_k = b/[(\Delta + k - 1)(\Delta + k)]$ for $k = 1, \ldots, n$, and $t_{n+1} = 0$.

EXAMPLE 2: Let the cost function $c(x)$ satisfy

$$c(x) = \begin{cases} \nu_2, & 0 \leq x < a, \\ \nu_1, & a \leq x \leq \infty, \end{cases} \tag{5.4}$$

where $\nu_2 > \nu_1$. It follows from Corollary 2 that there exists an optimal design with m_1 tampering points at $x = 0$ and $n - m_1$ tampering points at $x = a$. To find the value of m_1, define

$$t_{m+1}^* = b(e^{\theta_0 a} - 1)\left[\frac{1}{\Delta + m} - E\left\{\frac{1}{\Delta + m + B_m}\right\}\right] \Big/ (m = 0, 1, \ldots, n - 1), \tag{5.5}$$

where the distribution of B_m is binomial, with parameters $(n - m)$ and $\exp(-\theta_0 a)$. It can be proved that t_m^* is decreasing in m and that the optimal value of m_1 is given by

$$m_1 = k \text{ for } t_{k+1}^* \leqslant \nu_2 - \nu_1 < t_k^* \ (k = 0, 1, \ldots, n), \tag{5.6}$$

where $t_0^* = \infty$ and $t_{n+1}^* = 0$.

Now suppose that the cost function $c(x)$ has the form $c(x) = \nu_1 + (\nu_2 - \nu_1) \exp(-\delta x)$ for some value of $\delta < \theta_0$. Then the optimal design will contain some points x_i with $0 < x_i < \infty$. For this function, it can be proved that the optimal design consists of at least k tampering points at $x = 0$ if $\nu_2 - \nu_1 \leqslant b\theta_0/[\delta(\Delta + k - 1)(\Delta + k)]$. However, the explicit solution is not known.

Estimation of θ only: Suppose now that both the parameters θ and β are unknown. Then it follows from Theorem 2 that the function h is of the form

$$h(m) = b/(\Delta + n - m), \tag{5.7}$$

where b and Δ are appropriate positive constants.

EXAMPLE 3: Suppose that the cost function $c(x)$ satisfies the following property:

$$c(0) = \nu_1, \ c(\infty) = \nu_2 > \nu_1, \text{ and}$$

$$c(x) \geqslant \nu_2 - (\nu_2 - \nu_1)\left(\frac{s_0}{s_0 + x}\right)^{r_0} \text{ for } 0 < x < \infty, \tag{5.8}$$

where r_0 and s_0 are the parameters of the prior gamma distribution of θ. It follows from Theorem 4 that there exists an optimal design with m_0 tampering points at $x = \infty$ and the remaining $n - m_0$ tampering points at $x = 0$, where m_0 is defined by (5.3).

Estimation of β or of both β and θ when θ is unknown: It follows from Theorem 3 and the discussion thereafter that the function h is of the form

$$h(m) = \frac{b_1}{\Delta_1 + m} + \frac{b_2}{\Delta_2 + (n - m)}, \tag{5.9}$$

where b_i and Δ_i are appropriate positive constants. Furthermore, it follows from Theorem 4 that, it the cost function satisfies (5.8), then there exists an optimal design with m_2 tampering points at $x = 0$ and the remaining $(n - m_2)$ tampering points at $x = \infty$. It can be proved that the optimal value of m_2 is given by

$$m_2 = k \text{ for } s_{k+1} \leqslant \nu_2 - \nu_1 < s_k \ (k = 0, 1, \ldots, n), \tag{5.10}$$

where $s_0 = \infty$, $s_{n+1} = 0$, and

$$s_k = \frac{b_1}{(\Delta_1 + k - 1)(\Delta_1 + k)} - \frac{b_2}{(\Delta_2 - k)(\Delta_2 - k + 1)}, \quad k = 1, \ldots, n.$$

We terminate this section with an example of a cost function in which the optimal design is concentrated at more than two points.

EXAMPLE 4: Suppose again that β is to be estimated when θ is known to be θ_0. Therefore, the function h is defined by (5.1). Suppose that the cost function $c(x)$ is as follows:

$$c(x) = \begin{cases} 2\nu, & 0 \leqslant x \leqslant a_1, \\ \nu, & a_1 \leqslant x \leqslant a_2, \\ 0, & a_2 \leqslant x \leqslant \infty. \end{cases} \tag{5.11}$$

It follows from Corollary 2 that only the values 0, a_1, and a_2 need be considered for an optimal design. Let $p_i^* = \exp(-\theta_0 a_i)$ for $i = 1, 2$. For $n = 3$, it can be proved that the optimal design is to choose one tampering point at each of the three values $x = 0$, a_1, and a_2 if

$$(1 - p_1^*) < \frac{\nu(\Delta + 1)(\Delta + 2)(\Delta + 3)}{b(\Delta + 3 - 2p_2^*)} \leqslant (p_1^* - p_2^*).$$

Otherwise, the optimal design is concentrated on at most two tampering points.

REFERENCES

[1] Bessler, S., H. Chernoff, and A.W. Marshall, "An Optimal Sequential Accelerated Life Test," Technometrics *4*, 367-379 (1962).

[2] Chernoff, H., "Optimal Accelerated Life Designs for Estimation," Technometrics *4*, 381-408 (1962).

[3] DeGroot, M.H., *Optimal Statistical Decisions* (McGraw-Hill, New York, 1970).

[4] Epstein, B., "Estimation from Life Test Data," Technometrics *2*, 447-454 (1960).

[5] Goel, P.K., "Some Estimation Problems in the Study of Tampered Random Variables," Technical Report No. 50, Department of Statistics, Carnegie-Mellon University, Pittsburgh, Pennsylvania (1971).

[6] Goel, P.K., "Consistency and Asymptotic Normality of Maximum Likelihood Estimators," Scandinavian Actuarial Journal *2*, 109-118 (1975).

[7] Meeker, W.Q., and W. Nelson, "Optimum Accelerated Life-Tests for the Weibull and Extreme Value Distributions," IEEE Transactions on Reliability *R-24*, 321-332 (1975).

[8] Nelson, W., and T.J. Kielpinski, "Theory for Optimum Accelerated Life Tests for Normal and Lognormal Life Distributions," Technometrics *18*, 105-114 (1976).

SOME BAYES TESTS AND THEIR ASYMPTOTIC PROPERTIES FOR THE MULTIVARIATE, MULTISAMPLE GOODNESS-OF-FIT PROBLEM

K. V. Ramani

Indian Institute of Management
Ahmedabad, India

ABSTRACT

Independent samples are taken from c multivariate populations with continuous but unknown cumulative distribution function (c.d.f.). The problem is to test the hypothesis that the c population c.d.f's are identical to a specified c.d.f. We approach this problem by first transforming the data so that the hypothesis being tested is that the common distribution is uniform over a unit hypercube. We then construct some Bayes tests and investigate their asymptotic properties. These tests are based on the asymptotic normality of the number of observations falling in the "asymptotically sufficient groupings."

1. INTRODUCTION

Let $\{X_{ij};\ j = 1, 2, \ldots n_i\}$ be n_i independent observations from a population with continuous but unknown c.d.f. $= F_i$ for $i = 1, 2, \ldots, c$. Let us denote the p components of the observation X_{ij} by X_{ij}^a, $a = 1, 2, \ldots\ldots, p$. We assume the c samples to be independent and are interested in testing the hypothesis.

(1) $\quad H_o \ : \ F_1 = F_2 = \ldots\ldots = F_c = F$ (say),

where F is a completely specified c.d.f.

We approach this problem by transforming the original variables X by using the following transformation due to Rosenblatt [4]: Define

$$Z_{ij}^1 = F(x_{ij}^1) = P\ [X_{ij}^1 \leqslant x_1],$$

$$Z_{ij}^2 = F(X_{ij}^2 \mid X_{ij}^1) = P\ [X_{ij}^2 \leqslant X_2 \mid X_{ij}^1 = x_1],$$

$$\cdot$$

$$\cdot$$

$$\cdot$$

$$z_{ij}^p = F(X_{ij}^P \mid X_{ij}^{p-1}, \ldots\ldots, X_{ij}^1) = P[X_{ij}^p \leqslant x_p \mid X_{ij}^{p-1} = x_{p-1}, \ldots, X_{ij_1} = x_1].$$

It is then easily seen that, when $H_o^{(1)}$ is true, then n_i independent observations $\{X_{ij}^a\ ;\ a = 1, 2, \ldots\ldots, p\}$ are transformed into n_i independent and identically distributed uniform

random variables $\{Z_{ij}^a,\ a = 1, 2, \ldots, p\}$ lying over the p — dimensional unit cube C_p (say). Thus, we have reduced our problem of testing $H_o^{(1)}$ to that of testing

$$H_o : f_{n_i}^i (z^1, z^2, \ldots, z^p) \equiv \begin{cases} 1 \text{ if } (z^1, z^2, \ldots, z^p) \text{ is in } C_p, \\ 0, \text{ otherwise.} \end{cases}$$

It should be noted that, when $H_o^{(1)}$ is false, the above transformation transforms $\{X_{ij}\}$ into independent random variables $\{Z_{ij}\}$ laying over C_p. So for testing H_o, we consider, as neighboring alternatives contiguous alternatives of the form

$$H_1 : f_{n_i}^i (z^1, z^2, \ldots, z^p) = \begin{cases} 1 + r_{n_i}^i (z^1, z^2, \ldots, z^p) \\ 0, \text{otherwise,} \end{cases} \quad \text{if } (z^1, z^2, \ldots, z^p) \text{ is in } C_P,$$

where $r_{n_i}^i (z^1, z^2, \ldots, z^p)$ satisfies the following regularity conditions:

(1.1) (i) $\int_{C_p} \cdots \int r_{n_i}^i (z^1, z^2, \ldots, z^p)\, dz^1 \ldots dz^p = 0\ i = 1, 2, \ldots, c;$

(ii) $\dfrac{\partial r_{n_i}^i}{\partial z^a} (z^1, z^2, \ldots, z^p)$ and $\dfrac{\partial^2 r_{n_i}^i}{\partial z^a\, \partial z^b} (z^1, z^2, \ldots, z^p)$ exist

and are such that

(1.2) $| r_{n_i}^i (z^1, z^2, \ldots, z^p) | < B(n),$

$$\left| \frac{\partial r_{n_1}^i}{\partial z^a} (z^1, z^2, \ldots, z^p) \right| < B(n),$$

and

$$\left| \frac{\partial^2 r_{n_i}^i}{\partial z^a\, \partial z^b} (z^1, z^2, \ldots, z^p) \right| < B(n),$$

where

$n = \text{Min}_i\ (n_i)$ and $B(n)$ is a nonrandom sequence satisfying the property:

(1.3) $\lim_{n \to \infty} n\, [B(n)]^3 = 0$

Here we assume that max n_i is of the same order of magnitude as min n_i.

Let us now discuss the reasonableness of the assumptions we have made on our alternate hypothesis. Suppose we look at an alternate hypothesis of the form

$$f_n^i (z^1, z^2, \ldots, z^p) = \begin{cases} 1 + r^i \dfrac{(z^1, z^2, \ldots, z^p)}{n^\delta}, & \text{if } (z^1, z^2, \ldots, z^p) \text{ is in } C_p, \\ 0, \text{ otherwise.} \end{cases}$$

Then it is easily seen that when $\delta > \frac{1}{2}$ the asymptotic power of any sequence of tests approaches the asymptotic level of significance, and such alternatives are too close to the null hypothesis to be challenging. And when $\delta < \frac{1}{2}$ it is easy to construct tests for which the

asymptotic power tends to one, and such alternatives are too far from the null hypothesis to be challenging. Thus, $\delta = \frac{1}{2}$ gives us interesting alternatives, and our class discusssed above includes the case $\delta = \frac{1}{2}$.

We conclude this section by pointing out that a "natural measure of distance" between uniform density over C_p and the density $\{1 + r_n^i(z^1, z^2, \ldots, z^p)\}$ is $\int_{C_p} \cdots \int r_n^{(i)2}(z^1, z^2, \ldots, z^p)\, dz^1, \ldots dz^p$.

It is a consequence of this result that later we investigate tests with good asymptotic power with respect to the distance

$$\sum_{i=1}^{c} \int_{C_p} \cdots \int r_n^{(i)2}(z^1, z^2, \ldots, z^p) dz^1\, dz^2 \ldots dz^p.$$

2. A GENERAL TEST CRITERION

We now derive a general test criterion based on the chi-squared distribution for our problem of testing the goodness of fit hypothesis, which is equivalent to testing the hypothesis

$$H_o : r_{n_i}^i(z^1, z^2, \ldots, z^p) = 0$$

for $i = 1, 2, \ldots, c$ where $r_{n_i}^i(z^1, z^2, \ldots, z^p)$ satisfies the regularity conditions (1.1), (1.2), and (1.3). We first expand $r_{n_i}^i(z^1, z^2, \ldots, z^p)$ in a Fourier cosine series over C_p as

$$\sum_{a_1=0}^{\infty} \sum_{a_2=0}^{\infty} \sum_{a_p=0}^{\infty} A_n^{(i)}(a_1, a_2, \ldots, a_p)\, 2^{D(a_1, \ldots a_p)} \left[\prod_{j=1}^{p} \cos(a_j \Pi z_j)\right], \tag{2.1}$$

where $D(a_1, a_2, \ldots, a_p)$ is equal to one half of the number of positive values among $(a_1, a_2, \ldots, a_p)$,

$$A_n^{(i)}(a_1, a_2, \ldots, a_p) = \int_{C_p} \cdots \int r_{n_i}^i(z^1, z^2, \ldots, z^p)\, 2^{D(a_1, a_2, \ldots, a_p)} \left[\prod_{j=1}^{P} \cos(a_j \Pi z^j)\right] dz^1 dz^2 \ldots dz^p, \tag{2.2}$$

and $A_n^{(i)}(0, 0, \ldots, 0) = 0$.

For each n we now divide C_p into $k(n)$ equal subcubes $S_n(b)$, $b = 1, 2, \ldots, k(n)$, where $k(n)$ is chosen such that a number $N_n^i(b)$ of observations Z_{ij} falling in $S_n(b)$, $b = 1, 2, \ldots, k(n)$, can be considered to be asymptotically sufficient for all practical purposes; the $S_n(b)$'s are then said to be "asy:suff:groupings." For an understanding of asy:suff:groupings the reader should consult [5]. For any nonnegative integers $a_1, a_2, \ldots, a_p$, not all zero, define

$$B[a_1, a_2, \ldots, a_p; L_n(b)] = 2^{D(a_1, a_2, \ldots, a_p)} \prod_{j=1}^{p} \cos[a_j \Pi L_n^j,] \ldots, (b), \tag{2.3}$$

where $L_n(b) = [L_n^1(b), L_n^2(b) \ldots, L_n P(b)]$ is the center of $S_n(b)$. Also define

$$\bar{A}_n^{(i)}(a_1, a_2, \ldots a_p,) = n_i^{-\frac{1}{2}} \left\{\sum_{b=1}^{k(n)} N_n^i(b)\, B[a_1, a_2, \ldots, a_p; L_n(b)] - \frac{n_i}{k(n)} \sum_{b=1}^{k(n)} B[a_1, \ldots, a_p, L_n(b)]\right\}. \tag{2.4}$$

Notice that $\bar{A}_n^{(i)}(a_1, a_2, \ldots, a_p)$ is observable. We then have the following:

THEOREM 1: For each sample $i, i = 1, 2, \ldots, c$, let $\{a_1^k, a_2^k, \ldots, a_p^k, k = 1, 2, \ldots m_i\}$ be m_i different sets of p nonnegative integers, each set containing at least one positive value. Then, as n_i tends to infinity, the statistic T_n, where

$$T_n = \sum_{i=1}^{c} \sum_{k=1}^{m_i} \bar{A}_n^{(i)2}(a_1^k, \ldots, a_p^k), \tag{2.5}$$

has asymptotically a noncentral chi-squared distribution with m degrees of freedom and with noncentrality parameter Δ_n where

$$m = \sum_{i=1}^{c} m_i \tag{2.6}$$

and

$$\Delta_n = \sum_{i=1}^{c} n_i \sum_{k=1}^{m_i} A_n^{(i)}(a_1^k, a_2^k, \ldots, a_p^k). \tag{2.7}$$

PROOF. Since the c samples are assumed to be independent, it is enough to prove that $\sum_k \bar{A}_n^{(i)2}(a_1^k, a_2^k, \ldots, a_p^k)$ has, asympototically, a noncentrality chi-squared distribution with m_i degrees of freedom and with the noncentrality parameter $n_i \sum_{k=1}^{m_i} A_n^{(i)2}(a_1^k, a_2^k, \ldots, a_p^k)$. Hence, it is enough to prove that the asymptotic joint distribution of

$$\{\bar{A}_n^{(i)}(a_1^k, a_2^k, \ldots, a_p^k)\ k = 1, 2, \ldots, m_i\}$$

is that of m_i independent random normal variables with expected values equal to $\{\sqrt{n_i}\, A_n^{(i)}(a_1^k, a_2^k, \ldots, a_p^k)\ k = 1, 2, \ldots, m_i\}$ and variances equal to one. The theorem now follows if we use the generalizations to the multisample case [3] of results on the asymptotic normality of the number of observations falling in the asymptotically sufficient groupings given by Weiss [6] for the test-of-fit problem.

It is a consequence of the above theorem that the statistic T_n (2.5) can be used to test our hypothesis H_o, which is equivalent to testing the hypothesis that the noncentrality parameter (2.7) is zero.

3. SOME BAYES TESTS AND THEIR ASYMPTOTIC PROPERTIES

In this section we describe some Bayes tests based on T_n (2.5) and investigate their asymptotic properties. These tests are the multivariate and multisample extensions of those given for the univariate test of fit in [7].

From the Fourier series expansion (2.1) for $r_{n_i}^i(z^1, z^2, \ldots, z^p)$ we have $\int_{C_p} \cdots \int r_{n_i}^{i2}(z^1, z^2, \ldots, z^p)\, dz^1 dz^2 \ldots, dz^p = \sum_{a_1=0}^{\infty} \sum_{a_2=0}^{\infty} \sum_{a_p=0}^{\infty} A_n^{(i)2}(a_1, a_2, \ldots, a_p)$.

Suppose we choose an m_i and assume that $A_n^{(i)}(a_1^k, a_2^k, \ldots, a_p^k) = 0$ for all the vectors not among those chosen. Just which sets $(a_1^k, a_2^k, \ldots, a_p^k)$ are to be chosen and the choice of m_i depend on the alternatives of interest and the requirements on the power function. These are discussed for the numerical example given in Section 4. We then have the following testing-hypothesis problem:

(3.1) $$H_o^{(1)} : \sum_{i=1}^{c} \sum_{k=1}^{m_i} A_n^{(i)2} (a_1^k, a_2^k, \ldots, a_p^k) = 0,$$

and we let the alternate hypothesis be

(3.2) $$H_A^{(1)} : \sum_{i=1}^{c} \sum_{k=1}^{m_i} A_n^{(i)2} (a_1^k, a_2^k, \ldots, a_p^k) = c_1 (> 0).$$

To test $H_o^{(1)}$ (3.1) against $H_A^{(1)}$ (3.2) suppose we assume a prior distribution which assigns a probability b to the point

$$\{A_n^i (a_1^1, a_2^1, \ldots, a_p^1), A_n^i (a_1^2, a_2^2, \ldots, a_p^2), \ldots, A_n^i (a_1^{m_i}, a_2^{m_i}, \ldots, a_p^{m_i})\}$$
$$= (0, 0, \ldots, 0) \text{ for } i = 1, 2, \ldots c.$$

and assigns probability $(1 - b)$ spread uniformly over the region $\sum_{i=1}^{c} \sum_{k=1}^{m_i} A_n^{(i)2} (a_1^k, a_2^k, \ldots, a_p^k) = c_1 > 0$. It is then well known that (e.g., see [2] the test

(3.3) $$T^{(1)} : \text{Reject } H_0^{(1)} \text{ (3.1) if } T_n > c(\alpha, m),$$

where $c(\alpha,m)$ is chosen to guarantee a level of significance equal to α, has the following asymptotic properties:

(i) It is a minimax test of level of significance α; i.e., it maximizes the minimum asymptotic power against alternatives of the form $H_a^{(1)}$ (3.2) among all tests with level of significance α.

(ii) It is a uniformly most powerful unbiased test of level of significance α.

(iii) It is a uniformly most powerful invariant test of level of significance α.

(iv) Power considerations: In spite of the above desirable properties, it is seen that for $\sum_{i=1}^{c} \sum_{k=1}^{m_i} A_n^{(i)2} (a_1^k, a_2^k, \ldots, a_p^k)$ fixed, the asymptotic power of this test tends to the asymptotic level of significance α as m_i increases.

This means that there is no test for which the asymptotic power stays above α subject to the sole restriction $\Delta_n \geqslant c_1 > 0$ under $H_A^{(1)}$ (3.2). So we have to limit our class of alternatives in some sense in order to have the asymptotic power stay above the asymptotic level of significance. One natural thing to do is to bound $\int \left[\frac{\partial}{\partial z^a} r_{n_i}^i (z^1, \ldots, z^p)\right]^2 dz^1 \ldots, dz^p$ by a constant different from c_1 under $H_A^{(1)}$. Such a discussion is carried out in [5] for the univariate test-of-fit problem and it extends to our problem as well.

Now we describe a second Bayes test $T^{(2)}$ (say), for which the alternate hypothesis takes a special form. Suppose that for each sample i, $i = 1, 2, \ldots, c$, and for each n we have a set of $s_i(n)$ functions

$$r_{n_i}^i (z_n^2, \ldots, z^p; 1), r_n^i(z^1, z^2, \ldots, z^p; 2), \ldots, r_n^i [z^1, z^2, \ldots, z^p; s_i (n)]$$

and let our set of alternate hypotheses consist of distributions with densities of the form

$$1 + \sum_{j=1}^{s_i(n)} \theta_j^i r_n^i (z^1, z^2, \ldots, z^p; j)$$

for some unknown constants $\theta_1^i, \theta_2^i, \ldots, \theta_{s_i(n)}^i$; $i = 1, 2, \ldots, c$. Then we write $A_n^{(i)}(a_1, a_2, \ldots, a_p) = \sum_{j=1}^{s_i(n)} \theta_j^i A_n(a_1, a_2, \ldots, a_p; j)$, where

$$A_n^{(i)}(a_1, a_2, \ldots, a_p; j) =$$

$$\int_{C_p} \cdots \int r_n^i(z^1, z^2, \ldots, z^p; j) 2^{D(a_1, a_2 \cdots a_p)} \prod_{L=1}^{p} \cos(a_L \Pi z^L)\, dz^1 \ldots, dz^p.$$

Thus, we are testing the hypothesis that

$$E[\bar{A}_n^{(i)}(a_1^k, a_2^k, \ldots, a_p^k)] = 0 \quad i = 1, 2, \ldots, c,$$

against the alternative

$$E[\bar{A}_n^{(i)}(a_1^k, a_2^k, \ldots, a_p^k; j)] = \sqrt{n_i} \sum_{j=1}^{S_i(n)} \theta_j^i A_n^i(a_1^k, a_2^k, \ldots, a_p^k; j).$$

Now let the $t(n)$ vectors $V_n^i(q)$ $(q = 1, 2, \ldots, t(n))$, where

$$V_n^i(q) = V_n^i(a_1^1, a_2^1, \ldots, a_p^1; q), V_n^i(a_1^2, a_2^2, \ldots, a_p^2; q)$$

$$, \ldots, V_n^i(a_1^{m_i}, \ldots\ldots, a_p^{m_i}; q),$$

be an orthonormal basis for the vector space generated by the $s_i(n)$ vectors

$$(A_n^{(i)}(a_1^1, a_2^1, \ldots, a_p^1; j), A_n^{(i)}(a_1^2, a_2^2, \ldots, a_p^2, j),$$

$$\ldots, A_n^{(i)}(a_1^{m_i}, \ldots, a_p^{m_i}; j); j = 1, \ldots, s_i(n)).$$

Then under our alternate hypothesis,

$$E[\bar{A}_n^{(i)}(a_1^k, a_2^k, \ldots, a_p^k)] = \sqrt{n_i} \sum_{q=1}^{t(n)} b_q^i v_n^i(a_1^k, a_2, \ldots, a_p^k, q)$$

for some unknown constants $b_q^{(i)}$, $q = 1, 2, \ldots, t(n)$ and $i = 1, 2, \ldots, c$. Then

$$\sum_{k=1}^{m_i} A_n^{(i)2}(a_1^k, a_2^k, \ldots, a_p^k) = \sum_{q=1}^{t(n)} b_q^{(i)2},$$

which is obtained from the orthonormal property of the V_n's. Thus, we have reduced our problem to that of testing

(3.4) $$H_0^{(2)} : \sum_{i=1}^{c} \sum_{q=1}^{t(n)} b_q^{(i)2} = 0,$$

and we let the alternate hypothesis be

(3.5) $$H_a^{(2)} : \sum_{i=1}^{c} \sum_{q=1}^{t(n)} b_q^{(i)2} = c_2, \; c_2 > 0.$$

To test $H_0^{(2)}$ (3.4) against $H_A^{(2)}$ (3.5), let us assume a prior distribution which assigns a probability b to the point $\sum_i \sum_q b_q^{(i)2} = 0$ over the space of the $b_q^{(i)}s$ and distributes the remaining probability uniformly over the region $\sum_i \sum_q b_q^{(i)2} = c_2$. Then the usual simple calculations show that a Bayes Decision rule relative to the above prior distribution is given by

(3.6) $$T^{(2)} : \text{Reject } H_0^{(2)} \text{ (3.4) if } T_n^{(2)} > c\,[\alpha; ct(n)],$$

where

(3.7) $$T_n^{(2)} = \sum_{i=1}^{c} n_i \sum_{q=1}^{t(n)} \hat{b}_q^{(i)2}$$

$$\hat{b}_q^{(i)} = 1/\sqrt{n_i} \sum_{k=1}^{m_i} \bar{A}_n^{(i)}(a_1^k, \ldots, a_p^k)\, V_n^i(a_1^k, \ldots, a_p^k; q),$$

and $c[\alpha; ct(n)]$ is chosen to guarantee a level of significance equal to α. It is easily seen that $\sqrt{n_i}\,\hat{b}_q^{(i)}$ is asymptotically normally distributed with mean equal to $\sqrt{n_i}\, b_q^{(i)}$, covariances equal to zero, and variance equal to one. Since the c samples are assumed to be independent, we see that when $H_0^{(2)}$ (3.4) is true $T_n^{(2)}$ has a central chi-squared distribution with $ct(n)$ degrees of freedom. Thus, for the test $T^{(2)}$ (3.6), $c[\alpha; ct(n)]$ is the appropriate value from the chi-squared table with $ct(n)$ degrees of freedom. This test has the following asymptotic properties.

(i) It is a minimax test of level significance α against the class of alternatives of the form $H_A^{(2)}$ (3.5).

(ii) It is a uniformly most powerful unbiased test of level of significance α.

(iii) It is a uniformly most powerful invariant test of level of significance α.

(iv) Power considerations: Suppose that under the alternate hypothesis, the true density is given by

$$1 + r_n^*(i)\,(z^1, z^2, \ldots, z^p).$$

Define $A_n^{*(i)}(a_1, a_2, \ldots, a_p)$ to be the same function of $r_n^{*(i)}(z^1, z^2, \ldots, z^p)$ as $A_n^{(i)}(a_1, a_2, \ldots, a_p)$ is of $r_n^i(z^1, z^2, \ldots, z^p)$. We then write the vector

$$[A_n^{*(i)}(a_1^1, a_2^1, \ldots, a_p^1),\ A_n^{*(i)}(a_1^2, a_2^2, \ldots, a_p^2), \ldots,\ A_n^{*(i)}(a_1^{m_i}, \ldots, a_p^{m_i})]$$

as equal to

$$\sum_{q=1}^{t(n)} b_q^{*(i)}\, V_n^i(q) + \bar{b}_n V_n^{*(i)}, \tag{3.8}$$

where $V_n^{*(i)}$ is orthogonal to $[V_n^i(1), V_n^i(2), \ldots, V_n^i(t(n))]$. Then $T_n^{(2)}$ has a noncentral chi-squared distribution with $ct(n)$ degrees of freedom and noncentrality parameter $\Delta_n^{(2)}$, where

$$\Delta_n^{(2)} = \sum_{i=1}^{c} n_i \sum_{q=1}^{t(n)} b_q^{*(i)2}. \tag{3.9}$$

From the definition of $T_n^{(2)}$ (3.7) and $\Delta_n^{(2)}$ (3.9) we have, as $t(n)$ increases,

$$\frac{[T_n^{(2)} - ct(n) - \Delta_n^{(2)}]}{\sqrt{2ct(n) + 4\,\Delta_n^{(2)}}}$$

is asymptotically distributed as a normal random variable with zero mean and unit variance, and hence we conclude that $[\Delta_n^{(2)}/\sqrt{ct(n)}]$ must be bounded away from zero in order to have the asymptotic power stay above the asymptotic level of significance, against alternatives of the form $\{1 + r_n^{*(i)}(z^1, z^2, \ldots, z^p)\}$.

We now construct a third Bayes test, which we call an "all-purpose" test, in the sense that it does not depend on any special set of alternatives. We assume a prior distribution which assigns a probability $q\,(q > 0)$ to the point

$$E[\bar{A}_n^{(i)}(a_1^k, a_2^k, \ldots, a_p^k)] = 0,\ k = 1, 2, \ldots m_i,\ i = 1, 2, \ldots, c,$$

and assigns probability $\dfrac{1-q}{m(m-1)}$ to each of the $m(m-1)$ points with

$$E[\bar{A}_n^{(i)}(a_1^{k_1}, a_2^{k_1}, \ldots, a_p^{k_1})] = -\Delta_n^{(3)},$$

$$E[\bar{A}_n^{(i)}(a_1^{k_2}, a_2^{k_2}, \ldots, a_p^{k_2})] = \Delta_n^{(3)},$$

$$E[\bar{A}_n^{(i)}\ (a_1^{k_3}, a_2^{k_3}, \ldots, a_p^{k_3})] = 0,$$

for $k_3 \neq k_1, k_2$, that we get as we vary k_1 and k_2 from 1 to m_i with $k_1 \ k_2$, where $m = \sum_{i=1}^{c} m_i$. The usual simple calculations then show that a Bayes Decision rule relative to the above prior distribution is given by

(3.10) $T^{(3)}$: Reject $H_0^{(1)}$ (3.1) if $T_n^{(3)}$ is "too large," where

(3.11) $$T_n^{(3)} = \sum_{i=1}^{c} n_i \sum_{\substack{k_1 \neq k_2 \\ k_1, k_2 = 1}}^{m_i} \exp\{\Delta_n^{(3)} [\bar{A}_n^{(i)}\ (j_1^{k_1}, \ldots, j_p^{k_1}) - \bar{A}_n^{(i)}\ j_1^{k_2}, \ldots, j_p^{k_2})]\}.$$

To investigate the properties of $T^{(3)}$ (3.10) we expand $T_n^{(3)}$, using the expansion $e^x = 1 + x + x^2/2 + \ldots$, and use the fact that

$$\sum_{k_1 \neq k_2} [\bar{A}_n^{(i)}\ (j_1^{k_1}, \ldots, j_p^{k_1}) - \bar{A}_n^{(i)}\ (j_1^{k_2}, \ldots, j_p^{k_2})]^S$$

is zero if s is an odd integer. Then, if $\Delta_n^{(3)}$ tends to zero as n increases, we find that the asymptotic properties of the test $T^{(3)}$ (3.10) are the same as those of $T^{(1)}$ (3.3).

Now let the true density under the alternate hypothesis be

$$\{1 + \bar{r}_n^{(i)}\ (z^1, \ldots, z^n)\}.$$

Define $\bar{B}_n^{(i)}\ (a_1, a_2, \ldots, a_p)$ to be the same function of $\bar{r}_n^{\,i}\ (z^1, z^2, \ldots, z^n)$ as $A_n^{(i)}\ (a_1, a_2, \ldots, a_p)$ is of $r_n^i\ (z^1, z^2, \ldots, z^n)$. Then $T_n^{(3)}$, given by (3.11) with $\bar{A}_n^{(i)}$'s replaced by $\bar{B}_n^{(i)}$'s has asymptotically a noncentral chi-squared distribution with m degrees of freedom and noncentrality parameter $\bar{\Delta}_n^{(3)}$ (say), where

$$\bar{\Delta}_n^{(3)} = \sum_{i=1}^{c} n_i \sum_{k=1}^{m_i} B_n^{(i)2}\ (a_1^k, \ldots, a_p^k);$$

and hence the distribution of $\dfrac{(T_n^{(3)} - m - \bar{\Delta}_n^{(3)})}{\sqrt{2m + 4\,\bar{\Delta}_n^{(3)}}}$ is asymptotically normal with zero mean and unit variance.

For the same level of significance, the asymptotic power of the all purpose test $T^{(3)}$ (3.10) will be greater than the asymptotic power of $T^{(2)}$ (3.6) if

$$\frac{\sum_{i=1}^{c} n_i \sum_{k=1}^{m_i} B_n^{(i)2}\ (a_1^k, a_2^k, \ldots, a_p^k)}{\sum_{i=1}^{c} m_i} > \frac{\sum_{i=1}^{c} n_i \sum_{q=1}^{t(n)} b_q^{*(i)2}}{ct(n)}.$$

We see that if $\bar{b}_n$ in (3.8) is nonzero, then

$$\sum_{q=1}^{t(n)} b_q^{*(i)2} < \sum_{k=1}^{n_i} B_n^{(i)2}\ (a_1^k, a_2^k, \ldots, a_p^k).$$

But, on the other hand, $m = (\sum_{i=1}^{c} m_i)$ is likely to be higher than $ct(n)$. Thus, even though the test $T^{(3)}$ based on $T_n^{(3)}$ is an all-purpose test, it is less powerful than the test $T^{(2)}$ based on $T_n^{(2)}$ against alternatives used in its construction.

4. A NUMERICAL EXAMPLE

We will now give a numerical example to illustrate a test procedure based on T_n (2.5) for the case when $c = 3$ and $p = 2$. Following the notations used so far, $\{X_{ij},\ j = 1,2,\ldots,\ n_i\}$ are n_i independent observations from a population with continuous cumulative distribution function F_i for $i = 1,2$, and 3. Let the null hypothesis be

$$H_0 : F_1 = F_2 = F_3\ (= F \text{ say}),$$

where F is the cumulative distribution function of a bivariate normal random variable with zero mean, unit variance, and correlation coefficient equal to 0.4.

Since the multivariate normal distributions differ from one another only in their location parameters or covariance matrices, we will consider an alternate hypothesis where F_i's have the same location parameters but have different covariance matrices. So let our alternate hypothesis be

H_1 : F_i is the cumulative distribution function of a bivariate normal random variable with zero mean, unit variance, and correlation coefficient equal to ρ_i for $i = 1,2$, and 3, where $\rho_1 = 0.6$, $\rho_2 = 0.4$, and $\rho_3 = 0.2$.

For convenience let us take $n_i = 100$ for $i = 1,2$, and 3. By using the method given by Box and Muller [1], we then generate a Monte Carlo sample of 100 independent pairs $\{(X_{ij}^1, X_{ij}^2),\ j = 1,2,\ldots,\ 100\}$ from each of the three bivariate normal distributions considered under H_1. Application of Rosenblatt's transformation based on F to $\{X_{ij}\}$ then gives

$$Z_{ij}^1 = 1 - \phi(X_{ij}^1), \tag{4.1}$$

$$Z_{ij}^2 = 1 - \phi\left[\frac{X_{ij}^2 - 0 \cdot 4X_{ij}^1}{\sqrt{1 - 0.16}}\right] \tag{4.2}$$

for $i = 1,2,3$ and $j = 1,2,\ldots,\ 100$, where

$$\phi(t) = \int_t^\infty (2\Pi)^{-1/2} \exp(-y^2/2)\, dy.$$

We then divide the unit square C_2 into nine (say) equal parts. The center

$$L_n(b) = [L_n^1(b),\ L_n^2(b)]$$

of each subsquare $S_n(b)$ and the number $N_n^i(b)$ of observations Z_{ij} from the i^{th} sample falling in the subsquare $S_n(b)$ are given in Table 1. Our next step is to compute $\bar{A}_n^{(i)}(a_1,a_2)$ for $i = 1,2,3$, and we proceed as follows. First we have to decide on which sets (a_1,a_2) are to be chosen. Since we are interested in alternatives which differ from the null hypothesis only in their values for the correlation coefficients, and since the above transformations (4.1) and (4.2) use the correlation coefficient only for transforming the second co-ordinate, it is clear that we should emphasize a_2 more than a_1. So let us take (a_1,a_2) for each sample as follows:

Sample 1 : $(a_1,a_2) = \{(0,1), (1,0), (1,1), (0,2)\}$

Sample 2 : $(a_1,a_2) = \{(0,1), (1,0), (1,1)\}$

Sample 3 : $(a_1,a_2) = \{(0,1), (1,0), (1,1), (0,2)\}$.

We then compute $\bar{A}_n^{(i)}(a_1,a_2)$ by using (2.4), and the results are given in Table 2. Table 3 gives the value of our test statistic T_n.

Table 1

b	$[L_n^1(b),\ L_n^2(b)]$	$N_n^{1(b)}$ $(\rho_1 = 0.6)$	$N_n^{2(b)}$ $(\rho_2 = 0.4)$	$N_n^{3(b)}$ $(\rho_3 = 0.2)$
1	$\left(\frac{1}{6}, \frac{1}{6}\right)$	9	8	10
2	$\left(\frac{1}{2}, \frac{1}{6}\right)$	7	7	6
3	$\left(\frac{5}{6}, \frac{1}{6}\right)$	6	6	16
4	$\left(\frac{1}{6}, \frac{1}{2}\right)$	15	10	18
5	$\left(\frac{1}{2}, \frac{1}{2}\right)$	19	30	16
6	$\left(\frac{5}{6}, \frac{1}{2}\right)$	16	12	12
7	$\left(\frac{1}{6}, \frac{5}{6}\right)$	11	12	7
8	$\left(\frac{1}{2}, \frac{5}{6}\right)$	5	7	10
9	$\left(\frac{5}{6}, \frac{5}{6}\right)$	12	8	5

Table 2

(a_1,a_2)	$\bar{A}_n^{(1)}(a_1,a_2)$	$\bar{A}_n^{(2)}(a_1,a_2)$	$\bar{A}_n^{(3)}(a_1,a_2)$
(0,1)	$-\frac{6}{10}\frac{\sqrt{3}}{2}$	$-\frac{6}{10}\frac{\sqrt{3}}{2}$	$\frac{\sqrt{3}}{2}$
(1,0)	$\frac{1}{10}\frac{\sqrt{3}}{2}$	$\frac{4}{10}\frac{\sqrt{3}}{2}$	$\frac{1}{5}\frac{\sqrt{3}}{2}$
(1.1)	$-\frac{6}{10}$	$-\frac{3}{10}$	$-\frac{6}{5}$
(0,2)	$-\frac{5}{2}\sqrt{2}$	–	$-\frac{19}{10}\sqrt{2}$

Table 3

$\bar{A}_n^{(1)^2}(a_1,a_2)$	$\bar{A}_n^{(2)^2}(a_1,a_2)$	$\bar{A}_n^{(3)^2}(a_1,a_2)$	$T_n = \sum_{i=1}^{3} \bar{A}_n^{i^2}(a_1,a_2)$
13.90	0.87	10.62	25.39

But from Theorem 1 we know that T_n (2.5) has a central chi-squared distribution under H_0 with 11 degrees of freedom. Referring to the chi-squared tables, we get

$$X^{2}_{11,\ 0.005} = 26.75$$

and

$$X^{2}_{11,\ 0.01} = 24.72.$$

Hence, using T_n as our statistic to test H_0 against H_1, we get a level of significance between 0.005 and 0.01.

REFERENCES

[1] Box, G.E.P., and M. E. Muller, "A Note on the Generation of Random Normal Deviates," Annals of Mathematical Statistics *29* 610-611 (1958).

[2] Ferguson, T.S., *Mathematical Statistics: A Decision Theoretic Approach* (Academic Press, New York, 1967).

[3] Ramani, K.V., "Some Bayes Tests and their Asymptotic Properties for the Multivariate Multisample Goodness of Fit Tests," Technical Report 319(b), Department of Operations Research, Cornell University, Ithaca, N.Y. (1977).

[4] Rosenblatt, M. "Remarks on a Multivariate Transformation," Annals of Mathematical Statistics, *23*, 470-472 (1952).

[5] Weiss, L, "The Asymptotic Sufficiency of a Relatively Small Number of Order Statistics in Tests of Fit," Annals of Statistics, *2*, 795-802 (1974).

[6] Weiss, L., "Multivariate Tests of Fit Using Asymptotically Sufficient Grouping," Naval Research Logistics Quarterly, *23*, 629-638 (1976).

[7] Weiss, L., "Asymptotic Properties of Bayes Tests of Nonparametric Hypothesis," in *Statistical Decision Theory and Related Topics II: Proceedings of a Symposium held at Purdue University, May 17-19, 1976*, S.S. Gupta and D. S. Moore, eds., pp. 439-450 (Academic Press, New York, 1977).

MULTIPLE-ATTRIBUTE DECISION MAKING WITH PARTIAL INFORMATION: THE EXPECTED-VALUE CRITERION

Johnnie R. Charnetski

College of Administration and Business
Louisiana Tech University
Ruston, Louisiana

Richard M. Soland

Department of Operations Research
The George Washington University
Washington, D. C.

ABSTRACT

We consider the multiple-attribute decision problem with finite action set and additive utility function. We suppose that the decision maker cannot specify nonnegative weights for the various attributes which would resolve the problem, but that he/she supplies ordinal information about these weights which can be translated into a set of linear constraints restricting their values. A bounded polytope W of feasible weight vectors is thus determined. Supposing that each element of W has the same chance of being the "appropriate one," we compute the expected utility value of each action. The computation method uses a combination of numerical integration and Monte Carlo simulation and is equivalent to finding the center of mass of the bounded polytope W. Comparisons are made with another criterion already presented, the comparative hypervolume criterion, and two small examples are presented.

1. INTRODUCTION

The work reported here builds upon that in our previous paper [3], in which both problems of multiple-attribute decision making with partial information and the comparative hypervolume criterion for resolving such problems were defined and discussed. Our objectives here are to present a second decision criterion, the expected-value criterion, which may be used in such problems, and to contrast it with the aforementioned comparative hypervolume criterion.

For completeness, and because not much space is required, we shall not rely on a knowledge of the material contained in [3]; all that is needed will be developed here. All necessary definitions and background will be presented in this introductory section. The expected-value criterion (EVC) and its relation to the comparative hypervolume criterion (CHC) will be presented in the following section, and in Section 3 we will discuss the numerical procedure needed to compute the values required by the EVC. In the final section two numerical examples will be presented. We now move to a development of the notation and a framing of the problem.

We are concerned with finite-action decision problems in which the decision maker(s) (DM) must choose one action from a finite set A of feasible actions. Each action $a_i \in A$, $i = 1, \ldots, m$, has been evaluated with respect to each criterion (or attribute) c_j in a finite set $C \equiv \{c_1, \ldots, c_n\}$ of criteria. Define s_{ij} to be the raw score of action a_i with respect to attribute c_j; for each c_j the scores s_{ij} may be on either an ordinal or an interval scale.

Our supposition is that the DM wishes to choose an action on the basis of maximum utility (or value), where the utility of action a_i is $\mathbf{u}^*(a_i) = \mathbf{u}(s_{i1}, \ldots, s_{in})$. We further assume that an additive form for the utility function $\mathbf{u}$ is appropriate (see Keeney and Raiffa [5] for necessary and sufficient conditions for such additivity) so that

$$\mathbf{u}^*(a_i) = \mathbf{u}(s_{i1}, \ldots, s_{in}) = \sum_{j=1}^{n} u_j(s_{ij}). \tag{1}$$

As indicated by Keeney and Raiffa [5, p. 116], the functions $\mathbf{u}^*$ and $u_j (j = 1, \ldots, n)$ may be scaled to the interval [0, 1] so that $u_j(s_{ij})$ may be written as $w_j v_{ij}$, where v_{ij} is the relative value of the raw score s_{ij} based on the set of all scores s_{kj} with respect to attribute c_j, and each w_j is a positive weight. If, as is natural, the v_{ij} are also scaled from zero to one for each attribute c_j, then the weights w_j must sum to unity, and the utility of action a_i now has the form

$$\mathbf{u}^*(a_i) = \sum_{j=1}^{n} w_j v_{ij}. \tag{2}$$

A significant advantage of this form is the relative ease with which the DM can determine the v_{ij} values (see Keeney and Raiffa [5, Section 3.7] for further discussion and an illustrative example). Let us define $\mathbf{V}$ as the m by n matrix of v_{ij} values; for each column j the values v_{ij} lie between 0 and 1 inclusive, with at least one of them taking each of these extreme values. With V_i the i th row of $\mathbf{V}$ and $\mathbf{w}$ a column vector of the weights w_j, expression (2) becomes

$$\mathbf{u}^*(a_i) = V_i \mathbf{w}. \tag{3}$$

If the DM provides a weight vector $\mathbf{w} = \mathbf{w}^*$, then (3) provides the basis for selection of an action. In many cases, however, the DM may not be willing (or able) to provide a particular $\mathbf{w}^*$. In the case of an individual decision maker this may be simply due to the fact that he/she cannot articulate his/her preferences with the needed precision. In the case of a group of decision makers, there may be considerable disagreement about the appropriate weight vector $\mathbf{w}^*$ (see Sengupta, *et al.* [6] for further treatment of this case).

Our hypothesis in the following, therefore, is that the DM does not provide a particular $\mathbf{w}^*$, but rather that he/she provides information which allows the construction of a set $\mathbf{W}$ of vectors to be considered. More formally, we define $\mathbf{W}$ as the set of weight vectors that the DM deems feasible, i.e., may be appropriate ones in light of his/her subjective feelings (or in light of their range of agreement if there are several decision makers). We shall confine ourselves to the case in which $\mathbf{W}$ is characterized by a number of linear equality and/or inequality constraints on the components w_j of $\mathbf{w}$; see [3] for a discssion of how appropriate linear constraints may be construcuted. These constraints must include the normalization constraint $\sum w_j = 1$ and the nonnegativity constraints $w_j \geqslant 0$, $j = 1, \ldots, n$ (each w_j must actually be positive, but nothing is lost, in a numerical sense, by writing $w_j \geqslant 0$). In a general sense, therefore, we may write

$$\mathbf{W} = \{\mathbf{w} \in \mathbf{E}^n | \mathbf{A}\mathbf{w} \leqslant \mathbf{b}\}, \tag{4}$$

where $\mathbf{A}$ is an s by n matrix of constraint coefficients and $\mathbf{b}$ is the s by 1 vector of right-hand-side values. $\mathbf{W}$ is thus a bounded polytope in $\mathbf{E}^n$.

We mention two special cases: (1) that in which $\mathbf{W} = \{\mathbf{w}^*\}$, i.e., it is known that $\mathbf{w} = \mathbf{w}^*$, which we call the case of *complete information*; and (2) that in which

$$\mathbf{W} = \bar{\mathbf{W}} \equiv \{\mathbf{w} \in \mathbf{E}^n | \mathbf{w} \geqslant 0, \sum w_j = 1\}, \tag{5}$$

which we call the case of *no information*. We refer to the general case, which falls between these two extremes, as the case of *partial information*, and thus characterize the decision problem with which we deal as one of multiple-attribute decision making with partial information.

The problem we address here is formally equivalent to that considered by Fishburn, Murphy, and Isaacs [4] in the context of decision making under uncertainty with incomplete knowledge of the probabilities. They mention six possible approaches, several of which have elements in common with the development which follows. Also see Charnetski [1].

2. THE EXPECTED-VALUE CRITERION

Given $\mathbf{W}$, it is not immediately clear as to how it should be used to aid the DM in selecting an action a_i for implementation. In [3] we suggested the use of the comparative hypervolume criterion (CHC), which effectively determines for each action $a_i \in A$ the relative measure r_i of that subset $\mathbf{H}_i$ of $\mathbf{W}$ in which a_i has the highest utility value of all actions. We then select as "best" the action with the highest comparative hypervolume r_i. In effect, this treats all elements $\mathbf{w} \in \mathbf{W}$ as equally probable, and then choses an action which is "most likely" to have the highest utility value. Precise definitions of r_i and $\mathbf{H}_i$ will be given below.

A deficiency of the CHC is the possibility that no action a_i may have a very high value of r_i (e.g., $r_1 = 0.26$, $r_2 = 0.24$, $r_3 = 0.20$, $r_4 = 0.18$, $r_5 = 0.12$), so that the basis for choice becomes less clear. Moreover, the relative measures r_i are clearly affected by the number of action choices. Another argument against the CHC is that it does not account for the fact that an action may not be very likely to yield the highest utility value, but may nevertheless yield a high utility value for many $\mathbf{w} \in \mathbf{W}$, or "on the average." If this sounds somewhat like an argument in favor of expected value as opposed to maximum likelihood, in the context of a simple decision problem with a finite number of possible states of nature, it is, because the situation *is* directly analogous. Again treating all $\mathbf{w} \in \mathbf{W}$ as equally probable, one could compute the expected utility value for each action, and then choose the action with the highest such expected utility value. This is the expected-value criterion (EVC), which we shall now proceed to formalize.

Let $K = \int_{\mathbf{W}} d\mathbf{w}$ by the hypervolume of $\mathbf{W}$. Then $K^{-1}(d\mathbf{w})$ is the probability measure over $\mathbf{W}$ with which we must deal. We may obtain the expected utility value $\bar{\mathbf{u}}^*(a_i)$ of action a_i as

$$\begin{aligned}
\bar{\mathbf{u}}^*(a_i) &= K^{-1} \int_{\mathbf{W}} \mathbf{u}^*(a_i)\, d\mathbf{w} \\
&= K^{-1} \int_{\mathbf{W}} (V_i \mathbf{w})\, d\mathbf{w} \\
&= K^{-1} \int_{\mathbf{W}} \left(\sum_j v_{ij} w_j \right) d\mathbf{w} \\
&= K^{-1} \sum_j v_{ij} \int_{\mathbf{W}} w_j\, d\mathbf{w} \\
&= \sum_j u_i \bar{w}_j \\
&= V_i \bar{\mathbf{w}},
\end{aligned} \tag{6}$$

where $\bar{\mathbf{w}} = (\bar{w}_1, \ldots, \bar{w}_n)^T$ is the mean weight vector (or center of mass of $\mathbf{W}$). The EVC says to select the action a_i with the highest expected utility value $\bar{\mathbf{u}}^*(a_i)$. In the next section we shall deal with the actual computation of the $\bar{\mathbf{u}}^*(a_i)$ values, but first we will explore some relationships between the EVC and the CHC.

For the CHC we use the definitions

$$\mathbf{H}_i \equiv \{\mathbf{w} \in \mathbf{W} \mid V_i\mathbf{w} \geq V_k\mathbf{w},\ k = 1, \ldots, m\}, \tag{7}$$

$$r_i \equiv K^{-1} \int_{\mathbf{H}_i} d\mathbf{w}. \tag{8}$$

According to the EVC, we select action a_i if and only if $V_i\bar{\mathbf{w}} \geq V_k\bar{\mathbf{w}}$, $k = 1, \ldots, m$. From (7) we see that this is equivalent to choosing a_i if and only if $\bar{\mathbf{w}} \in \mathbf{H}_i$, i.e., the EVC says to choose an action a_i such that $\bar{\mathbf{w}} \in \mathbf{H}_i$.

The fact that $\bar{\mathbf{u}}^*(a_i) \geq \bar{\mathbf{u}}^*(a_k)$ for all k if and only if $\bar{\mathbf{w}} \in \mathbf{H}_i$ leads one to conjecture that perhaps the EVC and the CHC are equivalent in that $\bar{\mathbf{u}}^*(a_i) \geq \bar{\mathbf{u}}^*(a_k)$ for all k if and only if $r_i \geq r_k$ for all k. Example 1 of the appendix shows this conjecture to be false. A weaker, but similar, conjecture involves the comparison of only 2 actions. Defining

$$\mathbf{H}_{ik} = \{\mathbf{w} \in \mathbf{W} \mid V_i\mathbf{w} \geq V_k\mathbf{w}\} \text{ and } r_{ik} = K^{-1} \int_{\mathbf{H}_{ik}} d\mathbf{w}, \tag{9}$$

the conjecture is that $\bar{\mathbf{u}}^*(a_i) \geq \bar{\mathbf{u}}^*(a_k)$ if and only if $r_{ik} \geq r_{ki}$. This is false in general (see Example 2 of the appendix), but it is true when the dimension d of $\mathbf{W}$ is 1, i.e., whenever $\mathbf{W}$ is a line segment in $\mathbf{E}^n$. If we neglect the case of perfect information, this is always the case if $n = 2$.

3. COMPUTATION OF THE $\bar{\mathbf{u}}^*(a_i)$

The value of the $\bar{\mathbf{u}}^*(a_i)$, as defined by (6), cannot be calculated analytically because of the integration over $\mathbf{W}$ required. Although $\mathbf{W}$ has an explicit characterization as a polytope, only in very exceptional situations will it be possible to perform the required integration exactly. (Note that the cases of complete information ($\bar{\mathbf{w}} = \mathbf{w}^*$) and no information ($\bar{\mathbf{w}} = (1/n, \ldots, 1/n)$) are two such situations. For the general case we turn to Monte Carlo simulation as a means of obtaining accurate approximations of the $\bar{\mathbf{u}}^*(a_i)$.

From line 2 or 3 of (6), as well as from its interpretation, it is seen that $\bar{\mathbf{u}}^*(a_i)$ is the expected value of a linear function defined over the bounded polyhedral subset $\mathbf{W}$ of $\mathbf{E}^n$. In [2] we gave a detailed Monte Carlo procedure for estimating such expected values. This approach would require one Monte Carlo simulation for each $a_i \in A$, but this is less efficient than estimating the mean weight vector $\bar{\mathbf{w}}$ and then using the result $\bar{\mathbf{u}}^*(a_i) = V_i\bar{\mathbf{w}}$. The procedure of [2] is easily adapted to give $\bar{\mathbf{w}}$ directly and with only one Monte Carlo run; we give this adaptation here because it then provides a simple procedure for determining the center of mass $\bar{\mathbf{w}}$ of any bounded polyhedral set $\mathbf{W}$ defined by a finite set of linear inequalitites.

Recall that $\mathbf{W} = \{\mathbf{w} \in \mathbf{E}^n \mid \mathbf{A}\mathbf{w} \leq \mathbf{b}\}$, where $\mathbf{A}$ is s by n. Let A_l be the lth row of $\mathbf{A}$. Let γ be a nondegenerate vertex of $\mathbf{W}$ (obtained after some simplex pivots or by perturbation of a degenerate vertex if necessary), and let $\{\mathbf{v}^p \mid p = 1, \ldots, d\}$ be the set of unit vectors defining the directions of the edges of $\mathbf{W}$ incident to γ. Because γ is nondegenerate, d is the dimension of $\mathbf{W}$ ($d \leq n - 1$). Thus, the set Φ of all unit vectors ϕ originating at γ and directed into $\mathbf{W}$, may be written as

$$\Phi = \left\{\phi \mid \phi = \sum_{p=1}^{d} e_p\mathbf{v}^p / \| \sum_{p=1}^{d} e_p\mathbf{v}^p \|,\ \text{all } e_p \geq 0\right\}, \tag{10}$$

and for arbitrary $\phi \in \Phi$ we define the scalar

$$\lambda(\phi) = \min_i\{\lambda_i(\phi) | \lambda_i(\phi) = (b_i - A_i\gamma)/A_i\phi, \text{ for } A_i\phi > 0\}. \tag{11}$$

$\lambda(\phi)$ is just the "width" of $\mathbf{W}$ in the direction ϕ. With a sequence of uniform (pseudo)random numbers we can generate random unit vectors $\phi \in \Phi$ by use of (10). Numbering these $\phi^1, \phi^2, \ldots$, we obtain a sample of size M, and may estimate $\bar{\mathbf{w}}$ by

$$\bar{\mathbf{w}} = \gamma + \left[d \sum_{k=1}^{M} \{\lambda(\phi^k)\}^{d+1} \phi^k\right] / (d+1) \sum_{k=1}^{M} \{\lambda(\phi^k)\}^d. \tag{12}$$

Care must be taken in the generation of the random unit vectors $\phi \in \Phi$ through use of (10) in order that this sampling of elements of Φ be unbiased. The most natural procedure is that in which, if we take successive outputs of a (pseudo)random-number generator, the successive e_p are independent and uniformly distributed on the interval (0, 1). If we take

$$\phi = \sum_{p=1}^{d} e_p \mathbf{v}^p / \| \sum_{p=1}^{d} e_p \mathbf{v}^p \|, \tag{13}$$

the sampling technique actually yields a biased choice for ϕ. We may avoid this bias, however, if only sequences $(e_1, \ldots, e_d)$ with a special property are used, and the others discarded. If all pairs of unit vectors $(\mathbf{v}^p, \mathbf{v}^q)$, $1 \leqslant p < q \leqslant d$, form acute angles (verified by the inner product $\langle \mathbf{v}^p, \mathbf{v}^q \rangle \geqslant 0$), then it is necessary to discard all sequences $(e_1, \ldots, e_d)$ for which $\| \sum_p e_p \mathbf{v}^p \| > 1$. This follows from the argument that all points of the form $\gamma + \sum_p e_p \mathbf{v}^p$ are equally likely, so that if only those within the unit hypersphere centered at γ (for which $\| \sum_p e_p \mathbf{v}^p \| \leqslant 1$) are retained, then all unit vectors ϕ defined by (13) are equally likely. On the other hand, if any pair of unit vectors $(\mathbf{v}^p, \mathbf{v}^q)$ forms an obtuse angle, then it is necessary to further restrict the generated sequences $(e_1, \ldots, e_d)$ to those for which $\| \sum_p e_p \mathbf{v}^p \| \leqslant t$, where $t = (1 - \cos^2\beta)^{1/2}$ and β is the largest angle between pairs of unit vectors $(\mathbf{v}^p, \mathbf{v}^q)$. This is because in this case of an obtuse angle not all points within the unit hypersphere centered at γ can be obtained as $\gamma + \sum_p e_p \mathbf{v}^p$ for some sequence $(e_1, \ldots, e_d)$. But $\gamma + \sum_p e_p \mathbf{v}^p$ *can* generate all points within the hypersphere of radius t centered at γ.

Clearly, the necessity of discarding some sequences $(e_1, \ldots, e_d)$ decreases the overall computational efficiency of the procedure, perhaps significantly in those cases for which $t \ll 1$. In the event that $\{\mathbf{v}^p\}$ yields some obtuse angles at γ and that $t \ll 1$, it might be computationally more efficient to reject γ as the origin and search for another vertex. One might limit the search to those vertices adjacent to γ, stopping if the new sets $\{\mathbf{v}^p\}$ form acute angles only, or, failing that, choosing the vertex which allows the maximum value of $t < 1$. Hence, additional research on generating the $\phi \in \Phi$ is clearly desirable. For the two example problems presented in the next section the generation method indicated above was used along with the detailed methods given in [2] and [3].

4. EXAMPLE PROBLEMS AND COMPUTATIONAL RESULTS

We consider the following decision problem (from Sengupta, *et al* [6]). Table 1 describes the action choices and their evaluations s_{ij} with respect to three criteria. (The s_{ij} were left unscaled in this example to facilitate comparisons with the results given in [6].)

Table 1

Action \ Criteria	c_1	c_2	c_3
a_1	8.460	0.000	11.280
a_2	9.646	-5.250	8.194
a_3	2.118	10.000	4.824

Additionally, we assume the set **W** to be specified by:

$$w_1 \geqslant 0.4, \qquad 0.1 \leqslant w_2 \leqslant 0.7,$$

$$\sum_{j=1}^{3} w_j = 1, \text{ and } w_j \geqslant 0 \quad 0 \text{ for every } j.$$

A FORTRAN code executed on an IBM 370/148 machine at the Louisiana Tech Computational Center produced the results shown in Table 2.

Table 2

Action	Expected Value	Hypervolume
a_1	6.656	0.771
a_2	5.400	0.043
a_3	4.687	0.186

These expected values and hypervolumes are based on a sample of 1500 random vectors. The CPU time was 16.66 s, or approximately 1.11 s per 100 vectors generated.

For the second example problem, we have the decision matrix **V** shown in Table 3.

Table 3

Action \ Criteria	c_1	c_2	c_3	c_4	c_5
a_1	1.00	0	0.48	0.66	0.11
a_2	0.65	1.00	0	0.72	0.31
a_3	0.83	0.35	1.00	0.28	0
a_4	0	0.05	0.43	1.00	0.18
a_5	0.73	0.53	0.94	0	1.00
a_6	0.14	0.03	0.43	0.37	0.77

We assume the following constraints specify the set **W**:

$$\begin{aligned}
w_1 - w_2 - w_3 &\leqslant 0,\\
-w_2 + w_3 + w_4 + w_5 &\leqslant 0,\\
-w_1 + w_2 &\leqslant 0,\\
-w_2 + w_3 &\leqslant 0,\\
-w_3 + w_4 &\leqslant 0,\\
-w_4 + w_5 &\leqslant 0,\\
w_5 &\geqslant 0.05,
\end{aligned}$$

and

$$\sum_{j=1}^{5} w_j = 1, \quad w_j \geqslant 0, \quad \text{for every } j.$$

Again using a sample of 1500 random vectors, the results in Table 4 were obtained (the CPU time was 56.9 seconds).

Table 4

Action	Expected Value	Hypervolume
a_1	0.532	0.0
a_2	0.605	0.137
a_3	0.629	0.259
a_4	0.221	0.0
a_5	0.651	0.604
a_6	0.225	0.0

5. APPENDIX

Here we present the two counterexamples mentioned in Secton 2. Consider first the 3 by 2 adjusted payoff matrix **V**:

$$\mathbf{V} = \begin{bmatrix} 10 & 0 \\ 0 & 10 \\ 5 + 10\epsilon & 5 + 10\epsilon \end{bmatrix},$$

where ϵ is a small positive number. Let $\mathbf{W} = \overline{\mathbf{W}}$, the case of no information. Then, clearly, $\overline{\mathbf{w}} = (0.5, 0.5)^T$, so that $\bar{\mathbf{u}}^*(a_1) = \bar{\mathbf{u}}^*(a_2) = 5$, whereas $\bar{\mathbf{u}}^*(a_3) = 5 + 10\epsilon$. The $\mathbf{H}_i$ are:

$$\mathbf{H}_1 = \left\{(w_1, w_2)^T \mid 0.5 + \epsilon \leqslant w_1 \leqslant 1, \ w_1 + w_2 = 1\right\},$$

$$\mathbf{H}_2 = \left\{(w_1, w_2)^T \mid 0 \leqslant w_1 \leqslant 0.5 - \epsilon, \ w_1 + w_2 = 1\right\},$$

and

$$\mathbf{H}_3 = \left\{(w_1, w_2)^T \mid 0.5 - \epsilon \leqslant w_1 \leqslant 0.5 + \epsilon, \ w_1 + w_2 = 1\right\}.$$

Hence, $r_1 = r_2 = 0.5 - \epsilon$, whereas $r_3 = 2\epsilon$. Thus $\bar{\mathbf{u}}^*(a_3) \leqq \bar{\mathbf{u}}^*(a_k)$ for all k, but $r_3 << r_k$ for $k = 1, 2$.

For the second counterexample, take **V** as follows:

$$\mathbf{V} = \begin{bmatrix} 7 & 0 & 10 \\ 10 & 3 & 3 \\ 0 & 10 & 0 \end{bmatrix}.$$

Again take $\mathbf{W} = \overline{\mathbf{W}}$. Then $\overline{\mathbf{w}} = (1/3, 1/3, 1/3)^T$, so that $\bar{\mathbf{u}}^*(a_i) = 17/3$, $16/3$, and $10/3$ for $i = 1, 2, 3$, respectively. The definition (9) of $\mathbf{H}_{12}$ gives:

$$\mathbf{H}_{12} = \left\{(w_1, w_2, w_3)^T \mid w_1 + w_2 \leqslant 0.7; \ \text{all } w_j \geqslant 0\right\},$$

and then straightforward computaton yields $r_{12} = 0.49$. Hence $r_{21} = 0.51$ (since $r_{ik} + r_{ki} = 1$) and $r_{21} > r_{12}$ although $\bar{\mathbf{u}}^*(a_2) < \bar{\mathbf{u}}^*(a_1)$.

If $\mathbf{W}$ is a line segment in $\mathbf{E}^n$, with $\bar{\mathbf{w}}$ therefore at the midpoint, then it is clear that $\bar{\mathbf{u}}^*(a_i) \geqslant \bar{\mathbf{u}}^*(a_k)$ if and only if $r_{ik} \geqslant r_{ki}$, since $\bar{\mathbf{w}} \in \mathbf{H}_{ik}$ if and only if $r_{ik} \geqslant r_{ki}$.

REFERENCES

[1] Charnetski, J.R., "Bayesian Decision Making with Ordinal Information," Operations Research *25*, 889-892 (1977).

[2] Charnetski, J.R., and R.M. Soland, "Statistical Measures for Linear Functions on Polytopes," Operations Research *24*, 201-204 (1976).

[3] Charnetski, J.R., and R.M. Soland, "Multiple-Attribute Decision Making with Partial Information: The Comparative Hypervolume Criterion," Naval Research Logistics Quarterly *25*, 279-288 (1978).

[4] Fishburn, P.C., A.H. Murphy, and H.H. Isaacs, "Sensitivity of Decisions to Probability Estimation Errors: A Reexamination," Operations Research *16*, 254-267 (1968).

[5] Keeney, R.L., and H. Raiffa, *Decisions with Multiple Objectives: Preferences and Value Tradeoffs* (Wiley, New York, 1976).

[6] Sengupta, S.S., M.L. Podrebarac, and T.D.H. Fernando, "Probabilities of Optima in Multi-Objective Linear Programs," in *Multiple Criteria Decision Making*, J.L. Cochrane and M. Zeleny, eds. (University of South Carolina Press, Columbia, S.C., 1973), pp. 217-235.

OPTIMAL STATE-DEPENDENT PRICING POLICIES FOR A CLASS OF STOCHASTIC MULTIUNIT SERVICE SYSTEMS

R. K. Gupta

School of Business Administration and Commerce
Memorial University of Newfoundland
St. John's, Newfoundland, Canada

V. Srinivasan

Graduate School of Business
Stanford University
Stanford, California

P. L. Yu

School of Business
University of Kansas
Lawrence, Kansas

ABSTRACT

This paper models a k-unit service system (e.g., a repair, maintenance, or rental facility) with Poisson arrivals, exponential service times, and no queue. If we denote the number of units that are busy as the state of the system, the state-dependent pricing model formalizes the intuitive notion that when most units are idle, the price (i.e., the service charge per unit time) should be low, and when most units are busy, the price should be higher than the average. A computationally efficient algorithm based on a nonlinear programming formulation of the problem is provided for determination of the optimal state-dependent prices. The procedure ultimately reduces to the search on a single variable in an interval to determine the unique intersection point of a concave increasing function and a linear decreasing function. The algorithm takes, on the average, only about 1/2 second per problem on the IBM 360/65 (FORTRAN G Compiler). A discrete optimal-control approach to the problem is shown to result in essentially the same procedure as the nonlinear-programming formulation. Several properties of the optimal state-dependent prices are given. Comparisons of the optimal values of the objective function for the state-dependent and state-independent pricing policies show that the former is, on the average, only about 0.7% better than the latter, which may explain partly why state-dependent pricing is not prevalent in many service systems. Potential generalizations of the model are discussed.

INTRODUCTION

This paper deals with the short-run pricing decisions for multiunit service systems such as repair or maintenance facilities, equipment rentals, car rentals, motel rentals, and the like. Due

to the stochastic nature of the requests for such services and the probabilistic nature of service times, the number of units that are busy, referred to as the *state* of the system, also exhibits stochastic variations. When most of the units are idle, it makes intuitive sense that the firm should lower its price (i.e., service charge per unit time) so as to attract more customers to the system. If, however, most of the units are already busy, the firm may want to charge a higher than normal price, since in such a case the firm can afford to wait for better-paying customers. Such a pricing policy, where the price depends on the state of the system, will be referred to as a *state-dependent* pricing policy. In contrast, a pricing policy in which the price does not depend on the state of the system will be referred to as a *state-independent* pricing policy.

The state-dependent pricing framework may also be potentially useful as an internal-control mechanism for many not-for-profit service systems (e.g., public-sector and/or military maintenance facilities). For such systems state-dependent pricing would have the effect of restricting services to more urgent needs when the system is very busy and encouraging less urgent needs (e.g., preventive maintenance) when the system is relatively idle. Consequently, a better and more balanced utilization of the system is likely to result. Other stochastic systems such as consulting firms, job shops, and bank loan services also fit the basic description of such service systems, i.e., in the short run there is a constrained amount of some resource which can be "rented"; because of the stochastic variations in the state of the system, it may be profitable to follow a state-dependent pricing policy and change the "rental rate" depending on the current state of the system. For brevity, we will hereafter refer to such a stochastic system as a "rental" system and refer to the prices as "rental rates" although, as discussed in the previous examples, the approach is general enough to include a variety of multiunit stochastic service systems.

In practice, state-dependent pricing is observed only for small rental systems, where the number of units that can be rented or the total amount of the constrained resource than can be rented is small, e.g., a consultant (or a small consulting firm) quoting different rates for his consulting services, depending on how busy he is, or small firms bidding for contracts. (In fact, the bid price in the model by Kortanek, Soden and Sodaro [16] depends explicitly on the "opportunity cost" of the constrained resource.) Generally, we do not find that big car-rental agencies or motels practice state-dependent pricing. One of the interesting results of this paper is that, as the number of units for rent becomes large, the optimal state-dependent pricing policy is only marginally better than the optimal state-independent policy. This, when considered in conjunction with the disadvantages of the state-dependent policy, such as greater inventory-information costs, potential customer dissatisfaction, and, in some cases, potential legal problems (such as in apartment renting; it is, however, perfectly legal in most other situations) may explain why we do not observe state-dependent pricing for "large" rental agencies.

For analytical tractability, we shall assume that the return time of each rental unit in our model follows an exponential distribution. It is also assumed that the customers arrive according to a Poisson process. However, depending on the rental rate (or price), the customers may elect not to rent a unit. We shall start our analysis with the assumption that the rental charge is proportional to the time for which the unit is out on rent. Later, we shall show that our analysis carries through to the case where the rental price is the sum of a fixed charge and a variable charge that is proportional to the rental time.

Our model may be characterized as a Markovian decision process (or Markov renewal program). A number of articles have been published in this area, although most of them do not deal with the optimal-pricing problem we are considering; for instance, see Blackwell [1,2], Denardo [6], Denardo and Fox [7], Derman [8], Fox [9], Howard [12,13], Jewell [14], Keilson [15], Low [17], Ross and Lippman [19], and references cited in these articles. Except for Keilson [15], who supplies an elegant integer-programming method for a single-unit rental system,

all the above works use mainly a dynamic-programming approach to obtain the optimal decisions. The present research started as an attempt to extend Keilson's results to multiunit rental systems. The results reported in this paper are related to those of Low [17]. The main differences between our research and that of Low are: (i) Low uses a lump sum (or fixed charge) payment for price, while we assume that the rental charge is proportional to the time for which the unit is out on rent. In most rental and similar situations (e.g., car rentals, computer processing, or bank loans) charges are proportional to the rental-time; i.e., a variable-charge scheme prevails.) In a later section, we also cover the general case of a fixed charge plus a variable charge that is proportional to the rental time. (ii) Low assumes that the lump-sum prices are chosen from a fixed, finite number of available prices, while we allow the prices to be chosen from an interval (infinitely many choices). (Low, later in his paper, does extend his analysis to an interval as the decision space, but he avoids the discussion of computational problems.) (iii) Low basically uses a dynamic-programming approach, while our approaches are based on nonlinear programming and discrete optimal-control. Our resulting bounded-value difference equations can be solved easily to yield the optimal pricing policies. (iv) We obtain stronger results of strict monotonicity of the optimal prices, while Low's results are limited to monotonicity only. (v) We provide economic insights into various results.

As is usual in the Markovian decision problems, we are interested in stationary policies, i.e., policies which, although state-dependent, do not change with time. The existence of such policies is assured in our case, because our assumptions satisfy the sufficient conditions as described by Fox [9].

The emphasis of this paper is on providing a computationally efficient algorithm for the determination of optimal state-dependent rental rates as well as on offering qualitative insights into the problem by characterizing some important properties of the optimal solution. The basic mathematical programming model for the determination of optimal state-dependent prices is formulated in Section 1. Some preliminary properties of the optimal solution are derived in Section 2. Section 3 outlines some alternate solution strategies. In Section 4, the optimality conditions are derived by relaxation of one of the constraints of the mathematical program. These optimality conditions are equivalent to the solution of a nonlinear difference equation with boundary constraints, and they ultimately reduce to the unique intersection of two curves monotonic in opposite directions. The algorithm determines this intersection by the bisection, or Bolzano, search [21, p. 122]. Some important properties of the optimal solution are derived. An alternate solution strategy using the discrete-maximum principle of optimal-control theory [3,4] is developed in the appendix and is shown to result in the same set of optimality conditions as are found in Section 4. In Section 5, the optimal state-independent solution is explored, and the results are compared to the optimal state-dependent solution. Finally, some extensions to the basic model are explored in Section 6.

1. FORMULATION OF THE STATE-DEPENDENT PRICING MODEL

We consider a rental system with k units. Since we are interested only in the short-term pricing decisions, the number of units k can be assumed to be fixed. Let us define the index sets:

$$K = \{0, 1, 2, \ldots, k\},$$

$$K' = \{0, 1, 2, \ldots, k-1\}, \text{ and}$$

$$K'' = \{1, 2, \ldots, k\}.$$

The return of the rented units is assumed to have an exponential distribution with parameter η; i.e., the service time for a rented unit is exponential, with mean $1/\eta$. The customers arrive to inquire about the rates according to a Poisson process with arrival rate λ. We define the parameter ρ to be

$$\rho = \lambda/\eta, \quad (\rho > 0).$$

Customers who arrive when all the k units are busy leave (balk) the system and go to a competitor. Thus, there is never any queue, and the system can be fully described by the number of units i ($i \in K$) that are out on rent. We will refer to i as the state of the system. Let v_i, $i \in K'$, be the rental rate charged for units rented when the system is at state i; i.e., if a unit is rented when the state is i, and the unit is kept for t time units, then the rental charge will be $v_i t$. Let $\mathbf{v}$ denote the k-component vector with elements $\{v_i\}$. The probability that a customer who arrives will, in fact, rent a unit is modelled as

$$p(v_i) = 1 - av_i, \quad \text{for } i \in K' \ (a > 0). \tag{1}$$

Consequently, if $v_i = 0$ the customer will rent it with probability one, whereas if v_i becomes as large as $1/a$, then this probability drops to zero. Thus, (1) models the traditional downward-sloping demand curve. In order that $0 \leqslant p(v_i) \leqslant 1$, we require that

$$0 \leqslant v_i \leqslant 1/a, \quad \text{for } i \in K'. \tag{2}$$

The problem is to determine the optimal values for the k decision variables $v_0, v_1, v_2, \ldots, v_{k-1}$ so as to maximize D, the expected steady-state revenue of the rental system per unit time.

To express the objective D in terms of the decision variables $\{v_i\}$, we first define π_i to be the steady-state probability that the system will be in state i ($i \in K$) corresponding to the decision vector $\mathbf{v}$. Let $\vec{\pi}$ denote the $(k+1)$-component vector with elements π_i. Thus,

$$\pi_i \geqslant 0, \quad \text{for } i \in K, \tag{3}$$

and

$$\sum_{i \in K} \pi_i = 1. \tag{4}$$

To express the π_i's in terms of the v_i's, we use the result that, in steady state, the upward transition rate from i to $i+1$ should be the same as the downward transition rate from $i+1$ to i for $i \in K'$. Since λ is the arrival rate and $p(v_i)$ is the probability of renting a unit given an arrival, the effective arrival rate is $\lambda\, p(v_i)$. (Note that probabilistic independence has been assumed so that the multiplication operation is justified.) Thus,

$$\pi_i \lambda(1 - av_i) = \pi_{i+1}(i+1)\eta, \quad \text{for } i \in K', \text{ or} \tag{5}$$

$$\pi_{i+1} = \rho(1 - av_i)\pi_i/(i+1), \quad \text{for } i \in K'.$$

Consequently, given a set of $\{v_i\}$ and ρ, the $(k+1)$ values $\{\pi_i\}$ can be determined by solving the k equations (5) together with (4).

To determine D, we first find $G(\Delta t)$, the expected revenue accruing from the rentals made in the time interval $(t, t + \Delta t)$ as $t \to \infty$ (note that in steady state G depends only on Δt and not on t), so that

$$D = \lim_{\Delta t \to 0} G(\Delta t)/\Delta t. \tag{6}$$

To compute $G(\Delta t)$, let us assume for the moment that the system is in state i at time t. Then the expected number of units rented in the time interval $(t, t + \Delta t)$ is $(\lambda \Delta t)\ (1 - \mathrm{av}_i)$, and for each unit rented at rental rate v_i the expected service time is $1/\eta$, so that the revenue accruing from the rentals made during $(t, t + \Delta t)$ is $(\lambda \Delta t)\ (1 - \mathrm{av}_i)\ (\mathrm{v}_i/\eta)$. Since in steady state the probability that the system will be in state i is π_i, the expected revenue accruing from the rentals made in $(t, t + \Delta t)$ is

$$G(\Delta t) = \sum_{i \in K'} \rho \mathrm{v}_i (1 - \mathrm{av}_i)\ \pi_i \Delta t. \tag{7}$$

Finally, using (6) we obtain

$$D = \sum_{i \in K'} \rho \mathrm{v}_i (1 - \mathrm{av}_i)\ \pi_i. \tag{8}$$

(A more rigorous derivation of (8) is given in [10, pp. 18-23].) Consequently, the problem of determining the optimal $\{\mathrm{v}_i\}$ can be stated as Problem P_0,

$$\text{Maximize } D(\mathbf{v}) = \sum_{i \in K'} \rho \mathrm{v}_i (1 - \mathrm{av}_i)\ \pi_i \tag{9}$$

subject to

$$\pi_{i+1} = \rho(1 - \mathrm{av}_i)\ \pi_i/(i + 1), \quad \text{for } i \in K', \tag{10}$$

$$\sum_{i \in K} \pi_i = 1, \tag{11}$$

$$\pi_i \geqslant 0, \text{ for } i \in K, \tag{12}$$

and

$$0 \leqslant \mathrm{v}_i \leqslant 1/a, \text{ for } i \in K'. \tag{13}$$

To reduce the number of parameters of the system, let us consider the transformed decision variables $\mathbf{u} = a\mathbf{v}$, i.e.,

$$u_i = \mathrm{av}_i, \text{ for } i \in K', \tag{14}$$

and maximize $F = aD$ rather than the objective D (these are equivalent to changing the unit of measurement of the rental rate and the objective function by the same factor a) so as to obtain Problem P_1,

$$\text{(15)} \qquad \text{Maximize } F(\mathbf{u}) = \sum_{i \in K'} \rho u_i (1 - u_i) \pi_i$$

subject to

$$\text{(16)} \qquad \pi_{i+1} = \rho(1 - u_i)\pi_i/(i + 1), \text{ for } i \in K',$$

$$\text{(17)} \qquad \sum_{i \in K} \pi_i = 1,$$

$$\text{(18)} \qquad \pi_i \geqslant 0, \text{ for } i \in K,$$

and

$$\text{(19)} \qquad 0 \leqslant u_i \leqslant 1, \text{ for } i \in K'.$$

2. SOME PRELIMINARY PROPERTIES OF THE OPTIMAL SOLUTION

THEOREM 1: There exists an optimal solution $\mathbf{u}^*$ to Problem P_1 with $F^* = F(\mathbf{u}^*) > 0$.

PROOF: The constraints (16) through (19) define a compact set (closed and bounded) in the variables $(\mathbf{u}, \vec{\pi})$. The constraint set is nonempty, since $u_i = 1/2$, for $i \in K'$, defines a feasible solution. Consequently, since the objective function (15) is continuous in $(\mathbf{u}, \vec{\pi})$, there exists an optimum solution to Problem P_1. Now it is easily verified that for the feasible solution $u_i = 1/2$, for $i \in K'$, $\vec{\pi} > \vec{0}$ and $F > 0$. Consequently $F^* > 0$.

COROLLARY 1: For any optimal solution $\mathbf{u}^*$, $u_0^* < 1$.

PROOF: Assume the contrary, that $u_0^* = 1$. From (16) through (18) it is easily verified that $\pi_i^* = 0$ for $i = 1, 2, \ldots, k$. From (15) it then follows that $F^* = 0$, which contradicts Theorem 1.

To prove the other important properties of the optimal solution, we reformulate Problem P_1 in terms of only the decision variables $\mathbf{u}$. From (16),

$$\text{(20)} \qquad \pi_i = \pi_0 (\rho^i/i!) \prod_{j=0}^{i-1} (1 - u_j) \text{ for } i \in K,$$

where we follow the usual convention that $\prod_{j=0}^{-1} (1 - u_j) = 1$.

Consequently, from (17) it follows that

$$\text{(21)} \qquad \pi_0 = 1/\left[\sum_{i \in K} (\rho^i/i!) \prod_{j=0}^{i-1} (1 - u_j)\right].$$

Substituting (20) and (21) in the objective function (15), Problem P_1 is equivalent to Problem P_2,

$$\text{(22)} \qquad \underset{\mathbf{u}}{\text{Maximize}}\ F(\mathbf{u}) = \frac{\sum_{i \in K'} (\rho^{i+1}/i!)\, u_i \prod_{j=0}^{i} (1 - u_j)}{\sum_{i \in K} (\rho^{i}/i!) \prod_{j=0}^{i-1} (1 - u_j)}$$

subject to

$$\text{(23)} \qquad 0 \leqslant u_i \leqslant 1, \text{ for } i \in K'.$$

THEOREM 2: Any optimal solution $\mathbf{u}^*$ to Problem P_1 (or P_2) has the property that $u_i^* < 1$ for $i \in K'$.

PROOF: Assume the contrary. Let m be the smallest index for which $u_m^* = 1$. Corollary 1 implies that $m \geqslant 1$. Thus $u_i^* < 1$ for $0 \leqslant i \leqslant m-1$. From (22) it is seen that the objective F is independent of the values of u_i^* for $i > m$. Therefore, without loss of generality the values $u_i^* = 1$ can be assigned to the states $i > m$. For this $\mathbf{u}^*$, (22) becomes

$$\text{(24)} \qquad F(\mathbf{u}^*) = \frac{\sum_{i=0}^{m-1} (\rho^{i+1}/i!)\, u_i^* \prod_{j=0}^{i} (1 - u_j^*)}{\sum_{i=0}^{m} (\rho^{i}/i!) \prod_{j=0}^{i-1} (1 - u_j^*)}.$$

Defining

$$A = \sum_{i=0}^{m-1} (\rho^{i+1}/i!)\, u_i^* \prod_{j=0}^{i} (1 - u_j^*)$$

and

$$B = \sum_{i=0}^{m} (\rho^{i}/i!) \prod_{j=0}^{i-1} (1 - u_j^*),$$

we can rewrite (24) as

$$\text{(25)} \qquad F(\mathbf{u}^*) = A/B.$$

In the above expression for B, the term for $i = 0$ simplifies to unity. Renumbering the remaining indices $i = 1, 2, \ldots, m$ to $i = 0, 1, 2, \ldots m-1$, we get

$$B = 1 + \sum_{i=0}^{m-1} [\rho^{i+1}/(i+1)!] \prod_{j=0}^{i} (1 - u_j^*).$$

Now consider a solution $\mathbf{u}'$ which is identical to $\mathbf{u}^*$ in all its components, except

$$\text{(26)} \qquad u'_m = [1/(m+1)] + \left\{ \max_{0 \leqslant i \leqslant m-1} [u_i^* (i+1)/(m+1)] \right\}.$$

Since $0 \leqslant u_i^* < 1$ for $0 \leqslant i \leqslant m-1$, it follows that the term within the braces is, strictly, less than $m/(m+1)$, so that $0 < u'_m < 1$. For the solution $\mathbf{u}'$,

$$F(\mathbf{u}') = (A + C)/(B + D), \tag{27}$$

where A and B are as defined before, and

$$C = (\rho^{m+1}/(m!)\ u'_m\ (1 - u'_m) \prod_{j=0}^{m-1} (1 - u_j^*)$$

and

$$D = [\rho^{m+1}/(m+1)!]\ (1 - u'_m) \prod_{j=0}^{m-1} (1 - u_j^*).$$

Note that $F(\mathbf{u}') > F(\mathbf{u}^*)$ iff $(A + C)/(B + D) > A/B$ or iff $C/D > A/B$ or iff $B\ (C/D) - A > 0$, since A,B,C, and D are all strictly positive. Since $C/D = (m+1)\ u'_m$, it follows that

$$B(C/D) - A = (m+1)\ u'_m + \sum_{i=0}^{m-1} \{[\rho^{i+1}/(i+1)!]\ (m+1)u'_m - (\rho^{i+1}/i!)u_i\} \prod_{j=0}^{i} (1 - u_j^*), \tag{28}$$

which is strictly positive, since $[u'_m\,(m+1)/(i+1)] - u_i^* > 0$, from (26), for $0 \leqslant i \leqslant m-1$. Thus we obtain $F(\mathbf{u}') > F(\mathbf{u}^*)$, which contradicts the statement that $\mathbf{u}^*$ is an optimal solution.

Intuitively, a charge $u_i = 1$ would mean that no customers would rent when the system is at state i (recall that the probability that a customer who arrives will, in fact, rent is $p = 1 - u_i$), so that $(k - i)$ rental units will *always* be idle. In other words, a firm with i units would make the same average revenue as another firm with the same λ and η but with $k > i$ units, which does not make sense in a stochastic environment.

COROLLARY 2: For any optimal solution to P_1 (or P_2), $\pi_i^* > 0$, for $i \in K$.

PROOF: The proof follows directly from Theorem 2 and equations (20) and (21).

3. SOLUTION STRATEGIES

Several solution strategies suggest themselves for obtaining the optimal $\mathbf{u}^*$. Problem P_1 can be solved directly as a nonlinear program by the gradient-projection method [11, pp. 328-331]. But since the nonlinear constraints (16) have to hold as strict equalities, one can not guarantee that the procedure would find the global optimum. Similarly, if we solve Problem P_2 by the gradient method [11, pp. 296-315], there is no guarantee that we will obtain the global optimum solution, since we have not been able to prove that the objective (22) is at least pseudoconcave in $\mathbf{u}$.

Problem P_1 can be formulated as a dynamic program if we define π_i as a state variable. To incorporate the constraints (17) and (18), we define $s_i = \sum_{j=i}^{k} \pi_j$ as the second state variable with $s_0 = 1$. At stage i, $0 \leqslant i \leqslant k-1$,

(i) State-Space: $\{(\pi_i, s_i): 0 \leqslant \pi_i \leqslant s_i \text{ and } 0 \leqslant s_i \leqslant 1\}$.
(ii) Stage-Transformation: $\pi_{i+1} = \rho(1-u_i)\,\pi_i/(i+1)$ and $s_{i+1} = s_i - \pi_i$.
(iii) Decision Space: $\{u_i: 0 \leqslant u_i \leqslant 1\}$.
(iv) Performance Index: $f(u_i, \pi_i) = \rho u_i (1-u_i)\pi_i$.
(v) Functional Equation:

$$T_i(\pi_i, s_i) = \max_{0 \leqslant u_i \leqslant 1} [f(u_i, \pi_i) + T_{i+1}(\pi_{i+1}, s_{i+1})],$$

$$\text{where } T_k(\pi_k, s_k) = \begin{cases} 0 & \text{if } \pi_k = s_k, \\ -\infty & \text{otherwise.} \end{cases}$$

The dynamic programming procedure starts from stage $k-1$ and proceeds backward to stage 0. The optimum solution to Problem P_1 is given by the maximum of $T_0(\pi_0, 1)$ over all π_0 such that $0 \leqslant \pi_0 \leqslant 1$.

Although the dynamic programming procedure is easily formulated, it is computationally tedious, since it involves two continuous state variables. A more efficient computational procedure is presented in the next section.

4. PROPERTIES OF THE OPTIMAL SOLUTION AND AN ALGORITHM

4.1. Optimality Conditions

In this section, we first reformulate Problem P_1 using only the $\{\pi_i\}$ variables. By Corollary 2, we can restrict our attention to only those solutions for which $\vec{\pi} > \vec{0}$ Given $\pi_i > 0$, we can rewrite (16) as

$$\rho(1-u_i)\,\pi_i = (i+1)\,\pi_{i+1}, \text{ for } i \in K', \tag{29}$$

so that

$$u_i = 1 - [(i+1)\,\pi_{i+1}/(\rho\pi_i)], \text{ for } i \in K'. \tag{30}$$

Substituting (29) and (30) in (15), we get

$$F(\mathbf{u}) = R(\vec{\pi}) = \sum_{i \in K'} (i+1)\,\pi_{i+1}\{1 - [(i+1)\pi_{i+1}/(\rho\pi_i)]\}.$$

Consequently, under the conditions that $\pi_i > 0$ for $i \in K$, we can rewrite (15) through (19) as Problem P_3,

$$\text{(31)} \qquad \underset{\vec{\pi}}{\text{Maximize}}\ R(\vec{\pi}) = \sum_{i \in K''} i\pi_i \{1 - [i\pi_i/(\rho\pi_{i-1})]\}$$

subject to

$$\text{(32)} \qquad \sum_{i \in K} \pi_i = 1,$$

$$\text{(33)} \qquad \pi_i > 0 \text{ for } i \in K,$$

and

$$\text{(34)} \qquad \pi_i/\pi_{i-1} \leqslant \rho/i \text{ for } i \in K''.$$

The constraints (34) correspond to $u_i \geqslant 0$ in (19). The constraints $\pi_i/\pi_{i-1} \geqslant 0$ for $i \in K''$, corresponding to $u_i \leqslant 1$ in (19), have been omitted, since they are implied by (33).

From the above transformation and Corollary 2, we obtain the result:

THEOREM 3: Every optimum solution $\mathbf{u}^*$ to P_1 defines an optimal solution $\vec{\pi}^*$ to P_3, and conversely. The optimal value of the objective function for P_1 and P_3 is the same.

Theorem 4 shows that the objective function for P_3 is concave.

THEOREM 4: The objective function $R(\vec{\pi})$ is concave in $\vec{\pi}$ over $\{\vec{\pi} | \vec{\pi} > \vec{0}\}$.

PROOF: Consider the i th term in the objective (31),

$$r_i(\pi_{i-1}, \pi_i) = i\pi_i \{1 - [i\pi_i/(\rho\pi_{i-1})]\}.$$

The Hessian matrix for r_i is given by

$$\frac{2i^2\pi_i}{\rho\pi_{i-1}} \begin{bmatrix} -\pi_i/\pi_{i-1}^2 & 1/\pi_{i-1} \\ 1/\pi_{i-1} & -1/\pi_i \end{bmatrix}.$$

Note that the first principal minors are negative and the determinant vanishes. Thus, r_i is concave in (π_i, π_{i-1}), and hence r_i is concave in $\vec{\pi}$. Since $R(\vec{\pi})$ is a sum of concave functions r_i, for $i \in K''$, the theorem follows.

Our solution strategy for solving P_3 is as follows: We will, for the time being, ignore the constraints (34) and obtain the solution to (31) through (33) by finding a solution which satisfies the Kuhn-Tucker necessary conditions to (31) through (33). Since $R(\vec{\pi})$ is concave, the solution which satisfies the necessary conditions is also globally optimal to (31) through (33). We will then show that the solution so obtained also satisfies (34), so that it is, in fact, an optimal solution to P_3.

If we denote the Lagrange multiplier associated with (32) as θ (no multipliers need be associated with (33), since these constraints hold as strict inequalities), the Lagrangian function for maximizing (31) subject to (32) and (33) becomes

AD-A070 635 OFFICE OF NAVAL RESEARCH ARLINGTON VA F/G 15/5
NAVAL RESEARCH LOGISTICS QUARTERLY. VOLUME 26, NUMBER 2.(U)
JUN 79

UNCLASSIFIED

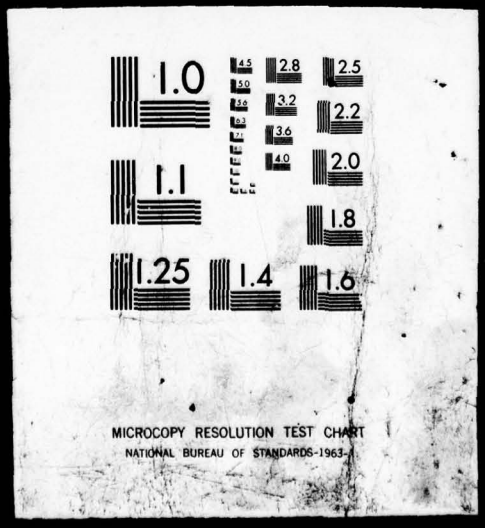
NL

2 OF 3
AD A070635

IFIED

2 OF 3

AD
A070635



$$L(\vec{\pi}, \theta) = \sum_{i \in K''} i\pi_i \{1 - [i\pi_i/(\rho\pi_{i-1})]\} - \theta \left[\left(\sum_{i \in K} \pi_i\right) - 1\right].$$

From Kuhn-Tucker theory [18] we obtain the following necessary and sufficient (because of concavity) conditions for optimality:

(35) (i) $\partial L/\partial \pi_i = 0$ for $i \in K$,

(36) (ii) $\sum_{i \in K} \pi_i = 1$,

(37) (iii) $\pi_i > 0$ for $i \in K$,

and

(iv) θ unrestricted in sign.

Inequalities (37) (the same as (33)) permit us to define $z_i = \pi_i/\pi_{i-1}$, for $i \in K''$, so that the $(k+1)$ conditions of (35) become

$$z_1^2 = \theta\rho \tag{38}$$

$$i\rho - 2i^2 z_i + (i+1)^2 z_{i+1}^2 = \theta\rho, \text{ for } i = 1, 2 \ldots, k-1, \tag{39}$$

and

$$k\rho - 2k^2 z_k = \theta\rho. \tag{40}$$

Thus, we seek a solution satisfying (36) through (40). We now explore some properties of such a solution.

4.2. Properties of the Optimal Solution

THEOREM 5. Suppose that $(\vec{\pi}, \theta)$ satisfies the conditions (36) through (40). Then $\theta = R(\vec{\pi})$.

PROOF: Multiplying (38) through (40) by the respective π_i's (recall that $z_i = \pi_i/\pi_{i-1}$) and summing, we obtain

$$(\pi_1^2/\pi_0) + \sum_{i=1}^{k-1} [i\rho\pi_i - 2i^2(\pi_i^2/\pi_{i-1}) + (i+1)^2(\pi_{i+1}^2/\pi_i)]$$

$$+ k\rho\pi_k - 2k^2(\pi_k^2/\pi_{k-1}) = \theta\rho \sum_{i \in K} \pi_i.$$

Expanding and rearranging the terms on the left-hand side, we obtain

$$\sum_{i \in K''} \rho i \pi_i \{1 - [i\pi_i/(\rho \pi_{i-1})]\} = \theta \rho \sum_{i \in K} \pi_i = \theta \rho.$$

Dividing throughout by ρ and using (31), we obtain the desired result.

Q.E.D.

To find the solution to (36) through (40), we first observe that (37) can be rewritten as

$$z_i = \pi_i/\pi_{i-1} > 0, \text{ for } i \in K''. \tag{41}$$

In the appendix we show that conditions (38) through (41), (30), and (34) also result from a discrete optimal-control approach [3,4] to solving Problem P_1.

From (36) and (41) it follows that

$$\pi_i = \pi_0 \prod_{j=1}^{i} z_j, \text{ for } i \in K, \tag{42}$$

and

$$\pi_0 = 1/\left(\sum_{i \in K} \prod_{j=1}^{i} z_j\right), \tag{43}$$

where we use the convention $\prod_{j=1}^{0} z_j = 1$. Consequently, the k-component vector $\mathbf{z} = \{z_i\}$ satisfying (38) through (41) (and the corresponding $\{\pi_i\}$ defined by (42) and (43)) satisfies the Kuhn-Tucker conditions for the problem of maximizing (31) subject to (32) and (33). From (38) and (41) it is clear that θ has to be nonnegative for the conditions (38) through (41) to be satisfied.

To solve the system (38) through (41), we first define

$$z_0 = 0, \tag{44}$$

so that, by setting $w_i = z_i^2$ in (38) and (39), we obtain

$$w_{i+1} = [(\theta - i)\rho + 2i^2 z_i]/(i+1)^2, \text{ for } i = 0, 1, 2, \ldots, k-1. \tag{45}$$

If the values of θ and z_i are such that the value for w_{i+1} obtained in (45) is nonnegative, then

$$z_{i+1} = +(w_{i+1})^{1/2} \text{ for } i = 0, 1, 2, \ldots, k-1. \tag{46}$$

Consequently, the system of equations (38) through (41) may be solved as follows: Given a value of $\theta \geq 0$, we set $z_0 = 0$ (44) and determine w_1 from (45) and z_1 from (46). In general, for $i = 0, 1, 2, \ldots, k-1$, given z_i we compute w_{i+1} using (45), and if w_{i+1} is nonnegative, we compute z_{i+1} from (46). From now on, this process of generating $z_1, z_2, \ldots, z_k$

will be referred to as the *forward-recursion procedure*. In Theorem 6 below we show that for each $i = 1, 2, \ldots k$ there exists a θ_i such that $w_i(\theta) \geqslant 0$ for $\theta \geqslant \theta_i$. Thus, $z_i(\theta)$ is well defined for $\theta \geqslant \theta_i$. Furthermore, it will be shown that for $\theta > \theta_i$, $z_i(\theta)$ is strictly positive, strictly increasing, and strictly concave in θ for $i = 1, 2, \ldots, k$. In particular, the above statements hold for $z_k(\theta)$, $\theta > \theta_k$, so that $z_k(\theta)$ can be pictorially represented as curve A in Figure 1.

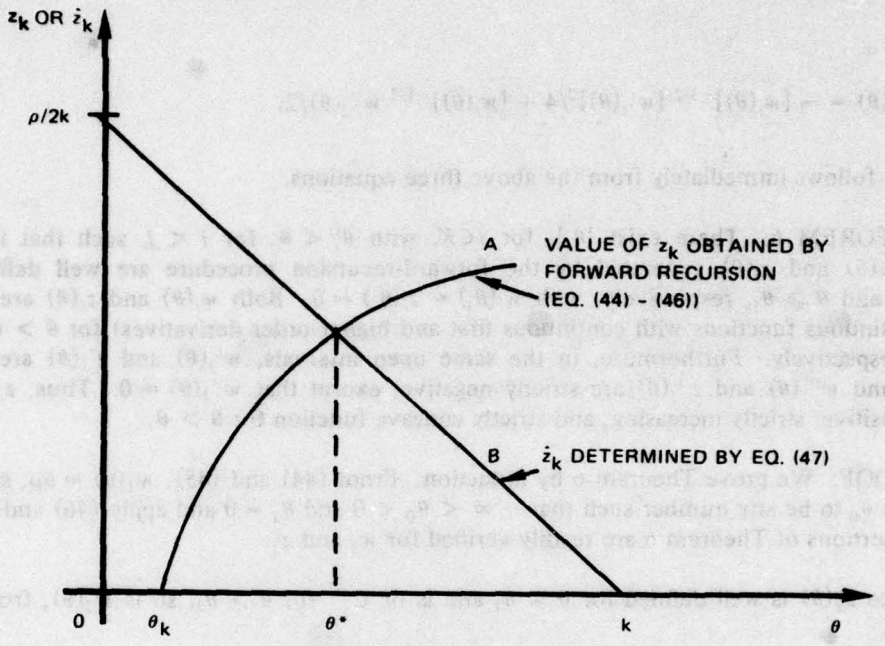


FIGURE 1. Determination of θ^*

Now let us consider (40). To distinguish the z_k obtained by the forward-recursion procedure from the z_k obtained from (40), we will call the latter $\hat{z}_k(\theta)$, so that

$$\hat{z}_k(\theta) = (k - \theta)\rho/(2k^2). \tag{47}$$

Thus, $\hat{z}_k(\theta)$ is a linear and strictly decreasing function of θ, as shown by line B in Figure 1. The value of $\theta = \theta^*$ satisfying (38) through (41) is thus obtained as the *unique* intersection point of these two monotone functions of θ monotonic in opposite directions (curve A and line B), as shown in Figure 1. The point θ^* can be found by the bisection (or Bolzano) search [21, p. 122] by searching in the region $\theta \in [0, k]$. Later, in Theorem 7, we shall prove the unique existence of θ^* and determine tighter bounds on θ^* so as to reduce the search effort. Once the value θ^* is determined, the z_i's for $i \in K$ can be determined from (44) through (46), the π_i's for $i \in K$ from (42) and (43) and the u_i's for $i \in K'$ from (30). In the lemmas and theorems below we denote the first and second derivatives with respect to θ by $'$ and $''$ respectively.

LEMMA 1: Consider the $w_i(\theta)$ and $z_i(\theta)$, for $i = 1, 2, \ldots, k$, obtained by the forward-recursion procedure. Suppose there exists a θ_i such that, for $\theta > \theta_i$, $w_i(\theta) > 0$, $w'_i(\theta) > 0$, and $w''_i(\theta) \leqslant 0$. Then, for $\theta > \theta_i$, $z_i(\theta) > 0$, $z'_i(\theta) > 0$, and $z''_i(\theta) < 0$.

PROOF: From (46),

$$z_i(\theta) = + [w_i(\theta)]^{1/2},$$

$$z'_i(\theta) = [w_i(\theta)]^{-1/2} w'_i(\theta)/2,$$

and

$$z''_i(\theta) = - [w_i(\theta)]^{-3/2} [w'_i(\theta)]^2/4 + [w_i(\theta)]^{-1/2} w''_i(\theta)/2.$$

Lemma 1 follows immediately from the above three equations.

THEOREM 6: There exist $\{\theta_i\}$, for $i \in K$, with $\theta_i < \theta_j$ for $i < j$, such that for each $i \in K''$, $w_i(\theta)$ and $z_i(\theta)$ generated by the forward-recursion procedure are well defined for $\theta \geqslant \theta_{i-1}$ and $\theta \geqslant \theta_i$, respectively, with $w_i(\theta_i) = z_i(\theta_i) = 0$. Both $w_i(\theta)$ and $z_i(\theta)$ are of C^∞ (i.e., continuous functions with continuous first and higher-order derivatives) for $\theta > \theta_{i-1}$ and $\theta > \theta_i$ respectively. Furthermore, in the same open intervals, $w'_i(\theta)$ and $z'_i(\theta)$ are strictly positive and $w''_i(\theta)$ and $z''_i(\theta)$ are strictly negative, except that $w''_1(\theta) = 0$. Thus, $z_i(\theta)$ is a strictly positive, strictly increasing, and strictly concave function for $\theta > \theta_i$.

PROOF: We prove Theorem 6 by induction. From (44) and (45), $w_1(\theta) = \theta\rho$, so that if we define θ_0 to be any number such that $-\infty < \theta_0 < 0$ and $\theta_1 = 0$ and apply (46) and Lemma 1, the assertions of Theorem 6 are readily verified for w_1 and z_1.

Since $z_1(\theta)$ is well defined for $\theta \geqslant \theta_1$ and is of C^∞ for $\theta > \theta_1$, so is $w_2(\theta)$, from (45). Note that

$$w'_2(\theta) = [\rho + 2z'_1(\theta)]/4$$

and

$$w''_2(\theta) = z''_1(\theta)/2.$$

Since $z'_1(\theta) > 0$ and $z''_1(\theta) < 0$ for $\theta > \theta_1$, it follows that $w'_2(\theta) > 0$ and $w''_2(\theta) < 0$ for $\theta > \theta_1$. Thus, $w_2(\theta)$ is a continuous, strictly increasing, and strictly concave function for $\theta > \theta_1$. From (45), $w_2(\theta) > 0$ when θ is sufficiently large, but $w_2(\theta_1) < 0$, since $\theta_1 = 0$ and $z_1(\theta_1) = 0$. Thus, there exists a unique $\theta_2 > \theta_1$ such that $w_2(\theta_2) = 0$ and $w_2(\theta) > 0$ for $\theta > \theta_2$. Consequently, from (46), $z_2(\theta)$ is well defined for $\theta \geqslant \theta_2$ and is of C^∞ for $\theta > \theta_2$. Furthermore, $w'_2(\theta) > 0$ and $w''_2(\theta) < 0$ for $\theta > \theta_2$, since these statements have already been shown to hold for $\theta > \theta_1$, and $\theta_2 > \theta_1$. From Lemma 1 it follows that $z_2(\theta) > 0$, $z'_2(\theta) > 0$, and $z''_2(\theta) < 0$ for $\theta > \theta_2$.

So far, we have established the existence of θ_i for $i = 0, 1, 2$ satisfying the conditions of Theorem 6. For the general induction step, we assume that we have proved the existence of θ_i for $i = 0, 1 \ldots, m < k$ $(m \geqslant 2)$ satisfying the assertions of the theorem. To complete the proof, we want to show that $\theta_{m+1} > \theta_m$ also exists, satisfying the conditions of the theorem.

Since $z_m(\theta)$ is well defined for $\theta \geqslant \theta_m$ and is of C^∞ for $\theta > \theta_m$, so is $w_{m+1}(\theta)$ (see (45)). From (45) and the induction hypothesis on $z_m(\theta)$, we have, for $\theta > \theta_m$,

$$w'_{m+1}(\theta) = [\rho + 2m^2 z'_m(\theta)]/(m+1)^2 > 0$$

and

$$w''_{m+1}(\theta) = 2m^2 z''_m(\theta)/(m+1)^2 < 0.$$

Thus, $w_{m+1}(\theta)$ is continuous, strictly increasing, and strictly concave for $\theta > \theta_m$. Furthermore, from (45), and since $z_m(\theta_m) = 0$,

$$w_{m+1}(\theta_m) = [(\theta_m - m)\rho]/(m+1)^2. \tag{48}$$

Applying (45), for $i = m-1$, we obtain,

$$w_m(\theta_m) = 0 = [(\theta_m - m + 1)\rho + 2(m-1)^2 z_{m-1}(\theta_m)]/m^2, \tag{49}$$

or

$$(\theta_m - m)\rho = -[\rho + 2(m-1)^2 z_{m-1}(\theta_m)] < 0,$$

since $z_{m-1}(\theta_m) > 0$ [$z_{m-1}(\theta_{m-1}) = 0$, $\theta_m > \theta_{m-1}$ and $z_{m-1}(\theta)$ is strictly increasing for $\theta > \theta_{m-1}$]. From (48) and (49), $w_{m+1}(\theta_m) < 0$. From (45) it is clear that $w_{m+1}(\theta) > 0$ for sufficiently large θ. Since $w_{m+1}(\theta)$ is continuous and strictly increasing for $\theta > \theta_m$, there exists a unique $\theta_{m+1} > \theta_m$ such that $w_{m+1}(\theta_{m+1}) = 0$ and $w_{m+1}(\theta) > 0$ for $\theta > \theta_{m+1}$. Consequently, from (46), $z_{m+1}(\theta)$ is well defined for $\theta \geq \theta_{m+1}$ and is of C^∞ for $\theta > \theta_{m+1}$. Furthermore, $w'_{m+1}(\theta) > 0$ and $w''_{m+1}(\theta) < 0$ for $\theta > \theta_{m+1}$, since these statements have already been shown to hold for $\theta > \theta_m$ and since $\theta_{m+1} > \theta_m$. From Lemma 1 it follows that $z_{m+1}(\theta) > 0$, $z'_{m+1}(\theta) > 0$, and $z''_{m+1}(\theta) < 0$ for $\theta > \theta_{m+1}$, thus completing the induction proof.

LEMMA 2: The $\{\theta_i\}$ for $i \in K$ defined in Theorem 6 satisfy the condition that $\theta_i < i$.

PROOF: The values $\theta_0 < 0$ and $\theta_1 = 0$ obviously satisfy the condition for $i = 0$ and $i = 1$, respectively. For $i = 2$,

$$w_2(\theta) = [(\theta - 1)\rho + 2z_1]/4,$$

so that

$$w_2(2) = [\rho + 2z_1(2)]/4.$$

Since $2 > \theta_1 = 0$, from Theorem 6 it follows that $z_1(2) > 0$, and hence $w_2(2) > 0$. Since $w_2(\theta_2) = 0$ and $w_2(\theta)$ is strictly increasing for $\theta > \theta_2$, it follows that $\theta_2 < 2$. Let us now assume that the assertion $\theta_i < i$ has been established for $i = 0, 1, 2, \ldots, m < k$ ($m \geq 2$). To complete the induction proof, we show that $\theta_{m+1} < (m+1)$. From (45),

$$w_{m+1}(m+1) = [\rho + 2m^2 z_m(m+1)]/(m+1)^2.$$

Now, from Theorem 6, $z_m(\theta_m) = 0$ and $z_m(\theta) > 0$ for $\theta > \theta_m$. Consequently, $z_m(m+1) > 0$, since $(m+1) > m > \theta_m$ by the induction hypothesis. Thus, $w_{m+1}(m+1) > 0$. Since

$w_{m+1}(\theta_{m+1}) = 0$ and $w_{m+1}(\theta)$ is strictly increasing for $\theta > \theta_{m+1}$, it follows that $\theta_{m+1} < (m+1)$, thus completing the proof.

LEMMA 3: In the forward-recursion procedure

$$z_i(\rho/4) = \rho/2i, \text{ for } i \in K''. \tag{50}$$

PROOF: For $i = 1$, from (44) and (45), $w_1(\rho/4) = (\rho^2/4)$, so that, from (46),

$$z_1(\rho/4) = + (\rho^2/4)^{1/2} = \rho/2,$$

thus verifying the assertion for $i = 1$. For $i = 2$, (45) with $\theta = \rho/4$ yields

$$h(\rho/4) = (k - \rho/4)\,(\rho/2k^2) - \rho/2k = -\rho^2/8k^2 < 0.$$

so that, from (46), $z_2(\rho/4) = \rho/4$, verifying (50) for $i = 2$.

To prove the general induction step, we assume that (50) has been shown to hold for $i = 1, 2, \ldots, m < k$, where $m \geqslant 2$. We now prove that (50) holds for $i = m+1$ also. For $i = m+1$, (45), with $\theta = \rho/4$, yields

$$w_{m+1}(\rho/4) = \{[(\rho/4) - m]\,\rho + 2m^2(\rho/2m)\}/(m+1)^2$$

$$= \rho^2/[4(m+1)^2].$$

Consequently, from (46), $z_{m+1}(\rho/4) = \rho/2(m+1)$, as was to be proved.

We now prove the (unique) existence of θ^* satisfying (38) through (41) and show that the solution corresponding to θ^* solves P_1.

THEOREM 7: (A) There exists a unique $\theta^* < \min\{\rho/4, k\}$ which satisfies (38) through (41). (B) The corresponding $\{z_i\}$ satisfying (38) through (41) also satisfy (34). (Recall that $z_i = \pi_i/\pi_{i-1}$.) (C) The associated π_i^* for $i \in K$ (see (42) and (43)) is the unique optimal solution to Problem P_3, (31) through (34). (D) The corresponding solution $\{u_i\}$ for $i \in K'$ obtained from (30) is optimal to P_1, (15) through (19).

PROOF: We first note that once (A) and (B) are proved, (C) and (D) follow immediately from our earlier development.

In order to prove (A), we first show the (unique) existence of $\theta^* < k$. Let us define

$$h(\theta) = \hat{z}_k(\theta) - z_k(\theta), \tag{51}$$

where $\hat{z}_k(\theta)$ is given by (47) and $z_k(\theta)$ is obtained by the forward recursion. Recall from Theorem 6 that $z_k(\theta)$ is well defined for $\theta \geqslant \theta_k$ and $z_k(\theta)$ is of C^∞, strictly positive, and strictly increasing for $\theta > \theta_k$. Since $\hat{z}_k(\theta)$ is well defined and strictly decreasing for all θ, $h(\theta)$ is well defined and strictly decreasing for $\theta > \theta_k$. Note also that $h(\theta)$ is continuous for $\theta \geqslant \theta_k$, since both of its component functions are continuous. Conditions (38) through (41) will be satisfied at θ^* iff $\theta^* > \theta_k$ and $h(\theta^*) = 0$.

From Lemma 2, $\theta_k < k$. Consequently, $h(\theta_k) = \hat{z}_k(\theta_k) > 0$, since $\hat{z}_k(\theta)$ is strictly decreasing and $\hat{z}_k(k) = 0$. Similarly $h(k) = -z_k(k) < 0$, since z_k is strictly increasing and $z_k(\theta_k) = 0$. By continuity and the strictly decreasing nature of $h(\theta)$, it follows that there exists a unique θ^* such that $h(\theta^*) = 0$, where $\theta_k < \theta^* < k$. Such a θ^* then satisfies (38) through (41).

To prove that $\theta^* < \min\{\rho/4, k\}$, it is enough to show that $\theta^* < \rho/4$, since we have just shown that $\theta^* < k$. By Lemma 3, $z_k(\rho/4) = \rho/2k$, so that

$$h(\rho/4) = (k - \rho/4)(\rho/2k^2) - \rho/2k = -\rho^2/8k^2 < 0.$$

Since h is strictly decreasing and $h(\theta^*) = 0$, it must be that $\theta^* < \rho/4$, thus completing the proof of (A).

To prove (B), we first note from Lemma 3 that $z_i(\rho/4) = \rho/2i$. From part (A) of the theorem, $\theta^* < \rho/4$. Consequently, by the strictly increasing nature of $z_i(\theta)$ (see Theorem 6) it follows that

$$\pi_i^*/\pi_{i-1}^* = z_i^* = z_i(\theta^*) < z_i(\rho/4) = \rho/2i < \rho/i, \tag{52}$$

thus satisfying (34). Q.E.D.

Theorems 6 and 7 establish that the solution procedure pictorially represented in Figure 1 does yield an optimum solution to Problem P_3, and hence to P_1. This solution procedure is detailed as Algorithm 1 below. Before presenting the algorithm, however, we prove the intuitively pleasing result that $u_0^* < u_1^* < u_2^* < \ldots < u_{k-1}^*$; i.e., when more units are out on rent, we would like to charge a higher price. The theorem also shows that $u_i^* < 1$ (cf. Theorem 2) and that $u_i^* > 1/2$. The result that $u_i^* > 1/2$ makes intuitive sense, since we will later show (cf. Theorem 9) that, as $k \to \infty$ (unconstrained resource), the optimal solution is $u_i^* = 1/2$, so that when k is finite (i.e., resource is constrained) the opportunity cost of the resource (cf. [16]) would make $u_i^* > 1/2$.

THEOREM 8: The optimal solution $\{u_i^*\}$, for $i \in K'$, to P_1 satisfies the property

$$1/2 < u_0^* < u_1^* < \ldots < u_{k-1}^* < 1. \tag{53}$$

PROOF: We prove this theorem in two parts. We first show that $1/2 < u_i^* < 1$ for $i \in K'$. We then prove that $u_i^* \leqslant u_{i+1}^*$ for $i = 0, 1, 2, \ldots, k-2$.

From (30), recalling that $z_{i+1}^* = \pi_{i+1}^*/\pi_i^*$,

$$u_i^* = 1 - (i+1)z_{i+1}^*/\rho = 1 - (i+1)z_{i+1}(\theta^*)/\rho. \tag{54}$$

From Theorem 7(A), $\theta_k < \theta^* < \min\{\rho/4, k\}$. Furthermore, from Theorem 6, $z_{i+1}(\theta)$ is strictly increasing in θ for $\theta > \theta_{i+1}$. Consequently,

$$u_i^* < 1 - (i+1)z_{i+1}(\theta_k)/\rho \leqslant 1 - (i+1)z_{i+1}(\theta_{i+1})/\rho = 1, \tag{55}$$

since $\theta_{i+1} \leqslant \theta_k$, from Theorem 6. (The equality holds when $i = k - 1$.) Futhermore, from the fact that $\theta^* < \rho/4$ and Lemma 3 it follows that

$$u_i^* > 1 - (i + 1)\, z_{i+1}(\rho/4)/\rho = 1 - 1/2 = 1/2. \tag{56}$$

From (55) and (56) it follows that $1/2 < u_i^* < 1$ for $i \in K'$.

To show that $u_{i-1}^* < u_i^*$ for $i = 1, 2 \ldots, k - 1$, we first note that, from (54), the condition holds if and only if

$$iz_i(\theta^*) > (i + 1)\, z_{i+1}(\theta^*), \text{ for } i = 1, 2, \ldots, k - 1. \tag{57}$$

We prove (57) by induction. For $i = 1$, we get from (38) and (39)

$$\rho - 2z_1^* + 4z_2^{*2} = \theta\rho = z_1^{*2},$$

or

$$z_1^{*2} - 4z_2^{*2} = \rho - 2z_1^* = 2[(\rho/2) - z_1^*] > 0,$$

since, from (52), $z_1(\theta^*) < \rho/2$. Consequently, $z_1^{*2} > 4z_2^{*2}$, or $z_1^* > 2z_2^*$, since $z_i^* > 0$, for $i \in K''$, at the optimum (see (41)). Thus, (57) holds for $i = 1$.

For $i = 2$, we get from (39)

$$\rho - 2z_1^* + 4z_2^{*2} = \theta\rho = 2\rho - 8z_2^* + 9z_3^{*2},$$

or

$$4z_2^{*2} - 9z_3^{*2} = (\rho - 4z_2^*) + (2z_1^* - 4z_2^*) > 0,$$

since, from (52), $z_2(\theta^*) < \rho/4$ and $z_1^* > 2z_2^*$, as shown earlier. Consequently, $4z_2^{*2} > 9z_3^{*2}$, or $2z_2^* > 3z_3^*$, since all the z_i^*'s are strictly positive at the optimum. Thus, (57) holds for $i = 2$.

For the general induction step, we assume that (57) holds for $i = 1, 2, \ldots, m - 1 < (k - 2)$, where $(m - 1) \geqslant 2$. We prove below that $m\, z_m^* > (m + 1)\, z_{m+1}^*$. From (39)

$$(m - 1)\, \rho - 2(m - 1)^2\, z_{m-1}^* + m^2\, z_m^{*2} = \theta\rho = m\rho - 2m^2 z_m^* + (m + 1)^2\, z_{m+1}^{*2},$$

or

$$m^2\, z_m^{*2} - (m + 1)^2\, z_{m+1}^{*2} = (\rho - 2mz_m^*) + 2(m - 1)\, [(m - 1)\, z_{m-1}^* - mz_m^*] > 0,$$

since $z_m^* < \rho/2m$, from (52), and $(m-1)\, z_{m-1}^* > m z_m^*$, by the induction assumption. Thus $m^2 z_m^{*2} > (m+1)^2\, z_{m+1}^{*2}$, or $m z_m^* > (m+1)\, z_{m+1}^*$, since the z_i^*'s are strictly positive at the optimum (see (41)), thus completing the proof.

4.3. The Algorithm

We now provide Algorithm 1, based on the forward-recursion procedure discussed earlier, using the bisection search to identify θ^*. From Theorem 7, $\theta^* < \theta_{\max}$, where

$$\theta_{\max} = \text{Min}\,\{\rho/4,\ k\}. \tag{58}$$

Furthermore, by Theorem 5 and from the fact that $\bar{\mathbf{u}} = \{\bar{u}_i = 1/2\}$ defines a feasible solution, we can set

$$\theta^* \geqslant \theta_{\min} = F(\bar{\mathbf{u}}) > 0. \tag{59}$$

ALGORITHM 1: To find the optimal $\mathbf{u}^*$ to P_1:

STEP 0: Define $\theta_{\max}$ and $\theta_{\min}$ as per (58) and (59). Set $\theta = (\theta_{\min} + \theta_{\max})/2$.

STEP 1: Set $z_0 = 0$. Using (45) and (46) recursively, compute $w_1, w_2, \ldots, w_k$ and $z_1, z_2, \ldots, z_k$. Each time a w_i is computed, check to see whether $w_i < 0$. If $w_i < 0$, replace $\theta_{\min}$ by the current value of θ, and go to Step 2. If $w_i \geqslant 0$ for all i, go to Step 3.

STEP 2: Set $\theta = (\theta_{\min} + \theta_{\max})/2$. Go to Step 1.

STEP 3: Calculate $\hat{z}_k(\theta)$ from (47). Compare it with $z_k(\theta)$ computed from Step 1. If either of the stopping rules below is satisfied, go to Step 4. If not, go to Step 5.

Stopping Rules

(A) $|z_k - \hat{z}_k| < E_1$, where E_1 is a prespecified small number.

(B) $(\theta_{\max} - \theta_{\min}) < E_2$, where E_2 denotes the maximum precision for computational purposes.

STEP 4: If stopped by Rule (A), compute $\{\pi_i^*\}$ from the $\{z_i\}$, using (42) and (43) and u_i^* from (30). The "optimal" value $F(\mathbf{u}^*) = \theta^*$, where θ^* is the current value of θ. STOP. If stopped by Rule (B), higher precision is required for the computations. Multiprecision routines can be used for this purpose. STOP.

STEP 5: If $z_k > \hat{z}_k$, replace $\theta_{\max}$ by the current θ, and go to Step 2. If $z_k < \hat{z}_k$, replace $\theta_{\min}$ by the current θ, and go to Step 2.

It is easily seen that at each iteration the search interval is reduced to half of its original size, and hence the algorithm will terminate in a finite number of steps for any strictly positive precision parameter E_2.

4.4 Computational Experience

A total of 48 test problems were solved by Algorithm 1 for each combination of the parameter values given below:

(i) $k = 2, 5, 10, 15, 20, 40$;

(ii) $\rho/k = 0.2, 0.5, 1, 2, 3, 4, 5, 6$.

The values of ρ/k above were chosen with a view to keeping the problem within realistic limits. Any real system, it was felt, would not have too high an "all idle" probability π_0 or too high an "all rented out" probability π_k. The values of ρ/k given above approximately meet an upper limit of 0.2 on the probabilities π_0 and π_k. The parameter E_1 was set equal to 10^{-5}, and E_2 was set to correspond to double-precision arithmetic on the IBM 360/65 computer ($E_2 \simeq 10^{-15}$). The 48 problems were all solved in a total of 24.7 seconds ($\simeq 0.5$ s per problem; the program was compiled on FORTRAN G). However, two of the 48 problems had to be terminated because of stopping rule (B), i.e., the double precision was not sufficient.

5. STATE-INDEPENDENT PRICING POLICIES.

Let u denote the rental rate that is independent of the state of the system. Problem P_4 for determining the optimal u is a special case of Problem P_2, (22) and (23). Replacing $\{u_i\}$, for $i \in K'$, by u, we obtain Problem P_4:

$$\text{(60)} \qquad \underset{u}{\text{Maximize}}\ Q(u) = \frac{\sum_{i \in K'} (\rho^{i+1}/i!)\ u(1-u)^{i+1}}{\sum_{i \in K} (\rho^i/i!)\ (1-u)^i}$$

$$\text{(61)} \qquad \text{subject to} \qquad 0 \leqslant u \leqslant 1.$$

The objective $Q(u)$ in (60) can be rewritten as

$$\text{(62)} \qquad Q(u) = \rho u\ (1-u)\ (E_{k-1}/E_k),$$

where

$$\text{(63)} \qquad E_j = \sum_{i=0}^{j} (\rho^i/i!)\ (1-u)^i.$$

The following theorem gives the main properties of the optimal solution u^* and its relationship to the optimal state-dependent pricing solution $\{u_i^*\}$.

THEOREM 9: (A) $0 < u^* < 1$.

(B) If $k \to \infty$, then $u^* = 1/2$ uniquely solves P_4.

(C) If $k \to \infty$, then $u_i^* = 1/2$, for $i \in K'$, is also an optimal state-dependent pricing policy, i.e., it solves P_3.

PROOF: (A) When $u \to 0$ or 1, numerator of $Q(u)$ in (60) also tends to zero, while the denominator tends to some strictly positive number. Consequently, $Q(u) \to 0$ and is

poorer than the feasible solution $u = 1/2$, which yields a strictly positive $Q(u)$. Consequently, $0 < u^* < 1$.

(B) For the optimal solution, observe that from (62) and (63) we have

$$\lim_{k\to\infty} E_{k-1}/E_k \to 1, \text{ so that}$$

$$\lim_{k\to\infty} Q(u) = \rho u\,(1-u).$$

By setting the derivative of $\rho u\,(1-u)$ to zero, we find that $u^* = 1/2$ uniquely maximizes $Q(u)$ and that $Q(u^*) = \rho/4$.

(C) From Theorem 5 and Theorem 7 (A), for any k, the optimal objective function value for P_3 is less than or equal to $\rho/4$. With $u_i^* = u^* = 1/2$ for all $i \in K'$, the objective function for Problem P_4 attains the value $\rho/4$ as $k \to \infty$. Since P_4 is a more constrained problem than P_3, it follows that the solution $u_i^* = 1/2$, for $i \in K'$, is also optimal to P_3.

Thus, from Theorem 9 we find that if $k \to \infty$ an optimal state-dependent pricing policy is, in fact, state-independent. To get a feel for the closeness of the two policies for any finite k, the same 48 problems solved in Section 4.4 were resolved by the problem formulation P_4. In all the 48 problems the function $Q(u)$ was found to be unimodal. Using the (conjectured) unimodality property, the optimal u^* for the 48 problems were determined by Fibonacci search [20, pp. 24-30] in a total of 18.7 s (approximately 0.4 s/problem). It was found that, although the state-independent pricing policy was obviously poorer than the state-dependent policy in terms of the objective, the difference in the values of the objective function between the two policies was very small, ranging from 0 to 1.5%, with an average of 0.7% over the entire set of problems tested. For the case when $\rho/k = 1$ (a likely situation with $\lambda = k\eta$), the percent difference drops from 0.312%, for $k = 2$, to 0.00067% for $k = 40$. As explained in Section 1, this may account, at least partly, for the absence of state-dependent pricing in many large rental systems. For a given k, the percent difference is the largest for $\rho/k = 3$. It drops sharply as ρ/k decreases below 3, but drops only gradually as ρ/k increases above 3.

6. SOME GENERALIZATIONS OF THE PROBLEM FORMULATION P1

The rental policy considered in Section 1 was a rental rate of the form v_i dollars/h. Let us assume that the policy, in addition, involves a fixed charge of the form A_i dollars/rental, so that the expected total rental cost per rental is $A_i + (v_i/\eta)$. Now, if we assume that a customer's probability of renting decreases linearly with the expected total rental costs, i.e.,

$$p_i(A_i, v_i) = 1 - b\,[A_i + (v_i/\eta)], \tag{64}$$

and follow a line of reasoning similar to that in Section 1, Problem P_5 becomes:

$$\underset{\mathbf{A},\mathbf{v}}{\text{Maximize}}\ D(\mathbf{A}, \mathbf{v}) = \sum_{i\in K'} \lambda[A_i + (v_i/\eta)]\,\{1 - b[A_i + (v_i/\eta)]\}\,\pi_i$$

subject to

$$\pi_{i+1} = \lambda\{1 - b[A_i + (v_i/\eta)]\}\pi_i/[(i+1)\eta], \text{ for } i \in K',$$

$$\sum_{i \in K} \pi_i = 1,$$

$$\pi_i \geqslant 0, \text{ for } i \in K,$$

and

$$0 \leqslant A_i + (v_i/\eta) \leqslant 1/b, \text{ for } i \in K'.$$

Fortunately, Problem P_5 can be reduced to the mathematical form P_1 (equations (15) through (19)) if we use the transformation

$$u_i = b\,[A_i + (v_i/\eta)], \text{ for } i \in K'$$

and maximize $(b/\eta)D$ instead of D. (Since b/η is a strictly positive constant, the optimal u_i^* will not be affected by this change in the objective function.) The equivalence of the two problem formulations shows that the optimal A_i^* and v_i^* can be set arbitrarily, subject to $b[A_i^* + (v_i^*/\eta)] = u_i^*$.

Another worthwhile extension is to assume that the service time depends on the rate v_i at which the unit was rented, e.g., the parameter η may be modelled as

$$\eta_i = c + dv_i \ (c > 0,\ d > 0),$$

so that if the rental rate is high, the unit's expected rental time, $1/\eta_i$, will be smaller. The objective $D(\mathbf{v})$ (equation (9)) can be easily modified to take this change into account. However, the dependence of service times on the rental rate destroys the underlying Markovian character of the original model P_0 in terms of the transition equations (10). Now, when the process is in state i, some previous history of the system is also needed. For instance, when $k = 2$ and $i = 2$, it must be known whether both units were rented at u_1 or one at u_0 and the other at u_1. Correspondingly, the step down transition rate will be $2\eta_1$ or $(\eta_0 + \eta_1)$. One method of overcoming this difficulty is to augment the state-space and reindex the states, so that the process will assume the underlying Markovian character [5, p.18]. This, however, increases the state space from $(k+1)$ to $(k+1)(2k+1)/6$ and destroys the structure of the constraint matrix (10) through (13), which permitted the efficient solution procedure of Section 4.

Clearly, other extensions, such as a general nonlinear form for the probability of renting $p(v_i)$ and general arrival and service time distributions, would be useful. However, these extensions can be made only at the expense of greatly increased computational effort.

ACKNOWLEDGMENTS

The authors wish to thank Professor John B. Long, Jr., of the University of Rochester, for his valuable suggestions during early stages of the research reported here. Professor Julian Keilson of the University of Rochester suggested the present research as an extension of his earlier work, helped in the formulation of the problem, and provided a more rigorous validation of the objective function in Section 1. We are extremely grateful for his most valuable contributions to this research.

APPENDIX

In this appendix, we use a discrete optimal control approach [3,4] to obtain the necessary conditions for the optimal solution to P_1. The resulting conditions will be shown to be equivalent to (38) through (41), (30), and (34) of the nonlinear programming approach discussed in Section 4.

If we define $s_i = \sum_{j=i}^{k} \pi_j$ (cf. the dynamic-programming procedure in Section 3), Problem P_1 ((15) through (19)) can be rewritten as Problem P_{A1},

$$\text{(A1)} \qquad \underset{\mathbf{u}}{\text{Maximize}} \; F(\mathbf{u}) = \sum_{i \in K'} \rho u_i (1 - u_i) \pi_i$$

subject to

$$\text{(A2)} \qquad \pi_{i+1} - \pi_i = \pi_i \{[\rho/(i+1)] (1 - u_i) - 1\}, \text{ for } i \in K',$$

$$\text{(A3)} \qquad s_{i+1} - s_i = -\pi_i, \text{ for } i \in K',$$

$$\text{(A4)} \qquad s_0 = 1,$$

$$\text{(A5)} \qquad 0 \leqslant s_k = \pi_k,$$

$$\text{(A6)} \qquad 0 \leqslant u_i \leqslant 1, \text{ for } i \in K',$$

and

$$\text{(A7)} \qquad 0 \leqslant \pi_i \leqslant s_i \leqslant 1, \text{ for } i \in K.$$

Problem P_{A1} is in the form of a discrete optimal control problem, with $\mathbf{u}$ as the control variable and $(\vec{\pi}, \mathbf{s})$ as the state variables. We first show that the constraints (A7) are redundant.

THEOREM A1: For Problem P_{A1} (A1) through (A7), constraints (A7) are redundant.

PROOF: From (A5), $\pi_k \geqslant 0$.

From (A2), $\pi_i = (i+1)\pi_{i+1}/\rho(1 - u_i)$.

From (A6), $(1 - u_i) \geqslant 0$, so that by backward recursion, $\pi_i \geqslant 0$ for $i \in K$.

Equation (A3) may be rewritten as

$$s_i = s_{i+1} + \pi_i.$$

Since $s_k \geqslant 0$ and $\pi_{k-1} \geqslant 0$, it follows that $s_{k-1} \geqslant \pi_{k-1} \geqslant 0$ and $s_{k-1} \geqslant s_k$. By backward recursion, $s_i \geqslant \pi_i \geqslant 0$ and $s_i \geqslant s_{i+1}$, for $i \in K'$. Consequently, from (A4), $1 = s_0 \geqslant s_i$, for $i \in K$, thus completing the proof.

For a given state vector $(\vec{\pi}, s)$, the constraints (A2), (A3), and (A6) are linear in u_i and the ith term in (A1), $\rho u_i(1 - u_i)\pi_i$, is a negative definite quadratic form in u_i. Thus, the directional convexity and other requirements [4, pp. 86-87] are satisfied, and hence the necessary conditions for discrete optimal control [4, pp. 91-92] apply.

Let $\{y_i\}$ and $\{e_i\}$, for $i \in K''$, be the adjoint variables associated with (A2) and (A3) respectively. Then the Hamiltonian of the problem is given by

(A8)
$$H(\pi_i, s_i, u_i, y_{i+1}, e_{i+1}, i) =$$
$$\rho u_i(1 - u_i)\pi_i + \{[\rho/(i + 1)]\ (1 - u_i) - 1\}\ \pi_i y_{i+1} - \pi_i e_{i+1}.$$

The adjoint equations are given by

(A9) $$y_i - y_{i+1} = \rho u_i(1 - u_i) + \{[\rho/(i + 1)]\ (1 - u_i) - 1\}y_{i+1} - e_{i+1}, \text{ for } i \in K',$$

and

(A10) $$e_i - e_{i+1} = 0, \text{ for } i \in K'.$$

The transversality conditions corresponding to (A4) and (A5) are

(A11) $$y_0 = 0$$

and

(A12) $$y_k = -\ e_k.$$

From (A10) and (A12) we get

(A13) $$e_i = -\ y_k = E \text{ (say)}, \text{ for } i \in K.$$

By the discrete-maximum principle, if $\{u_i^*\}$ is optimal then, for each $i \in K'$ and for all u_i satisfying (A6),

(A14) $$H(\pi_i^*, s_i^*, u_i^*, y_{i+1}^*, e_{i+1}^*, i) \geqslant H(\pi_i^*, s_i^*, u_i, y_{i+1}^*, e_{i+1}^*, i),$$

where π_i^*, s_i^*, y_{i+1}^*, and e_{i+1}^* are derived from (A2) through (A5) and (A9) through (A12) with u_i^* substituted for u_i.

By (A14) and from (53) (i.e., $1/2 < u_i^* < 1$), u_i^* can be obtained by differentiation of $H(\pi_i^*, s_i^*, u_i, y_{i+1}^*, e_{i+1}^*)$ (equation (A8)). This yields

(A15) $$u_i^* = \{1 - [y_{i+1}/(i+1)]\}/2.$$

Substituting (A15) and (A13) for u_i and e_{i+1} in (A9) and simplifying the resulting expression, we obtain

(A16) $$y_i^* = (\rho/4)\,[y_{i+1}^*/(i+1)]^2 + (\rho/2)\,[y_{i+1}^*/(i+1)] + (\rho/4) - E.$$

Defining

(A17) $$x_i = y_i^*/i, \text{ or } y_i^* = ix_i,$$

we can simplify (A16) to

(A18) $$ix_i = (\rho/4)\,(1+x_{i+1})^2 - E, \text{ for } i \in K',$$

or

(A19) $$x_i = (\rho/4i)\,(1 + x_{i+1})^2 - (E/i), \text{ for } i = 1, 2, \ldots, k-1.$$

(The condition for $i = 0$ is considered later in (A21)).

The transversality condition (A12) together with (A13) and (A17), yields

(A20) $$x_k = -E/k.$$

The transversality condition (A11), together with (A17) and (A18), yields

(A21) $$(\rho/4)\,(1 + x_1)^2 - E = 0.$$

Furthermore, (A15) and (A17) yield

(A22) $$u_i^* = (1 - x_{i+1})/2, \text{ for } i \in K',$$

so that the conditions (A6) can be rewritten as

(A23) $$x_i \leqslant 1, \text{ for } i \in K'',$$

and

(A24) $$x_i \geqslant -1, \text{ for } i \in K''.$$

We can summarize the above results as follows:

THEOREM A2: Suppose that $\mathbf{u}^*$ is optimal to P_1. Then there exists an E such that (A19) through (A24) are satisfied.

We now show that the necessary conditions for P_1, as stated by Theorem A2, are precisely the same as the conditions derived for the nonlinear-programming approach to the problem detailed in Section 4.

THEOREM A3: The system (A19) through (A24) is equivalent to the conditions (38) through (41), (30), and (34) under the transformations

(A25) $$x_i = (2iz_i/\rho) - 1$$

and

(A26) $$E = \theta.$$

PROOF: By direct substitution of (A25) and (A26), it can be verified that the conditions (A19) through (A24) are the same as (39), (40), (38), (30), (34), and (41) respectively.

Thus, the discrete optimal-control approach leads to the same conditions, and hence the same solution procedure, as Algorithm 1 described in Section 4.

REFERENCES

[1] Blackwell, D., "Discrete Dynamic Programming," Annals of Mathematical Statistics *33*, 719-726 (1962).

[2] Blackwell, D., "Discounted Dynamic Programming," Annals of Mathematical Statistics *36*, 226-235 (1965).

[3] Blaquiére, A., F. Gérard, and G. Leitmann, *Quantitative and Qualitative Games* (Academic Press, New York, 1969).

[4] Canon, M.D., C.D. Cullum, Jr., and E. Polak, *Theory of Optimal Control and Mathematical Programming* (McGraw-Hill, New York, 1970).

[5] Cox, D.R., and H.D. Miller, *The Theory of Stochastic Processes* (Wiley, New York, 1968).

[6] Denardo, E.V., "Contraction Mappings in the Theory Underlying Dynamic Programming," SIAM Review *9*, 165-177 (1967).

[7] Denardo, E., and B. Fox, "Multi-Chain Markov Renewal Programs," SIAM Journal of Applied Mathematics *16*, 468-487 (1968).

[8] Derman, C., *Finite State Markovian Processes* (Academic Press, New York, 1970).

[9] Fox, B.L., "The Existence of Stationary Optimal Policies for Some Markov Renewal Programs," SIAM Review *9*, 573-576 (1967).

[10] Gupta, R.K., "A Class of Stochastic Bidding Problems," unpublished doctoral dissertation, The University of Rochester, Rochester, New York (1974).

[11] Hadley, G., *Nonlinear and Dynamic Programming* (Addison-Wesley, Reading, Massachusetts, 1964).

[12] Howard, R.A., *Dynamic Programming and Markov Processes* (Wiley, New York, 1960).

[13] Howard, R.A., *Dynamic Probabilistic Systems,* Vols. I and II (Wiley, New York, 1971).

[14] Jewell, W.S., "Markov-Renewal Programming, I and II," Operations Research *11*, 938-971 (1963).

[15] Keilson, J., "A Simple Algorithm for Contract Acceptance," Opsearch *7*, 157-166 (1970).

[16] Kortanek, K.O., J.V. Soden, and D. Sodaro, "Profit Analyses and Sequential Bid Pricing Models," Management Science *20*, 396-417 (1973).

[17] Low, D.W., "Optimal Dynamic Pricing Policies for an M/M/S Queue," Operations Research *22*, 545-561 (1974).

[18] Mangasarian, O.L., *Nonlinear Programming* (McGraw-Hill, New York, 1969).

[19] Ross, S., and S. Lippman, "The Streetwalker's Dilemma: A Job Shop Model," Western Management Science Institute Technical Report, University of California, Los Angeles, California (1969).

[20] Wilde, D.J., *Optimum Seeking Methods* (Prentice-Hall, Englewood Cliffs, New Jersey, 1964).

[21] Zangwill, W.I. *Nonlinear Programming — A Unified Approach* (Prentice-Hall, Englewood Cliffs, New Jersey, 1969).

ON THE FIRST TIME A SEPARATELY MAINTAINED PARALLEL SYSTEM HAS BEEN DOWN FOR A FIXED TIME*

Sheldon M. Ross

Department of Industrial Engineering
and Operations Research
University of California, Berkeley
Berkeley, California

Jack Schechtman

Institute of Pure and Applied Mathematics
Rio de Janeiro, Brazil

ABSTRACT

Consider a system consisting of n separately maintained independent components where the components alternate between intervals in which they are "up" and in which they are "down". When the i^{th} component goes up [down] then, independent of the past, it remains up [down] for a random length of time, having distribution F_i[G_i], and then goes down [up]. We say that component i is failed at time t if it has been "down" at all time points $s \in [t-A, t]$; otherwise it is said to be working. Thus, a component is failed if it is down and has been down for the previous A time units. Assuming that all components initially start "up," let T denote the first time they are all failed, at which point we say the system is failed. We obtain the moment-generating function of T when $n = 1$, for general F and G, thus generalizing previous results which assumed that at least one of these distributions be exponential. In addition, we present a condition under which T is an NBU (new better than used) random variable. Finally we assume that all the up and down distributions F_i and G_i, $i=1, \ldots, n$, are exponential, and we obtain an exact expression for $E(T)$ for general n; in addition we obtain bounds for all higher moments of T by showing that T is NBU.

INTRODUCTION AND SUMMARY

In considerating a system that works for a random time and when failed is fixed in a length of time that is also random, an important variable is the first time the system is not working for an interval of time longer than some prespecified value. For instance, in a nuclear reactor, when the safety system is out for some critical time, it is necessary to shut down the complete system, with all the problems this entails. In the food industry, where food must in general be kept at a certain temperature, an important question that arises when the refrigeration system goes down is how long this situation can be maintained before the food becomes spoiled.

*This research has been partially supported by the Air Force Office of Scientific Research (AFSC), USAF, under Grant AFOSR-77-3213 and the Office of Naval Research under Contract N00014-77-C-0299 with the University of California.

In this paper, we consider a system consisting of n separately maintained independent components, where the components alternate between intervals in which they are "up" and in which they are "down." When the i^{th} component goes up [down] then, independent of the past, it remains up [down] for a random length of time, having distribution $F_i[G_i]$, and then goes down [up]. We say that component i is *failed at time t* if it has been "down" at all time points $s \in [t-A,t]$; otherwise, it is said to be working at time t. Thus, a component is failed if it is down and has been down for at least previous A time units. Assuming that all components initially start "up," let T denote the first time they are all failed, at which point we say the system is failed.

In Section 1, we obtain the moment-generating function of T when $n = 1$, for general F and G, thus generalizing results in [5] and [6] which assumed that at least one of these distributions be exponential. In Section 2, we present a condition under which T is an NBU (new better than used) random variable. In Section 3, we assume that all the up and down distributions F_i and $G_i, i=1, \ldots, n$, are exponential, and we obtain an exact expression for $E(T)$ for general n; in addition, we obtain bounds for all higher moments of T by showing that T is NBU.

1. THE CASE $n = 1$

Let us denote by N the number of "up" intervals that occur before the component fails. Then, given $N = k$, we can represent T by

$$T = X_1 + \ldots + X_k + Y_1^A + \ldots + Y_{k-1}^A + A, \tag{1}$$

where X_i denotes the length of the i^{th} up cycle and Y_i^A the length of the i^{th} down cycle before failure. All the random variables in the representation (1) are independent, with the X_i having distribution F and the Y_i^A having distribution

$$P\{Y_i^A \leqslant x\} = P\{Y \leqslant x \mid Y < A\} = \begin{cases} \dfrac{G(x)}{G(A)}, & 0 \leqslant x < A, \\ 1, & x \geqslant A, \end{cases}$$

where F is the distribution of an up cycle and G that of a down cycle, and Y is a generic variable, having distribution G, and representing the unconditional length of a down cycle. As

$$P\{N = k\} = \bar{G}(A)[G(A)]^{k-1}, k = 1, \ldots,$$

where $\bar{G} = 1 - G$, we obtain the moment-generating function of T by conditioning on N as follows:

$$\begin{aligned} E[e^{sT}] &= E[E[e^{sT} \mid N]] \\ &= E\left[e^{sA}(\phi_X(s))^N \left(\phi_{Y^A}(s)\right)^{N-1} \right] \\ &= e^{sA}\phi_X(s)\bar{G}(A) \sum_{k=1}^{\infty} \left[\phi_X(s)\phi_{Y^A}(s)G(A)\right]^{k-1} \qquad (2) \\ &= \frac{e^{sA}\phi_X(s)\bar{G}(A)}{1 - G(A)\phi_X(s)\phi_{Y^A}(s)}, \end{aligned}$$

where

$$\phi_X(s) = E[e^{sX}] = \int_0^\infty e^{sx} dF(x)$$

and

$$\phi_{Y_A}(s) = E\left[e^{sY_A}\right] = \int_0^A \frac{e^{sx} dG(x)}{G(A)}.$$

For the special case in which X is exponential with mean $1/\lambda$ and Y is exponential with mean $1/\mu$, we have

$$E[e^{sT}] = \frac{\lambda(\mu - s)e^{-(\mu - s)A}}{s^2 - (\lambda + \mu)s + \lambda\mu e^{-(\mu - s)A}},$$

a result previously obtained in [5] and [6].

All of the moments can now be obtained by successive differentiation of (2) or by a direct conditioning argument. For instance, we obtain

$$\begin{aligned} E[T] &= E[E[T|N]] \\ &= E[NE[X] + (N-1)E[Y|Y < A] + A] \\ &= \frac{E[X]}{\bar{G}(A)} + \frac{\int_0^A x dG(x)}{\bar{G}(A)} + A. \end{aligned} \tag{3}$$

By viewing the working-failed system as an alternating renewal process, we deduce that the long-run proportion of time the component is failed is

$$\frac{E[Y - A|Y > A]}{E[T] + E[Y - A|Y > A]} = \frac{\int_A^\infty \bar{G}(y)\,dy}{E[Y] + E[X]}.$$

2. WHEN IS T NBU, $n = 1$

The nonnegative random variable W is said to be new better than used (written NBU) if

$$P\{W > s + t | W > s\} \leqslant P\{W > t\}, \forall s, t \geqslant 0.$$

If we think of W as representing the life of some object, then W NBU means that the additional remaining life of any s-year-old (i.e., used) item is stochastically smaller than that of a new item, for all s.

If W is NBU and has distribution function H, then we also say that H is NBU.

PROPOSITION 1: If X, the length of an up time, is NBU, then so is T.

PROOF: Suppose failure has not yet occurred by time s. Now there are two possibilities:

CASE 1: At time s the component is up and has been up for a time t. In this case the remaining time to failure has the distribution of the convolution of F_t and H, where F_t is the distribution of remaining up time for a component that has been up for a time t and H is the distribution of time to failure starting with the component initially down. But since F_t is stochastically smaller than F (the definition of X being NBU), this distribution is stochastically smaller than the convolution of F and H, which is the distribution of T.

CASE 2: At time s the component is down and has been down for a time t (necessarily, $t < A$). In this case the remaining time to failure has some distribution, call it D. However, the distribution of T can be written as the convolution of D and the distribution of the first time that the component has been down for t consecutive time units. This latter convolution distribution is clearly stochastically larger than D.

Thus, in all cases the distribution of T is stochastically larger than the distribution of remaining time until failure. Hence T is NBU.||

3. EXPONENTIAL LIFETIMES, GENERAL n

In this section we suppose there are n components and the distribution of up [down] time for the i^{th} component is exponential, with rate $\lambda_i[\mu_i]$, $i = 1, \ldots, n$. We start by deriving $E[T]$, the expected time until the system fails; that is, until all components are failed, starting with all components up.

We can write T as the sum of independent random variables as follows:

$$T = T_{A=0} + Z, \tag{4}$$

where $T_{A=0}$ denotes the first time that all components are down (it is thus equal to T in the special case $A = 0$) and Z the extra (or additional) time from $T_{A=0}$ until all components are failed. Now Brown [1] has computed $E[T_{A=0}]$ and shown that

$$E[T_{A=0}] = \sum_{k=1}^{n} \sum_{i_1<i_2<\ldots<i_k} \frac{\left[\prod_{j=1}^{k} \frac{\mu_{i_j}}{\lambda_{i_j}} - (-1)^k\right]}{\sum_{j=1}^{k} (\lambda_{i_j} + \mu_{i_j})}.$$

Thus, it remains to compute $E[Z]$. Let M denote an exponential random variable with rate $\mu \equiv \sum_{1}^{n} \mu_i$. Then, by conditioning on whether or not all components remain down in the A time units following time $T_{A=0}$, we obtain

$$E[Z] = Ae^{-\mu A} + (1 - e^{-\mu A})[E[M|M<A] + E[D] + E[Z]],$$

where D is the time until all components are down, given that they were all down and one has just gone up. Thus, from the above, we obtain

$$E[Z] = A + (e^{\mu A} - 1)\left[\frac{\int_0^A \mu x e^{-\mu x} dx}{1 - e^{-\mu A}} + E[D]\right]. \tag{5}$$

However, Ross [3] has shown that

$$E[D] = \frac{1 - \prod_{j=1}^{n} \frac{\lambda_j}{\mu_j + \lambda_j}}{\sum_{i=1}^{n} \mu_i \prod_{j=1}^{n} \frac{\lambda_j}{\mu_j + \lambda_j}}, \tag{6}$$

and thus the expression for $E[T]$ follows from (4), (5) and (6).

The next proposition partly characterizes the distribution of T and will enable us to obtain bounds on all higher moments of T.

PROPOSITION 2: T is NBU.

PROOF: Suppose that all components have never been simultaneously failed by time s. There are two cases:

CASE 1: At time s all components are down, the one that has been down for the shortest time having been down for a time t (where, necessarily, $t < A$). Since T can be expressed as T_{A-t} (the first time all components have been down for the past t time units) plus a random variable having the same distribution as the remaining time to failure of the system, it follows that T is stochastically larger than the remaining time to system failure in this case.

CASE 2: Not all components are down at time s. In this case the remaining time to system failure can be written as the time until all components are down plus an independent random variable having the same distribution as Z in the representation (4). Now Ross [4] has shown that the time until all components are down is stochastically larger when it starts with all being initially up than when it starts in any other position. Hence, from the representation (4), it follows that the remaining time to system failure at time s is stochastically smaller than T.

Hence, in all cases T is stochastically larger than the remaining time to system failure, thus proving the result.||

The above result is particularly useful, as it enables us to obtain bounds on $E[f(T)]$, whenever f is an increasing convex function, by use of the following special case of Theorem 4.6 of Marshall and Proschan [2].

PROPOSITION 3: If X is NBU with mean $1/\lambda$, then

$$E[f(X)] \leqslant \int_0^\infty f(x)\lambda e^{-\lambda x}dx$$

for all increasing convex functions f.

In words, Proposition 3 says that if X is NBU, then $E[f(X)] \leqslant E[f(M)]$ for all increasing convex f, where M is an exponential random variable having the same mean as X.

COROLLARY 1: $Var(T) \leqslant (E[T])^2$.

PROOF: Follows immediately from Propositions 2 and 3 by use of the function $f(x) = x^2$.

REFERENCES

[1] Brown, M., "The First Passage Time Distribution for a Parallel Exponential System with Repair," in *Reliability and Fault Tree Analysis,* (SIAM, Philadelphia, 1975).

[2] Marshall, A.W., and F. Proschan, "Classes of Distributions Applicable in Replacement, with Renewal Theory Implications," in *Proceedings of the Sixth Berkeley Symposium in Mathematical Statistics and Probability,* (University of California Press, 1972) Vol. 1, pp. 395-415.

[3] Ross, S., "On the Calculation of Asymptotic System Reliability Characteristics," in Reliability and Fault Tree Analysis (SIAM, Philadelphia, 1975).

[4] Ross, S., "On Time to First Failure in Multicomponent Exponential Reliability Systems," *Journal of Stochastic Processes and Its Applications* 4, 167-173 (1976).

[5] Von Ellenrieder, A., and A. Levine, "The Probability of an Excessive Non-Functioning Interval," Operations Research 14, No. 5 (1966).

[6] Von Ellenrieder, A., and J. Schechtman, "Sobre El Problema Del Intervalo Critico de No Functionamento," in *Atas do I Simposio Brasileiro de Pesquisa Operacional e suas Aplicacões* (Brazil, 1968), Vol. II, pp. 534-541, (in Portuguese).

A CAPACITY-EXPANSION MODEL FOR TWO FACILITY TYPES

Hanan Luss

Bell Laboratories
Holmdel, New Jersey

ABSTRACT

This paper describes a deterministic capacity-expansion model for two facility types with a finite number of discrete time periods. Capacity expansions are initiated either by new construction or by the conversion of idle capacity from one facility type to the other. Once converted, the capacity becomes an integral part of the new facility type. The costs incurred include construction, conversion, and holding costs. All cost functions are assumed to be nondecreasing and concave. Using a network flow approach, the paper develops an efficient dynamic-programming algorithm to minimize the total costs when the demands for additional capacity are nonnegative in each period. Thereafter, the algorithm is extended for arbitrary demands. The model is applied to a cable-sizing problem that occurs in communication networks, and numerical examples are discussed.

INTRODUCTION

Capacity-expansion models are needed to plan the expansion of facilities over time so as to satisfy given demands at the lowest possible cost. In this paper we describe a deterministic model for two facility types. The model assumes a finite number of discrete time periods, with known demands for each of the two facilities in each period. These demands must be satisfied immediately; i.e., shortages of capacity are not allowed. In each period, facility $i (i = 1,2)$ may be expanded, either by new construction, or by conversion of idle capacity associated with facility $j (j = 1,2$ and $j \neq i)$ to accommodate the demand for facility i. Conversion implies physical modification, so that the converted capacity becomes an integral part of the new facility and is not reconverted automatically at the end of the period. The costs incurred include construction and conversion costs, and holding costs of idle capacity. All cost functions depend on the time period and are assumed to be nondecreasing and concave, perhaps reflecting fixed charges and economies of scale. It is assumed that the operating costs depend only on the quantity of active capacity; hence, they are independent of the expansion policy and are omitted from the model. The capacity-expansion policy consists of timing and sizing decisions for new construction and conversion, and the objective is to find the policy which minimizes total costs.

The study has been stimulated by communication-network applications, especially a cable-sizing problem. Concentrating on a single link of a communication network, we assume that there are demands for two types of cables. Each cable type is characterized by the diameter of the wire pairs in that cable, and the cable size is simply the number of wire pairs included in the cable. The cable type needed to serve the demand for the link associated with any two endpoints of the network depends on the distance between those endpoints. Furthermore, the

more expensive cable type (the one which consists of the larger wire pairs) can serve both demands, whereas the cheaper cable can serve only its associated demand. Since the construction cost is a concave function of the cable size, it may be attractive to use only the expensive cable to satisfy the demands for both cables. Another similar application of the model for communication networks is the planning of capacity expansion associated with facilities which serve both digital and analog demands.

The model may also be useful for certain transportation problems. For example, it can be used to plan the capacity-expansion policy for two modes of transportation, such as passenger and freight trains, where passenger trains can be converted to handle the shipment of goods. In addition, the model can also be viewed as a production problem for two substitutable products or as an inventory problem for a single product produced and consumed in two separate regions. In the latter case, the demand for additional capacity in each period is defined as the demand for the product in each of the two regions. Furthermore, idle capacity is replaced by inventory, capacity construction is replaced by production, and capacity conversion is replaced by shipment of the product from one region to the other.

Many capacity-expansion and inventory models have been developed for the single facility problem with a finite number of discrete time periods. The first such model with time-dependent costs was proposed by Wagner and Whitin [13], who examined a dynamic version of the economic lot-size model. Many authors extended this model; for example, Zangwill [15,16], Manne and Veinott [10], Florian and Klein [4], and Rao [12].

Several models for two facilities, in which it was assumed that converted capacity is reconverted automatically, at no cost, at the end of each period, have been published. Appropriate references are Manne [9], Erlenkotter [2,3], Kalotay [7] and Fong and Rao [5]. In many applications (such as those mentioned before), it is more reasonable to assume that converted capacity is not reconverted at the end of each period. Kalotay [8] and Wilson and Kalotay [14] extended Kalotay's earlier work, and Merhaut [11] extended Erlenkotter's work [2] for this case. The model described in this paper is similar to [5] for the case when converted capacity is not reconverted at the end of each period.

In Section 1 we formulate the model, and in Section 2 a dynamic-programming approach is presented. In Section 3 some properties of an optimal solution are identified through a network-flow representation, and a computational procedure for nonnegative demand increments is developed. In Section 4 the procedure is modified for arbitrary demands. Finally, in Section 5 the model is applied to a cable-sizing problem.

1. FORMULATION

Let

i,j = indices for the two facilities.

t = an index for time period ($t = 1, 2, \ldots, T+1$, where T is the planning horizon).

r_{it} = The increment of demand for additional capacity of facility i at period t. For the present we assume that $r_{it} \geqslant 0$; however, this assumption is relaxed in Section 4. Also, for convenience, the r_{it}'s are assumed to be integers.

$$R_i(t_1,t_2) = \sum_{t=t_1}^{t_2} r_{it}, \quad \text{for } t_1 \leqslant t_2.$$

$$R_t = R_1(t,T) + R_2(t,T).$$

x_{it} = The amount of new construction of facility i at period t.

y_{it} = the amount of capacity associated with facility i converted at period t to satisfy demand for facility j. Once converted, the capacity becomes an integral part of facility j.

I_{it} = The amount of idle capacity of facility i at the beginning of period t (or, equivalently, the idle capacity at the end of period $t-1$, $t = 2, 3, \ldots, T+1$). We assume that the initial idle capacities are zero, $I_{i1} = 0$, and that shortages of capacity are not allowed; i.e., $I_{it} \geqslant 0$.

$c_{it}(x_{it})$ = the construction cost of x_{it}.

$g_{it}(y_{it})$ = the conversion cost of y_{it}.

$h_{it}(I_{i,t+1})$ = the holding cost of idle capacity $I_{i,t+1}$ from period t to period $t+1$.

All cost functions $c_{it}(\bullet)$, $g_{it}(\bullet)$, and $h_{it}(\bullet)$ are assumed to be nondecreasing and concave.

The problem can be stated as follows:

$$\text{(1.1)} \qquad \underset{x_{it},\, y_{it}}{\text{Minimize}} \left\{ \sum_{t=1}^{T} \sum_{i=1}^{2} [c_{it}(x_{it}) + g_{it}(y_{it}) + h_{it}(I_{i,t+1})] \right\}$$

$$\text{(1)} \quad \left.\begin{array}{ll} \text{(1.2)} & I_{i,t+1} = I_{it} + x_{it} + y_{jt} - y_{it} - r_{it}, \\ \text{(1.3)} & x_{it} \geqslant 0, \quad y_{it} \geqslant 0, \quad I_{it} \geqslant 0, \\ \text{(1.4)} & I_{i1} = 0, \\ \text{(1.5)} & I_{i,T+1} = 0, \end{array}\right\} \begin{array}{l} i = 1,2, \\ j = 1,2\ (j \neq i), \\ t = 1,2, \ldots, T, \end{array}$$

where, for convenience, we assume that $c_{it}(0) = g_{it}(0) = h_{it}(0) = 0$. The objective (1.1) is to minimize the total costs incurred, specifically the construction, conversion, and holding costs over all periods. The constraints (1.2) state that the idle capacity of facility i at the beginning of period $t+1$ is equal to the idle capacity at the beginning of period t plus the net change of the capacity of facility i at period t minus the demand increment for facility i at t. The constraints $I_{it} \geqslant 0$ and $I_{i1} = 0$ are introduced by the assumptions. Furthermore, the idle capacities at the end of period T, i.e., $I_{i,T+1}$, are fixed at zero, since for $r_{it} \geqslant 0$ any other solution can readily be modified to satisfy this constraint without increasing the total costs.

2. A DYNAMIC-PROGRAMMING APPROACH

The constraints (1.2) through (1.5) form a nonempty convex set ($x_{it} = r_{it}$, $y_{it} = I_{it} = 0$, $\forall i$ and t, is a feasible solution). Since each component of the objective function (1.1) is nondecreasing with a finite value at zero and all variables are required to be nonnegative, there exists a finite optimal solution to problem (1). Furthermore, since (1.1) is concave, there exists an extreme point solution which minimizes (1.1); we shall concentrate on finding such a solution.

We define a *capacity point* as a period t for which $I_{1t}I_{2t} = 0$. The constraints (1.2) can be shown to be totally unimodular (Hu [6]). Hence, since the r_{it}'s are integers, any extreme-point solution consists of integer values for all variables, and only the following idle-capacity values of capacity points need be considered:

(2)
$$\begin{aligned} &I_{11} = I_{21} = 0, \\ &\begin{cases} I_{1t} = I_{2t} = 0, \\ I_{1t} = 0 \text{ and } I_{2t} = 1, 2, \ldots, R_t, \\ I_{2t} = 0 \text{ and } I_{1t} = 1, 2, \ldots, R_t, \end{cases} \quad t = 2, 3, \ldots, T, \\ &I_{1,T+1} = I_{2,T+1} = 0. \end{aligned}$$

Since at most one idle-capacity value is positive at a capacity point, we conveniently define the values in (2) by a single parameter α_t:

(3)
$$\begin{aligned} &\alpha_t = 1 \text{ if } I_{1t} = I_{2t} = 0, \\ &\alpha_t = m + 1 \text{ if } I_{1t} = 0 \text{ and } I_{2t} = m \ (m = 1, 2, \ldots, R_t), \\ &\alpha_t = m + R_t + 1 \text{ if } I_{2t} = 0 \text{ and } I_{1t} = m \ (m = 1, 2, \ldots, R_t). \end{aligned}$$

Thus, α_t may take on any of the values $1, 2, \ldots, 2R_t + 1$.

We now describe a dynamic-programming approach that can be used to solve problem (1). Let

$d_{uv}(\alpha_u, \alpha_{v+1})$ = the minimal cost associated with an optimal policy during periods $u, u+1, \ldots, v$, when u and $v+1$ are two successive capacity points with idle-capacity values defined by α_u and α_{v+1}. More specifically:

(4)
$$d_{uv}(\alpha_u, \alpha_{v+1}) = \underset{x_{it}, y_{it}}{\text{minimum}} \left\{ \sum_{t=u}^{v} \sum_{i=1}^{2} [c_{it}(x_{it}) + g_{it}(y_{it}) + h_{it}(I_{i,t+1})] \right\}$$

where

(i) The constraints (1.2) and (1.3) are satisfied for $t = u, u+1, \ldots, v$,

(ii) $I_{1t}I_{2t} > 0$, $t = u+1, u+2, \ldots, v$,

(iii) I_{1u} and I_{2u} are given by α_u, and $I_{1,v+1}$ and $I_{2,v+1}$ are given by α_{v+1}.

Furthermore, let

$f_t(\alpha_t)$ = the cost of an optimal policy over periods $t, t+1, \ldots, T$, given that period t is a capacity point and I_{1t} and I_{2t} are specified by α_t.

Assume that all the *subproblem values* $d_{uv}(\alpha_u, \alpha_{v+1})$ are known. The following dynamic-programming formulation is then obtained:

$$f_{T+1}(\alpha_{T+1}) = 0, \quad \alpha_{T+1} = 1,$$

$$(5) \qquad f_u(\alpha_u) = \min_{\substack{u \le v \le T \\ 1 \le \alpha_{v+1} \le 2R_{v+1}+1}} [d_{uv}(\alpha_u, \alpha_{v+1}) + f_{v+1}(\alpha_{v+1})],$$

$$u = T, T-1, \ldots, 1,$$
$$\alpha_u = 1, 2, \ldots, 2R_u + 1 \ (u \neq 1),$$
$$\alpha_1 = 1.$$

The first term of the minimand is the minimum cost of the optimal policy during periods $u, u+1, \ldots, v$, given that u and $v+1$ are two successive capacity points with idle capacities α_u and α_{v+1}. The second term is the optimal cost for periods $v+1, v+2, \ldots, T$, given α_{v+1}. Thus, searching for the optimal values of v and α_{v+1} results in $f_u(\alpha_u)$. As shown in Figure 1, this formulation may also be viewed as the shortest-path problem for an acyclic network in which the nodes represent all possible values of capacity points. Each node is described by two values (t, α_t), where t is the time period and α_t is the associated capacity-point value. From each node (u, α_u) there emanates a directed link to any node $(v+1, \alpha_{v+1})$ for $v \geq u$ with an associated cost of $d_{uv}(\alpha_u, \alpha_{v+1})$. It can be shown by simple enumeration and algebraic manipulations that the total numer of links N in Figure 1 is

$$(6) \qquad N = \sum_{t=2}^{T-1} \left[2R_t \sum_{\tau=t+1}^{T} 2R_\tau \right] + T \sum_{t=2}^{T} 2R_t + \frac{(T+1)T}{2}.$$

Most of the computational effort involved in solving (5) is spent on the solution of the subproblems (4).

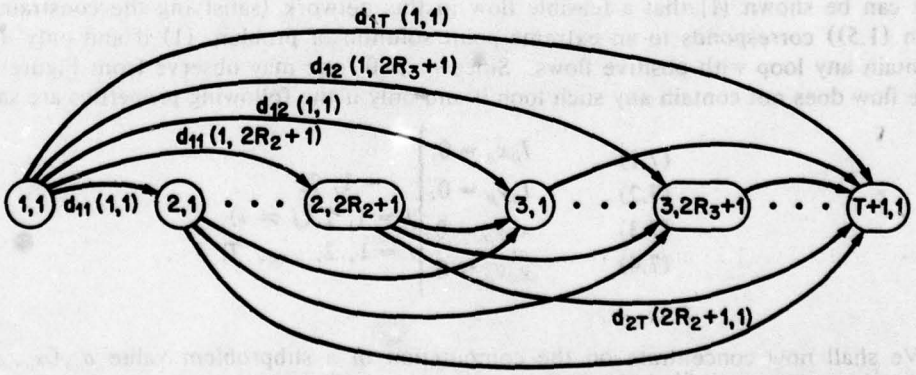


FIGURE 1. The shortest-path problem equivalent to the dynamic-programming formulation.

3. THE SOLUTION FOR NONNEGATIVE DEMAND INCREMENTS

We shall first characterize several properties of an extreme-point solution of problem (1) when $r_{it} \geq 0$. To derive these properties it is convenient to view (1) as a single-commodity network problem, as shown in Figure 2. The network includes a single source (node 0) with a supply of R_1. There are $2T$ additional nodes, each denoted by either (i,t) or (j,t), where i (or j) specifies the facility and t indicates the time period. At each node (i,t) there is an external demand of r_{it}. The nodes are connected by directed links, where the flows along these links represent the construction, conversion, and idle-capacity variables. Specifically, the nodes are connected as follows:

- A link from node 0 to each node (i,t), with flow x_{it};
- A link from each node (i,t) to $(i,t+1)$, with flow $I_{i,t+1}$;
- A link from each node (i,t) to (j,t), for $j \neq i$, with flow y_{it}.

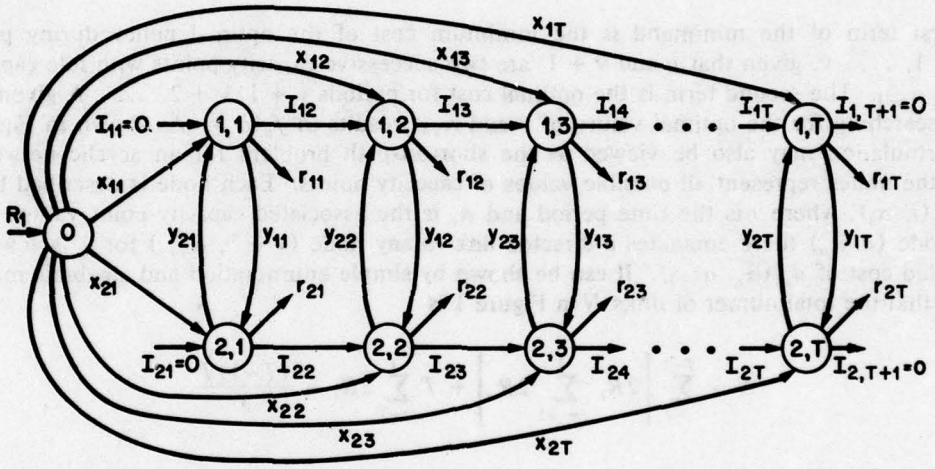


FIGURE 2. A network-flow representation of the capacity-expansion problem.

It can be shown [1] that a feasible flow in this network (satisfying the constraints (1.2) through (1.5)) corresponds to an extreme-point solution of problem (1) if and only if it does not contain any loop with positive flows. Since $r_{it} \geq 0$, one may observe from Figure 2 that a feasible flow does not contain any such loop if and only if the following properties are satisfied:

(7)

$$\left.\begin{aligned} &(7.1) \quad I_{it}x_{it} = 0, \\ &(7.2) \quad I_{it}y_{jt} = 0, \\ &(7.3) \quad x_{it}y_{jt} = 0, \\ &(7.4) \quad y_{it}y_{jt} = 0. \end{aligned}\right\} \begin{aligned} &i = 1, 2, \\ &j = 1, 2\ (j \neq i), \\ &t = 1, 2, \ldots, T, \end{aligned}$$

We shall now concentrate on the computation of a subproblem value $d_{uv}(\alpha_u, \alpha_{v+1})$ as defined by (4) which satisfies (7). The subnetwork associated with the subproblem is given in Figure 3. From (7) it follows that $x_{it} = y_{it} = 0$ for $i = 1, 2$ and $t = u+1, u+2, \ldots, v$. Therefore, only x_{1u}, x_{2u}, y_{1u}, and y_{2u} may be positive in the solution of (4).

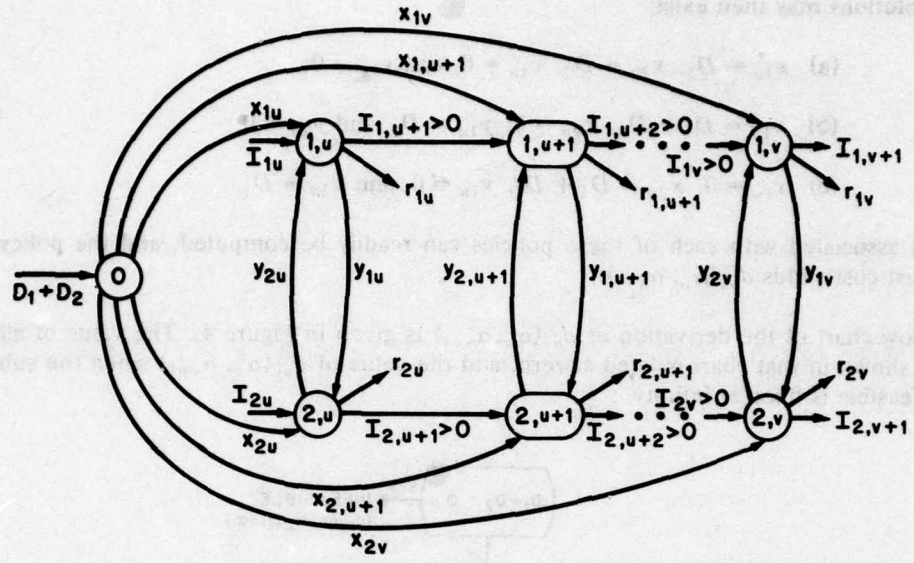


FIGURE 3. A network-flow representation of a subproblem.

Let D_i be the capacity change of facility i during periods $u, u+1, \ldots, v$, i.e.,

$$D_i = \sum_{t=u}^{v} (x_{it} + y_{jt} - y_{it}) = I_{i,v+1} + R_i(u,v) - I_{iu}, \quad i = 1, 2, \quad j = 1, 2 \ (j \neq i). \tag{8}$$

Since $D_1 + D_2 = \sum_{t=u}^{v} (x_{1t} + x_{2t})$, $D_1 + D_2 < 0$ implies that at least one of the variables x_{it} is negative, hence, the corresponding subproblem is infeasible. (An infeasible subproblem is defined as a subproblem for which there is no solution which satisfies constraints (1.2) through (1.5) and the properties of (7).)

Suppose that $D_1 + D_2 \geq 0$, $I_{1u} = 0$, and $I_{2u} > 0$. From (7.1), $x_{2u} = 0$, so that $x_{1u} = D_1 + D_2$. From (7.2), $y_{1u} = 0$, hence $D_2 = x_{2u} + y_{1u} - y_{2u} = -y_{2u}$, or $y_{2u} = -D_2$. The subproblem (4) is therefore infeasible for two cases: $D_2 > 0$, or, by (7.3), $D_2 < 0$ and $D_1 + D_2 \neq 0$. For all other values of D_1 and D_2 the subproblem has a unique feasible solution $x_{1u} = D_1 + D_2$, $y_{2u} = -D_2$, and $x_{2u} = y_{1u} = 0$. The optimal value of the subproblem is then

$$d_{uv}(\alpha_u, \alpha_{v+1}) = c_{1u}(D_1 + D_2) + g_{2u}(-D_2) + \sum_{t=u}^{v} \sum_{i=1}^{2} h_{it}(I_{i,t+1}),$$

where the idle capacities $I_{i,t+1}$ are obtained by appropriate substitutions in (1.2). When $I_{2u} = 0$ and $I_{1u} > 0$, the solution is obtained in the same manner, with the indices 1 and 2 interchanged.

Assume now that $I_{1u} = I_{2u} = 0$, in which case, by (8), $D_1 \geqslant 0$ and $D_2 \geqslant 0$. Up to three feasible solutions may then exist:

(a) $x_{1u} = D_1$, $x_{2u} = D_2$, $y_{1u} = 0$, and $y_{2u} = 0$;

(b) $x_{1u} = D_1 + D_2$, $x_{2u} = 0$, $y_{1u} = D_2$, and $y_{2u} = 0$;

(c) $x_{1u} = 0$, $x_{2u} = D_1 + D_2$, $y_{1u} = 0$, and $y_{2u} = D_1$.

The costs associated with each of these policies can readily be computed, and the policy with the smallest cost yields $d_{uv}(\alpha_u, \alpha_{v+1})$.

A flow chart of the derivation of $d_{uv}(\alpha_u, \alpha_{v+1})$ is given in Figure 4. The value of all variables not shown in that chart is fixed at zero, and the value of $d_{uv}(\alpha_u, \alpha_{v+1})$ when the subproblem is infeasible is fixed at infinity.

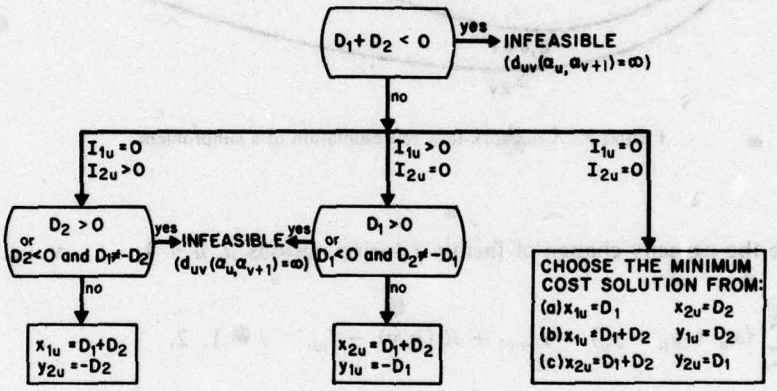


FIGURE 4. Derivation of an optimal policy for a subproblem.

As a final point, it can be shown from the network representation in Figure 2 that (5) need not always be computed for all possible values of α_u. Specifically, the values of interest for I_{it} $(i = 1, 2)$ are

$$(9) \qquad I_{it} = \begin{cases} 0, \\ R_i(t, \tau), \quad j = 1, 2\ (j \neq i), \\ R_i(t, \tau) + R_j(t', \tau'), \quad t \leqslant \tau \leqslant T, \quad t \leqslant t' \leqslant \tau, \quad t' \leqslant \tau' \leqslant T. \end{cases}$$

Since these values are sums of demands, this observation may reduce the computational effort, especially when the demands increase linearly.

4. MODIFICATIONS FOR ARBITRARY DEMAND INCREMENTS

When the r_{it}'s are allowed to be negative, $I_{i,T+1}$ may be positive in all optimal solutions of problem (1). However, since all cost functions are nondecreasing, there exists an optimal solution in which $I_{i,T+1} \leqslant \max_{1 \leqslant t \leqslant T} [R_i(1,t)] - R_i(1,T)$. Let $T' = T + 1$, with $c_{iT'}(\bullet) =$

$g_{iT'}(\bullet) = h_{iT'}(\bullet = 0$ and $r_{iT'} = \max_{1 \leqslant t \leqslant T} [R_i(1,t)] - R_i(1,T)$. Obviously, any optimal solution for the T' - period problem is also optimal for the original T-period problem. Furthermore, there exists an optimal solution in which $I_{i,T'+1} = 0$ $(i = 1,2)$. The problem is then solved by the dynamic-programming formulation (5), where R_t is redefined as $R_t = R_1(t, T') + R_2(t, T')$.

The difficulties in solving (5) arise from the computational effort involved in solving the subproblems $d_{uv}(\alpha_u, \alpha_{v+1})$. When the r_{it}'s are allowed to be negative, an extreme-point solution (or, equivalently, a feasible flow on the network given in Figure 2 which does not contain any loop with positive flows) does not imply that the properties of (7) are satisfied. However, if we examine a subproblem such as the one in Figure 3, the following properties must be satisfied to obtain a flow which does not contain any loop with positive flows:

(10)

$$(10.1) \quad x_{it_1} x_{it_2} = 0, \; u \leqslant t_1, t_2 \leqslant v \; (t_1 \neq t_2), \quad i = 1,2;$$

$$(10.2) \quad y_{it_1} y_{jt_2} = 0, \; u \leqslant t_1, \; t_2 \leqslant v, \quad i, j = 1,2,$$

and either $i \neq j$ or $t_1 \neq t_2$;

$$(10.3) \quad x_{1t_1} x_{2t_2} y_{it_3} = 0, \; u \leqslant t_1, \; t_2, \; t_3 \leqslant v, \quad i = 1, 2.$$

For example, if (10.1) is violated, then the flows x_{it_1}, I_{i,t_1+1}, ..., I_{it_2}, x_{it_2} (when $t_1 < t_2$) form a loop with positive flows. Thus, we need to consider only subproblems in which there is at most one new construction for each facility (10.1) and at most one conversion (10.2). Furthermore, if two constructions are being considered (one per facility), conversion is then not allowed (10.3).

In contrast to the case of $r_{it} \geqslant 0$, optimal construction and conversion may take place on any period t, $u \leqslant t \leqslant v$. We shall now summarize the possible policies (satisfying the constraints of (1) and the properties of (10)) which need to be examined in order to solve $d_{uv}(\alpha_u, \alpha_{v+1})$. These policies depend on the capacity change D_i, as defined by (8), and are summarized in Table 1. In this table it is assumed that $D_1 + D_2 \geqslant 0$, since otherwise the associated subproblems are infeasible. All the variables not mentioned for a given policy are fixed at zero, and t_1 and t_2 are time periods between u and v. The dashes represent infeasible realizations.

To solve a subproblem with given values of D_1 and D_2, all policies shown in the appropriate column have to be evaluated. Feasible values of t_1 and t_2 include all values which satisfy the constraints $I_{it} > 0$ for $i = 1, 2$ and $t = u + 1, u + 2, \ldots, v$. Hence, a significant amount of computation may be needed to obtain all feasible policies and compare the costs associated with these policies. The computational effort can be reduced in certain cases, for example, when the cost functions are uniformly decreasing with t (such as cost functions which depend on t only through a discount factor).

As a final comment, it may be of interest to examine problems where $I_{i1} > 0$. The dynamic-programming equations can readily be applied if we redefine $\alpha_1 = 1$, when I_{11} and I_{21} are equal to their initial nonzero values, and the derivation of $d_{uv}(\alpha_u, \alpha_{v+1})$ described in this section can be used. The algorithm can also be extended to similar models, for example, when capacity shortages are allowed.

TABLE 1. *Possible Policies for Arbitrary Demand Increments*

Policy \ D_1, D_2	$D_1 < 0$ $D_2 > 0$	$D_1 > 0$ $D_2 < 0$	$D_1 = 0$ $D_2 > 0$	$D_1 > 0$ $D_2 = 0$	$D_1 > 0$ $D_2 > 0$	$D_1 = 0$ $D_2 = 0$
x_{1t_1}	–	–	0	D_1	D_1	0
x_{2t_2}			D_2	0	D_2	0
x_{1t_1}	$D_1 + D_2$	–	D_2	D_1	$D_1 + D_2$	0
y_{1t_2}	D_2		D_2	0	D_2	0
x_{2t_1}	–	$D_1 + D_2$	D_2	D_1	$D_1 + D_2$	0
y_{2t_2}		D_1	0	D_1	D_1	0
x_{2t_1}	$D_1 + D_2$	–	D_2	–	–	0
y_{1t_2}	$-D_1$		0			0
x_{1t_1}	–	$D_1 + D_2$	–	D_1	–	0
y_{2t_2}		$-D_2$		0		0

5. APPLICATION TO A CABLE-SIZING PROBLEM

Cable-sizing problems, described in the introduction, often occur in network-planning applications. We assume that once an expensive cable (cable 1) is used for demand associated with a cheaper cable (cable 2), it cannot be reconverted to serve demand associated with the expensive cable. Given the demands for the two cables, one needs to plan the capacity expansions. The decisions to be made include what cables should be installed, when, and how large they should be.

We shall assume that all demand increments are nonnegative, so that the solution technique developed in Section 3 can be applied. Since cable 2 cannot be used for the demand associated with cable 1, $y_{2t} = 0$ for $t = 1, 2, \ldots, T$. Therefore $I_{2t} \leqslant R_2(t, T)$, and the number of possible values of capacity points (see (2)) can be reduced. Furthermore, the flow chart described in Figure 4 can be simplified as follows:

- When $I_{1u} = 0$ and $I_{2u} > 0$, $D_2 \neq 0$ implies that the subproblem is infeasible.
- When $I_{1u} = 0$ and $I_{2u} = 0$, possibility (c) need not be considered.

The model has been applied to pairs of cables chosen from four types of cables. The construction cost functions are approximated from discrete data points. We assume that these functions are composed of a fixed cost per installation (about the same for all cable types, slightly varied for sensitivity analyses) plus a constant cost per wire pair (different for each cable type). Specifically, we assume

(a) $3600 + 60x$, (b) $4000 + 33x$,

(c) $4000 + 24x$, (d) $3700 + 18x$,

where x is the number of wire pairs included in the cable. A single-year discount factor of 0.93 is assumed to account for the time value of money. Furthermore, we assume that $g_{it}(\bullet) = h_{it}(\bullet) = 0$. All examples are for a planning horizon of $T = 20$ years.

In Table 2 we show examples for a linearly growing demand (i.e., r_{it} is the same for all t), where the demand increments are chosen to represent areas with moderate growth. Under the "Installation" headings, 1(250) means that a cable with 250 wire pairs is installed in period 1. Under the "Conversions" heading, 1-20 means that the demand increments associated with cable 2 for periods 1-20 are satisfied by cable 1.

TABLE 2. *Examples for Linearly Growing Demands*

Cost Functions		Demands		Installations of the Expensive Cable (Cable 1)	Installations of the Cheap Cable (Cable 2)	Conversions from Cable 1 to Cable 2	Optimal Cost
Cable 1	Cable 2	r_{1t}	r_{2t}				
(a)	(d)	50	50	1(250),6(250),11(250),16(250)	1(500),11(500)		65,651
(a)	(d)	50	20	1(250),6(250),11(250),16(250)	1(240),13(160)		57,579
(a)	(d)	50	10	1(270),6(250),11(250),16(250)	3(180)	1-2	54,007
(b)	(c)	50	50	1(500),6(500),11(500),16(500)		1-20	51,586
(b)	(c)	50	30	1(480),7(560),14(560)		1-20	43,136
(b)	(c)	50	10	1(420),8(420),15(360)		1-20	34,356
(a)	(b)	50	50	1(250),6(250),11(250),16(250)	1(350),8(350),15(300)		76,743
(a)	(b)	50	40	1(290),6(250),11(290),16(250)	2(360),12(360)	1,11	72,282
(a)	(b)	50	25	1(275),6(250),11(275),16(250)	2(225),12(225)	1,11	64,798
(a)	(b)	50	20	1(350),6(350),11(350),16(350)		1-20	61,903
(b)	(d)	50	50	1(400),8(350),15(300)	2(450),11(500)	1	48,709
(b)	(d)	50	40	1(430),8(350),15(300)	3(360),12(360)	1-2	45,966
(b)	(d)	50	35	1(455),8(350),15(510)	4(385)	1-3,15-20	44,463
(b)	(d)	50	30	1(440),8(350),15(480)	4(330)	1-3,15-20	42,813
(b)	(d)*	50	30	1(480),7(560),14(560)		1-20	43,136
(b)	(d)†	50	30	1(350),8(350),15(300)	1(330),12(270)		43,432
(b)	(d)	50	25	1(450),7(525),14(525)		1-20	40,949

* "All conversion" policy enforced.

† "No conversion" policy enforced.

The results for the cables with cost functions (a) and (d) suggest that the demand associated with each cable should be satisfied primarily by installation of the appropriate cable. In contrast to these results, when cost functions (b) and (c) are assumed, only cable (b) is installed. The different policies result from the differences in the variable costs for an additional wire pair.

The examples for cost functions (a) and (b) reveal an interesting observation. When the demand associated with cable (b) is relatively high (about 25 wire pairs per period, or more), it is satisfied by installing cable (b) (except for minor conversions). However, when the demand for (b) is 20 wire pairs per period or less, all the demand is satisfied by cable (a). Let a *mixed policy* be one in which the demand for cable 2 is satisfied by each of the two cables for a substantial number of periods. Thus, a mixed optimal policy may only exist for a narrow range of demand increments for cable (b), somewhere between 20 to 25 wire pairs per period. The results for cost functions (b) and (d) also indicate that the range of demand increments for which mixed policies are optimal is quite narrow.

The examples in Table 2 suggest that a good heuristic is to choose the best from the following two strategies:

- The optimal policy when all the demand is satisfied by the more expensive cable;
- The optimal policy when all demand associated with the cheaper cable is satisfied by installation of the cheaper cable.

These optimal policies can be found by applying efficient algorithms, such as those given in [10], designed for problems with a single facility.

The model has also been applied to examples with convex and concave growing demands. The results again suggest that good heuristics can be designed, based on the solution of several single-facility problems per example. The examples were solved on an IBM 370/168 computer, and on the average it took about 10 s to solve each example. About 3000 subproblems with feasible solutions were computed in 1 s. Furthermore, only about 3% of the subproblems were feasible, and the time spent on infeasible subproblems was negligible.

Several approximations can be implemented to reduce the computational effort, for example,

- Consider larger increments of demand as one unit;
- Set $d_{uv}(\alpha_u, \alpha_{v+1}) = \infty$ whenever $v - u < I$ (where I is a positive integer); i.e., limit the number of constructions and conversions.

Our computational experience suggests that the additional objective-function cost incurred by implementing such approximations may be small.

REFERENCES

[1] Dantzig, G.B., *Linear Programming and Extensions* (Princeton University Press, Princeton, New Jersey, 1963) pp. 352-357.

[2] Erlenkotter, D., "Two Producing Areas — Dynamic Programming Solutions," in *Investments for Capacity Expansion: Size, Location, and Time Phasing*, A.S. Manne, ed., (MIT Press, Cambridge, Massachusetts, 1967) pp. 210-227.

[3] Erlenkotter, D., "A Dynamic Programming Approach to Capacity Expansion with Specialization," Management Science *21*, 360-362 (1974).

[4] Florian, M., and M. Klein, "Deterministic Production Planning with Concave Costs and Capacity Constraints," Management Science *18*, 12-20 (1971).

[5] Fong, C.O., and M.R. Rao, "Capacity Expansion with Two Producing Regions and Concave Costs," Management Science *22*, 331-339 (1975).

[6] Hu, T.C., *Integer Programming and Network Flows* (Addison Wesley, Reading, Massachusetts, 1969) pp. 124-127.

[7] Kalotay, A.J., "Capacity Expansion and Specialization," Management Science *20*, 56-64 (1973).

[8] Kalotay, A.J., "Joint Capacity Expansion without Rearrangement," Operational Research Quarterly *26*, 649-658 (1975).

[9] Manne, A.S., "Two Producing Areas — Constant Cycle Time Policies," in *Investments for Capacity Expansion: Size, Location, and Time Phasing*, A. S. Manne, (MIT Press, Cambridge, Massachusetts, 1967) pp. 193-209.

[10] Manne, A.S., and A.F. Veinott, "Optimal Plant Size with Arbitrary Increasing Time Paths of Demand," in *Investments for Capacity Expansion: Size, Location, and Time Phasing*, A.S. Manne, ed. (MIT Press, Cambridge, Massachusetts, 1967) pp. 178-190.

[11] Merhaut, J.M., "A Dynamic Programming Approach to Joint Capacity Expansion without Rearrangement," M. Sc. Thesis, Graduate School of Management, University of California, Los Angeles (1975).

[12] Rao, M.R., "Optimal Capacity Expansion with Inventory," Operations Research *24*, 291-300 (1976).

[13] Wagner, H.M., and T.M. Whitin, "Dynamic Version of the Economic Lot Size Model," Management Science *5*, 89-96 (1958).

[14] Wilson, L.O., and A.J. Kalotay, "Alternating Policies for Nonrearrangeable Networks," INFOR *14*, 193-211 (1976).

[15] Zangwill, W.I., "A Deterministic Multiperiod Production Scheduling Model with Backlogging," Management Science *13*, 105-119 (1966).

[16] Zangwill, W.I., "A Backlogging Model and a Multiechelon Model for a Dynamic Economic Lot Size Production System — A Network Approach," Management Science *15*, 506-527 (1969).

ON A SINGLE-SERVER QUEUE WITH STATE-DEPENDENT SERVICE

J.G. Shanthikumar

Department of Industrial Engineering
University of Toronto
Toronto, Ontario, Canada

ABSTRACT

This paper discusses a class of queueing models in which the service time of a customer at a single server facility is dependent on the queue size at the onset of its service. The Laplace transform for the wait in queue distribution is derived and the utilization of the server is given when the arrival is a homogeneous Poisson process.

INTRODUCTION

There is increasing attention in the queueing literature to the study of the systems in which the service characteristics change dynamically to accomodate variations in the system state. In this paper we wish to extend Harris' [5,6] two-state, state-dependent M/M/1 queueing model to the two-state, state-dependent M/G/1 model, where the service time of a customer is sampled from the arbitrary distributions $B(.)$ or $B_1(.)$, depending on whether there are any customers behind him or not at the onset of his service. Harris [5,6] studied the M/G/1 queueing system in which the service time parameter of a customer is a stochastic process dependent on the number in the queue at the moment his service is begun. Some general theory was developed and three special cases were also considered in [5,6]. One of them is the two-state, state-dependent M/M/1 model, where the service time of a customer is exponentially distributed with parameters μ or μ_1, depending on whether there are any customers behind him or not at the onset of his service. For this model, Harris obtained the probability distribution for the number in the system, and recently Brill and Posner [1] used system point theory and obtained the distribution function for the waiting time.

In this note, using the appropriate embedded Markov chain, the Laplace transform for the waiting time in the two-state, state-dependent M/G/1 queue will be obtained. It will also be shown that the result obtained by Harris [5] for the probability of no waiting is independent of the form of the distribution function $B(.)$, as long as $B_1(.)$ is exponential.

GENERAL MODEL

Consider arrivals of customers at a single server facility at times $\tau_1, \tau_2, \ldots$ $(\tau_0 = 0 < \tau_1 < \tau_2 < \ldots)$, and let $T_n = \tau_n - \tau_{n-1} (n = 1, 2, \ldots)$, so that T_n denotes the interarrival time between the $(n-1)^{th}$ and n^{th} customers. Customers are serviced in order of arrival.

Let S_n denote the service time and W_n denote the waiting time in the queue of the n^{th} arriving customer measured from the time he joined the queue until the instant he enters service. Throughout this paper we shall use $F_n(w) = \Pr\{W_n \leq w\}$.

Using the imbedded Markov chain approach with respect to the sequence $\{W_n\}$ of waiting times in the queue, we may easily write

$$W_{n+1} = [W_n + S_n - T_{n+1}]^+, \tag{1}$$

where $[y]^+ = \max(0,y)$.

A Queue with Service Time Depending on the Number in the System.

Customers arrive at a single server facility at times $\tau_1, \tau_2, \ldots (\tau_0 = 0 < \tau_1 < \tau_2 < \ldots)$, and $T_n = \tau_n - \tau_{n-1} (n=1, 2, \ldots)$ denotes the interarrival time between the $(n-1)^{th}$ and n^{th} customers. In this paper we shall assume that the sequence $\{T_n\}$ represents a set of mutally independent and identically distributed random variables with common distribution function $A(.)$. Customers are serviced in the order of arrival. The n^{th} arriving customer receives a service of S_n , where $\Pr\{S_n \leq x\} = B_1(x)$ if no newly arrived customers are behind him at the onset of his service, and $\Pr\{S_n \leq x\} = B(x)$ if at least one customer is behind him at the onset of his service.

This example with homogeneous Poisson arrival and general service time has been considered by Harris [5,6]. Brill and Posner [1] have considered the same problem with exponential service times.

From (1), for $w > 0$ and $n = 1, 2, \ldots$, we have

$$F_{n+1}(w) = \Pr\{W_n + S_n - T_{n+1} \leq w\}$$

and, rewriting,

$$F_{n+1}(w) = \Pr\{S_n \leq w + T_{n+1} - W_n\}.$$

Now, conditioning on the interarrival time $T_{n+1}(= x - w)$ and waiting time $W_n(= \alpha)$, we get

$$F_{n+1}(w) = \int_{x-w}^{\infty} \int_{\alpha=0^-}^{\infty} \Pr\{S_n \leq x - \alpha | W_n = \alpha, T_{n+1} = x-w\} dF_n(\alpha) dA(x-w).$$

Since, by definition of the model,

$$\Pr\{S_n \leq x-\alpha\} = \begin{cases} B_1(x-\alpha), & \alpha < x-w, \\ B(x-\alpha), & \alpha \geq x-w, \end{cases}$$

we can rewrite the above equation for $F_{n+1}(w)$ as follows:

$$\begin{aligned} F_{n+1}(w) &= \int_{x-w}^{\infty} \int_{\alpha=0-}^{x-w} B_1(x-\alpha)\, dF_n(\alpha) dA(x-w) \\ &+ \int_{x-w}^{\infty} \int_{\alpha=x-w}^{x} B(x-\alpha) dF_n(\alpha) dA(x-w), w > 0. \end{aligned} \tag{2}$$

$F_{n+1}(0)$ may be found from the condition $F_{n+1}(\infty) = 1$.

In more general cases, equation (2) is difficult to investigate thoroughly. Attention will therefore be focused on the analysis of more specific subclasses of models. If we assume that the arrivals represent a homogeneous Poisson process of rate λ, then $A(z) = 1 - \exp(-\lambda z), z > 0$.

THE M/G, G/1 MODEL

In the following analysis we will assume that the arrivals represent a homogeneous Poisson process with rate λ. For $w > 0$, let the probability density function $f_n(w) = dF_n(w)/dw$, assuming that $F_n(w)$ is continuous and differentiable for $w > 0$, and for $w = 0, f_{0,n} = F_n(0)$ is the probability that the n^{th} customer does not wait in the queue. Differentiating equation (2) with respect to w, with $dA(z) = \lambda \exp(-\lambda z)dz$, we get

$$
\begin{aligned}
(3)\quad f_{n+1}(w) = & -\int_{x=w}^{\infty} B_1(w) f_n(x-w)\,dA(x-w) - \lambda B_1(w) f_{0,n} \\
& + \lambda \int_{x=w}^{\infty} \int_{\alpha=0-}^{x-w} B_1(x-\alpha)\,dF_n(\alpha)\,dA(x-w) \\
& + \int_{x=w}^{\infty} B(w) f_n(x-w)\,dA(x-w) + \lambda B(w) f_{o,n} \\
& + \lambda \int_{x=w}^{\infty} \int_{\alpha=x-w}^{x} B(x-\alpha)\,dF_n(\alpha)\,dA(x-w) \\
& - \int_{\alpha=0-}^{w} \lambda B(w-\alpha)\,dF_n(\alpha).
\end{aligned}
$$

Substituting (2) in (3) for $F_{n+1}(w)$ and $L_n(\lambda) = \int_{0-}^{\infty} \exp(-\lambda\alpha)\,dF_n(\alpha) = (1/\lambda)\int_0^{\infty} f_n(\alpha)\,dA(\alpha) + f_{0,n}$, we get

$$
(4)\quad f_{n+1}(w) = \lambda F_{n+1}(w) - \int_{\alpha=0-}^{w} \lambda B(w-\alpha)\,dF_n(\alpha) + \lambda [B(w) - B_1(w)] L_n(\lambda).
$$

Assuming now that the limiting distribution exists, the sequence $\{F_n\}$ will converge uniformly to F. Let $F(w) = \lim_{n\to\infty} F_n(w)$ and $f(w) = dF(w)/dw$ for $w > 0$. The stationary waiting time distribution may then be written from (4) as and

$$
\begin{aligned}
(5)\quad f(w) &= \lambda F(w) - \int_{\alpha=0-}^{w} \lambda B(w-\alpha)\,dF(\alpha) + \lambda [B(w) - B_1(w)] L(\lambda) \\
&= \lambda \int_{\alpha=0-}^{w} [1 - B(w-\alpha)]\,dF(\alpha) + \lambda [B(w) - B_1(w)] L(\lambda),
\end{aligned}
$$

where

$$
L(\lambda) = \int_{0-}^{\infty} \exp(-\lambda\alpha)\,dF(\alpha).
$$

Now, multiplying both sides of equation (5) by $\exp(-sw)$ and integrating over w from 0 to ∞, we get $L(s) - f_0 = \{\lambda[1 - B^*(s)]L(s)\}/s + \{\lambda [B^*(s) - B_1^*(s)]L(\lambda)\}/s$, where $L(s) = \int_{w=0}^{\infty} \exp(-sw)\,dF(w)$ is the Laplace transform of the waiting-time distribution $B^*(s) = \int_{w=0}^{\infty} \exp(-sw)\,dB(w)$, $B_1^*(s) = \int_{w=0}^{\infty} \exp(-sw)\,dB_1(w)$ and $f_0 = F(0)$.

Now, solving for $L(s)$, we have

$$
(6)\qquad L(s) = \frac{\lambda [B^*(s) - B_1^*(s)] L(\lambda) + s f_0}{s - \lambda [1 - B^*(s)]}
$$

Substituting $s = \lambda$ in (6) and using the condition $\lim_{s\to 0} L(s) = 1$, we will get two equations for the unknowns $L(\lambda)$ and f_0. Solving these two equations, we get

$$
f_0 = (1 - \rho) B_1^*(\lambda)/[\rho_1 - \rho + B_1^*(\lambda)]
$$

and

$$L(\lambda) = (1-\rho)/[\rho_1 - \rho + B_1^*(\lambda)],$$

where

$$\rho = \lambda \int_0^\infty x dB(x) \text{ and } \rho_1 = \lambda \int_0^\infty x dB_1(x).$$

Let U be the utilization of the server; then

$$U = 1 - f_0$$
$$= \{\rho_1 - \rho[1 - B_1^*(\lambda)]\}/[\rho_1 - \rho + B_1^*(\lambda)].$$

The mean waiting time W_q in the queue is then

$$W_q = \lim_{s \to 0}[-dL(s)/ds]$$
$$= \frac{W_0 B_1^*(\lambda) - (1-\rho_1) W_0 + (1-\rho) W_0'}{(1-\rho)[\rho_1 - \rho + B_1^*(\lambda)]},$$

where

$$W_0 = (\lambda/2)\int_0^\infty x^2 dB(x) \text{ and } W'_0 = (\lambda/2)\int_0^\infty x^2 dB_1(x).$$

THE M/M, G/1 MODEL

When the distribution of service time $B_1(.)$ is exponentially distributed with parameter μ_1 such that $B_1(z) = 1 - \exp(\mu_1 z), z > 0$, and $B(.)$ is still an arbitrary distribution function, we get

$$f_0 = (1-\rho)/(1 - \rho + \rho_1 + \rho_1^2 - \rho_1\rho).$$

The above result was first obtained by Harris [5,6] with the exponential assumption for $A(.), B(.)$, and $B_1(.)$. The above result shows that it is independent of the form of distribution for $B(.)$. We also have

$$W_q = \frac{W_0 \mu_1 \rho_1^2 + (1-\rho)(1+\rho_1)\rho_1}{\mu_1(1-\rho)(1-\rho+\rho_1+\rho_1^2-\rho_1\rho)}$$

THE M/M, M/1 MODEL

When $B(x) = 1 - \exp(-\mu x)$ and $B_1(x) = 1 - \exp(-\mu_1 x)$, for $x > 0$, we can easily derive the waiting time distribution from equation (5) using differential operators, as done in [1].

ACKNOWLEDGMENT

The author would like to thank the Canadian Commonwealth Scholarship and Fellowship Committee for providing financial support for this research.

BIBLIOGRAPHY

[1] Brill, P.H., and M.J.M. Posner, "Level Crossings in Point Processes Applied to Queues: Single Server Case," Operations Research *25*, 662-674 (1977).

[2] Buzacott, J.A., "The Effect of Queue Discipline on the Capacity of Queues with Service Time Depending on Waiting Times," INFOR *12*, 174-185 (1974).

[3] Gupta, S.K., "Queues with Hyper-Poisson Input and Exponential Service Time Distribution with State-Dependent Arrival and Service Rates," Operations Research *15*, 847-856 (1967).

[4] Hadidi, N., "Busy Period of Poisson Queues with State-Dependent Arrival and Service Rates," Journal of Applied Probability *11*, 842-848 (1974).

[5] Harris, C.M., "Queues with State-Dependent Stochastic Service Rates," Operations Research, *15*, 117-130 (1967).

[6] Harris, C.M., "Queues with State-Dependent Stochastic Service Rates," Ph. D. Dissertation, Polytechnic Institute of Brooklyn, Brooklyn, New York (1966).

[7] Harris, C.M., "A Queueing System with Multiple Service Time Distributions," Naval Research Logistics Quarterly *14*, 231-239 (1967).

[8] Libura, M., "On A One-Channel Queueing System with Service Time Depending on Waiting Time," Archiwum Automatyki Telemechaniki *16*, 279-286 (1971).

[9] Posner, M.J.M., "Single Server Queues with Service Time Depending on Waiting Time," Operations Research *21*, 610-616 (1973).

[10] Rosenshine, M., "Queues with State-Dependent Service Times," Transportation Research *1*, 97-104 (1967).

APPROXIMATION TECHNIQUES IN THE SOLUTION OF QUEUEING PROBLEMS

U. Narayan Bhat and Mohamed Shalaby

Department of Operations Research and Engineering Management
School of Engineering and Applied Science
Southern Methodist University
Dallas, Texas

Martin J. Fischer

Defense Communications Engineering Center
Reston, Virginia

ABSTRACT

In the study of complex queueing systems, analysis techniques aimed at providing exact solutions become ineffective. Approximation techniques provide an attractive alternative in such cases. This paper gives an overview of different types of approximation techniques available in the literature and points out their relative merits. Also, the need for proper validation procedures of approximation techniques is emphasized.

INTRODUCTION

Queueing theory has passed through several stages in its growth. During the first three decades of this century pioneering work was done in its foundation. Major analysis techniques for the investigation into the behavior of Markovian systems were developed during the next two decades. The 1950's saw investigations extended into problems related to non-Markovian systems. This trend continued well into the middle of the sixties. Until then, queueing theory, having been developed by mathematicians, probabilists, and statisticians, had grown with minimal interaction with applications. During the past ten years, the trend has been more toward applications and making queueing-theory results applicable. The two major areas receiving maximum attention during this period are optimization problems in queues and approximation techniques in the solution of queueing problems.

As the complexity of the systems being considered by applied scientists increases, finding effective solution techniques leading to exact solutions is becoming a difficult task. Approximation techniques provide an attractive alternative in such cases. Over the years several types of approximation techniques have been developed for the solution of queueing problems. It is our intention here to provide an overview of these techniques and discuss their relative merits.

Three different stages may be identified in the modeling and analysis of a queueing system. At the first stage a suitable mathematical model for the system is developed. The second stage concerns the identification of and investigation into the basic process underlying the

Research work of the first two authors was supported by the ONR/DCEC Contract N00014-75-0597, NR042-324.

model. At the third stage numerical results are obtained from the analysis of the process. Noting that an approximation can be introduced at any one of these stages, we may identify three major categories of approximation techniques, system approximation, process approximation, and numerical approximation.

In the following sections we shall discuss different techniques used in approximations for the solutions of queueing problems based on the above categorization. Since justifying approximate results is an integral part of the process, techniques for validating approximations are also discussed. Finally, comments are made about future prospects in this direction.

It should be pointed out, however, that the objective of the paper is not the categorization, but the understanding of different types of approximating procedures. As will be clear later, even though the approximation is initiated at a certain stage, that the net result is to impact the system at all stages of analysis. Consequently, the distinction between the techniques sometimes becomes unclear.

For purposes of convenience we use Kendall notation suitably modified to include finite capacity. For instance, A/B/C/D represents a system in which the symbols A, B, C, D stand for the interarrival time distribution, service time distribution, number of servers, and system capacity respectively. When dealing with systems with no limitations on capacity, D is dropped. Also, the time dependence of an element is indicated by writing it as a function of t.

In compiling a bibliography, the intention of the authors has been to include a representative list of references. We have also tried to be exhaustive, so as to make it useful to the reader. All omissions of significant papers are inadvertent rather than intentional.

SYSTEM APPROXIMATION

A system approximation is mainly a simplification of the system under study such that the behavior of the new system is strongly related to the original system. The four main elements in a queueing system are the arrival process, the queue discipline, the service process, and the system structure. These elements are described by their properties or attributes. Also, due to the complexity of some applications, such as networks of queues, we need to add a set of relations that hold among these elements which are the results of various assumptions. Hence, a system simplification may be characterized either as simplifying the system elements or relaxing the relational assumptions.

Simplification of system elements is at the heart of the practice of queueing theory. Many times, results may not be available for the exact representation of the system element model (such as the distribution for interarrival time or service time). Then the best available model is used to arrive at the best approximate result. The predominant use of the exponential distribution in practice is due to this approximating process. In an attempt to incorporate more-general interarrival-time and service-time distributions, Erlangian distribution and Erlangian mixtures have been extensively used. In this regard the papers by Luchak [68], Wishart [103], Kotiah et al. [63], and Schassberger [95] are significant. The first three of the above papers supply the practicality of the approach, whereas the last paper provides the theoretical basis for the procedure.

A common technique in system approximation is the use of a simpler system either to derive an approximate measure of performance or suitable grounds for them. For instance, Maaloe [69] uses simple relations existing between mean waiting times of M/M/1 and M/M/s

systems to provide an approximate value of the mean waiting time in an M/E_k/s system. Gross [36] examines the effect of using an M/M/s model to approximate a G/G/s model. His results indicate that when one estimates mean value measures of congestion, the sensitivity to the exponential assumption is more pronounced, whereas it is not as pronounced for cost optimization models. Chandy et al. [16,17] study a queueing network with a direct application of Norton's theorem which implies that the properties of a subsystem in a queueing network can be obtained by replacing all queues that are not of interest by a single queue with equivalent load characteristics (see also Sauer and Chandy [94]). Another approach in the treatment of queueing networks occurring in computer systems is that of Avi-Itzhak and Heyman [3]. First, exact results are obtained for a closed-system model in terms of cycle times and server utilization. These results are then used to develop approximate results for an open-system model. For other examples of the use of simpler systems see Ghosal [34] and Rosenshine [90].

Nonstationarity of the arrival process can also be effectively handled through approximations. Moore [78] provides methods for partitioning the time axis into intervals with stationary characteristics and approximates an M(t)/G/1 queue by an M/G/1 queue during these periods.

Using simpler systems, upper and lower bounds for system performance measures have been derived in several cases. Brosh [13] derives mean total time spent by a customer in a priority queueing system by essentially changing the priority level of the customer so as to provide a worse case and a better case. Brumelle [14] obtains bounds for mean waiting time in a G/G/s system by constructing two single-server systems; one of them uses a share of the original load to give an upper bound and the second uses a service rate s times faster than the original one to give a lower bound. A further improvement on the upper bound for mean waiting time in the system G/M/s is obtained by Brumelle [15] by the waiting time in an associated G/M/1 queue. Yu [104] bounds a multiserver queue with recurrent input and Erlang service times by a simple G/E_k/1 queue. Kotiah [62] uses a linear programming technique to provide bounds in Markovian systems.

In a series of articles, Stoyan ([97] and references cited in it) has studied internal and external monotonicity properties of systems G/G/1 and G/G/s. The internal monotonicity property refers to the properties of characteristics such as waiting-time distribution functions over a discrete index parameter, and the external monotonicity property refers to the relationship of the monotonicities of element (input and service) distributions to the queue characteristic distributions. These properties can be used to provide approximations for complex systems by finding comparable ones that are easier to analyze. An excellent review of these results is given in Stoyan [97] which includes a bibliography of sixty-seven articles, more than half of which are not found in English language journals. Two other recent papers on this topic are Rolski and Stoyan [89] and Bergman and Stoyan [6].

Simpler bounding systems can be obtained by modifications to the queue discipline. Under low traffic situations, Bloomfield and Cox [12] obtain lower bounds for mean waiting time by ignoring the waiting times of customers other than the one being considered. In the context of a traffic queue at a signalized road intersection Bhat and Prabhu [9] obtain upper and lower bounds by sweeping the traffic arriving during a green period to the right and left extremities of the period (see also Bhat, Wheeler, and Fischer [11]).

Replacing a general distribution by one that has the same moments is an appealing approach. Kuczura [64] approximates the overflow process of an M/M/s/s system by an interrupted Poisson process which is alternatively turned on and off for exponentially distributed lengths of time. The approximation is obtained by matching the first two or three moments of the two processes. To study the mean waiting time of a G/G/1 queue, Marchal and Harris [73]

use an $E_k/E_l/1$ queue and match the first four moments of the random variable representing the difference (service time − interarrival time).

A problem of great interest in telephone work relates to predicting the blocking probability of an overflow stream of traffic in a group of channels operating as a loss system. An approximation widely used is the equivalent random method, which replaces the system under consideration by an equivalent loss system with a Poisson input. For details of this method see Wilkinson [101], Cooper [20], and Holtzman [46].

There are queueing systems in which more than one class of customers share the resources. A relatively simple procedure to derive the performance measures of such systems is to consider the two classes separately and improve the accuracy of approximation by successively using the most recent results for one class in the derivation of results for the other (see Bhat and Raju [10]).

For approximating more complex systems, many of these different characteristics could be used at different stages. Some examples of such efforts may be found in papers such as Leibowitz [67], Halfin [39], Willemain [102], and Rosenshine and Chandra [91].

Many system approximations are heuristic in nature. The quality of such procedures depends very much on intuition and creativity. The justification for the use of heuristic methods is not that they are analytically sound, but that experimentation has proved they are useful in practice. The basic approach is to observe the system, to relate it to some other system with known behavior, and then to make an educated guess about the behavior of the original system. For instance, Cosmetatos [21] derives approximate formulae for the steady-state queue size and waiting-time distribution in the system GI/M/s by observing the similarity of the mean waiting-time curves drawn against the coefficient of variation of the interarrival-time distribution, when the traffic intensity is kept constant for different numbers of servers. By this procedure he obtains approximate results that are within 5% of the actual value. Bhat and Fischer [8] have derived approximate results such as blocking probability and waiting time in a two-class heterogeneous multiserver system with Poisson arrivals, in which one class acts as a loss system but the second acts as a delay system. A key to this procedure is the observation that the probability of blocking is relatively insensitive to the ratio of the service rates of each class, which allows them to assume equal service rates. Conolly [19] considers Poisson queues belonging to the class of generalized birth and death processes as essentially renewal models with "effective" interarrival and service times (actual intervals may be dependent on queue size).

Nozaki and Ross [86] provide an approximation for mean waiting time in a multiserver queue M/G/s by assuming the equilibrium distribution form for the remaining service time of customers in service at the time of arrival. The expression involves the distribution of the number of busy servers, for which an approximate formula similar to the exact distribution in the queue M/M/s is derived.

Given above are only some examples of the use of heuristic approaches in approximations. To some extent all approximations can be considered to have some heuristic elements in it; but in system approximations they are in abundance.

PROCESS APPROXIMATION

Representation of a mathematical model follows the identification of the system model. Many times, the basic process underlying the mathematical model is so complex that a direct

analysis does not become worthwhile for the situation. One alternative would be to simplify the system model itself as described above. The second alternative is to identify a simpler process, whose analysis is either known or can be derived, that has properties similar to the basic process. Diffusion approximation, fluid approximation, and the use of asymptotic or limiting results are examples of such procedures. System-approximation techniques described in the previous section can also be looked upon as a form of process approximation when the availability of a simpler underlying process is the motivation for such an effort. System approximation techniques suggested by Moore [78] and Bhat and Prabhu [9] are examples of such situations.

Fluid approximation, as suggested by Newell [84] is mostly an engineering approach. It starts with some crude and naive estimates and relationships between system elements, and improvements are made in them as the analysis proceeds. The essential concept is to consider the arrival and departure processes in the system as fluid flowing in and out of a reservoir. Because of its deterministic nature, when the ouput rate (service rate) is in excess of the input rate (arrival rate), the fluid approximation results in an empty queue. In view of this, a proper setting for its application would be a short-term analysis of a queue or the behavior of an oversaturated queue (when the arrival rate exceeds the service rate). Also, the particular significance of its usage would be when the arrival and service rates are time dependent. Then, if $A(t)$ and $D(t)$ are the arrival and departure processes, with rates $\lambda(t) = dA(t)/dt$ and $\mu(t) = dD(t)/dt$, respectively, an approximate expression for the queue length $Q(t)$ at time t can be given as

$$\begin{aligned} Q(t) &\approx Q(0) + A(t) - D(t) \\ &= Q(0) + \int_0^t \lambda(\tau)\, d\tau - \int_0^t \mu(\tau) d\tau. \end{aligned}$$

A stochastic analogue of the fluid approximation is the diffusion approximation. In this procedure we replace a queueing process with jump transitions, or with continuous and jump transitions, by a continuous process which reflects the main characteristics of the original process. Diffusion processes are governed by stochastic differential equations incorporating the infinitesimal mean and variance of the process. Let

$$\begin{aligned} E\{Q(t+\tau) - Q(t)\} &= \int^{t+\tau} [\lambda(x) - \mu(x)]dx \\ &\approx [\lambda(t) - \mu(t)]\tau, \end{aligned}$$

where the arrival and departure rates $\lambda(t)$ and $\mu(t)$ are considered to be nearly constant over time as compared to τ. The quantity $\lambda(t) - \mu(t)$ is known as the infinitesimal mean of the process at time t. Also, let

$$\sigma^2(t) = \text{Var}\,\{Q(t+\tau) - Q(t)\}/\tau$$

be the infinitesimal variance of the process. If we denote the distribution of the process $Q(t)$ by $f(x, t)$ (note that $Q(t)$ is considered to be a continuous process in this approximation), under this approximation the function $f(x, t)$ is assumed to satisfy the Fokker-Planck equation

$$\frac{\partial f(x,t)}{\partial t} = -[\lambda(t) - \mu(t)] \frac{\partial f(x,t)}{\partial x} + \frac{\sigma^2(t)}{2} \frac{\partial^2 (x,t)}{\partial x^2}.$$

Diffusion approximation is usually related to heavy traffic (service rate close to the arrival rate), since we need the time variable to be large as compared to intervals between transitions. Under heavy traffic, idle periods occur very infrequently, and therefore one could use the zero state as a reflecting barrier of the process without degrading the approximation much further. (Note that a diffusion process can drift toward states below zero, whereas a queueing process remains on the nonnegative side of the axis.)

The equilibrium distribution of the resulting process can be derived in most cases as $f(x) = \lim_{t\to\infty} f(x, t)$. If necessary it can be discretized by integrating it over the unit interval $n < x \leqslant n + 1$, or $n - 0.5 < x \leqslant n + 0.5$.

Gaver's analysis [32] of the virtual waiting time of an M/G/1 queue is one of the initial efforts using diffusion approximation for queueing systems. In this case the infinitesimal mean and variance for the process are $\lambda E(S) - 1$ and $\lambda E(S^2)$, respectively, where S is the service time. Newell [83] gives an extensive treatment of a time-dependent arrival process using the Fokker-Planck equation. Heyman [42] has extended Gaver's results to study the busy period of the queue M/G/1. The transient behavior of the G/G/1 queue has been approximated by Heyman [43], and the approximation has been extended to the G/G/k system by Halachmi and Franta [38] by similar techniques. If we denote the interarrival time by A and the service time by S, for G/G/1 the infinitesimal mean and variance are

$$\lambda(t) - \mu(t) = \left\{\frac{1}{E(A)} - \frac{1}{E(S)}\right\}t$$

and

$$\sigma^2(t) = \left\{\frac{\text{Var}(A)}{[E(A)]^3} + \frac{\text{Var}(S)}{[E(S)]^3}\right\}t.$$

For the queue G/G/k, these take the form

$$\lambda(t) - \mu(t) = \left\{\frac{1}{E(A)} - \frac{\min(x, k)}{E(S)}\right\}t$$

and

$$\sigma^2(t) = \left\{\frac{\text{Var}(A)}{[E(A)]^3} + \frac{\min(x, k)\ \text{Var}(S)}{[E(S)]^3}\right\}t,$$

where x is the state of the system.

Some of the other applications of diffusion approximation in queues can be found in Newell [85], who provides a general setting for the analysis of the behavior of a sequence of servers in series with finite storage in between. Other references that suggest and elaborate earlier applications can be found in Kimura [52], Newell [82], Cox and Miller [22] and Feller [25].

Diffusion approximation has also been successfully employed in the analysis of queueing networks. Appropriate references in this area are Kobayashi [58,59] and Reiser and Kobayashi [87]. Fischer's use of the procedure in analyzing alternating priority queues [26] and Gaver and Shedler's (1973) application in obtaining the processor utilization in a multiprogramming computer system [33] are evidence to the effectiveness of this approximation technique (see also [27,28]).

A different approach will be to observe that the process under study converges in some sense to a diffusion process. Iglehart [47] has shown that in the M/M/n queue, if we let the mean interarrival time approach zero as $n\to\infty$, then the queue length process (after proper scaling) and normalizing tends to the Ornstein-Uhlenbeck process. McNeil [77] considers a sequence of nonstationary birth and death processes $\{x_N(t)\}$ with input and output rates dependent on N. He has shown that $\lim_{N\to\infty} x_N(t)$ (after normalizing) corresponds to a nonstationary Ornstein-Uhlenbeck process. An additional reference in this class of efforts is Harrison [40], who considers a sequence of systems with increasing traffic intensities.

NUMERICAL APPROXIMATION

Numerical approximation can be defined as a simplification which is brought in while one actually manipulates the arithmetic expressions, leading to an evaluation of a certain measure. If we identify an approximation $\hat{x}$ as $\hat{x} = x + \delta$, where x is the corresponding exact value and δ is an unknown small quantity, then we call $\hat{x}$ a "point approximation" if δ is unrestricted in sign, and we call $\hat{x}$ a "one-sided approximation" (or an interval approximation) if δ is retricted in sign. Clearly, the more we know about the properties of δ the more reliable the approximation will be, and we would like δ to be as small as possible.

The queue G/G/1 presents many difficulties in deriving exact results for its performance measures. Several attempts have been made to obtain approximations. The more successful of these are the heavy-traffic approximation (a point approximation) and upper and lower bounds (giving an interval approximation) for the mean waiting time given by Kingman [53-55], Marshall [74,75], and Suzuki and Yoshida [98]. All these efforts are based on the fundamental relation

$$W_{n+1} = \max[O,\ W_n + S_n + T_n],$$

where W_n is the waiting time of the n^{th} customer, S_n is his service time, and T_n, the time interval between the $(n-1)^{st}$ and the n^{th} cutomer. Writing $U_n = S_n - T_n$ and denoting the idle period by I, one gets the result

$$E[W] = \frac{E[U^2]}{-2E[U]} - \frac{\pi_0 E[I^2]}{-2E[U]}, \tag{1}$$

where π_0 is the probability that an arrival finds the system empty, $\lim_{n\to\infty} W_n \equiv W$, and $\lim_{n\to\infty} U_n \equiv U$. Since exact values for $E[I]$ and $E[I^2]$ are not available except in cases such as exponential interarrival times, an upper bound for $E[W]$ can be obtained as

$$E[W] \leqslant \frac{\mathrm{Var}[T] + \mathrm{Var}[S]}{2(E[T] - E[S])}. \tag{2}$$

when ρ is close to 1, Kingman [53] has shown that the upper bound for $E[W]$ is a good approximation for $E[W]$ itself. Furthermore, by using the central limit theorem on the basic random variables $\{U_n\}$, he has also shown that under heavy traffic, the waiting-time distribution under equilibrium conditions is exponential.

Lower bounds for $E[W]$ have been derived by both Kingman [55] and Marshall [74.75]. Marshall shows that

$$E[W] \geqslant l,$$

where l is the unique solution of the equation

$$x = \int_{-x}^{\infty} [1 - K(u)]du, \quad (x \geqslant 0),$$

where $P[U \leqslant u] = K(u)$. Kingman's alternate bound [55] can be given as

$$E[W] \geqslant \frac{E[(U^+)^2]}{2(E[T] - E[S])},$$

where $U^+ = \max[0, U]$. Comparing the bounds, Kingsman points out that Marshall's bound is sharper in light traffic ($\rho << 1$) whereas his bound is sharper in heavy traffic. Nevertheless, it should be noted that both lower bounds require the knowledge of the distribution of U, whereas the Kingman upper bound depends only on the first two moments of the interarrival-time and service-time distributions. For a concise discussion of bounds and approximations reference can be made to Gross and Harris [37], Chapter 6.

Another approximation for $E[W]$ can be obtained by writing $\pi_0 \approx 1 - \rho$ and $E[I^2] \approx E[U^2]$, where ρ is the traffic intensity of the system. Then we get, from (1),

$$E[W] \approx \rho\left(\frac{E[T^2] + E[S^2] - 2E[T]E[S]}{2(E[T] - E[S])}\right) \tag{3}$$

Comparing these approxmations for systems with one of the interarrival-time or service-time distribution exponentials, Bhat [7] has shown that the simple approximation given in (3) is in fact better than the heavy-traffic approximation given by (2) except when $C_v[S] >> C_v[T]$, where C_v stands for the coefficient of variation.

An additional effort in providing a better approximation for $E[W]$ is that of Marchal [70], who incorporates the coefficient of variation of the service time distribution $C_v(S)$ by suggesting

$$E[W] \approx \left(\frac{1 + C_v^2[S]}{\rho^{-2} + C_v^2[S]}\right)\left(\frac{\text{Var}[T] + \text{Var}[S]}{2(E[T] - E[S])}\right),$$

which is identical with the Kingman heavy-traffic approximation when $\rho = 1$. Marchal has also provided an alternate lower bound,

$$E[W] \geqslant \left[\frac{\rho^2 C_v^2(S) + \rho(\rho - 2)}{2(1 - \rho)}\right] E[T]$$

which incorporates only ρ and the coefficient of variation of the service-time distribution (see also Marchal [71,72] and Kleinrock [56], Chapter 2).

Using Martingale theory, Ross [92] has derived upper and lower bounds for the mean delay in the *G/G/1 queue*. Even though they are somewhat sharper than the ones described above, they are much harder to evaluate.

Extending the Kingman upper bound (2), we may give the following bounds for $E[W]$ in the multiserver queue G/G/s:

$$E[W] \leqslant \frac{\text{Var}[T] + \text{Var}(S/s)}{2(E[T] - E[S/s])},$$

which is essentially the G/G/1 result with a modified service time. This result, originally suggested by Kingman [54], has been studied by Suzuki and Yoshida [98]. A bound later suggested by Kingman [55] has the form

$$E[W] \leqslant \frac{s\text{Var}[T] + \text{Var}[S] + (1 - 1/s)\{E[S]\}^2}{2(sE[T] - E[S])}.$$

Bounds for some generalizations of the G/G/1 queue have been derived by Marshall [75]. Some of the cases discussed by him are queues with arrivals in batches of random size, queues with service in batches of fixed size, and queues with added delay for the first customer in a busy period. Marshall and Wolff [76] consider bounding the difference between the mean queue length found by an arriving customer and the arbitrary-time mean queue length in the G/G/1 system. It is also shown that, for G/G/1, the difference between the mean virtual wait and the mean actual wait does not exceed one half the mean interarrival time. Holtzman [44,45] derives an upper bound for mean waiting time in a Poisson input single server priority queue by considering waiting time as composed of four distinct parts and obtaining an upper bound for each of them.

Heathcote and Winer [41] take a somewhat different approach in deriving approximations for the moments of waiting times in the G/G/1 queue. Using an expansion related to the central limit theorem, they express $E[W_n] - E[W]$ as an infinite series. Now, knowing $E[W]$ one could estimate $E[W_n]$ by approximating the series. Other papers considering approximations and bounds for mean waiting time in G/G/1 and G/G/s queues or their special cases are Granot et al. [35] and Harrison [40].

Approximation techniques have been used for deriving information on other performance measures as well. Rider [88] approximates the emptiness probability to solve for the average queue size in a time-dependent M/M/1 queue. Natvig [80] approximates the transition probability $P_{10}(t)$ of the transition of the number in the system from 1 to 0 in time t, in a single-server Markovian queue with discouragement, by simplifying the expression derived through inversion. Benes [4] gives an approximation for p_n, the probability that an arriving customer finds n busy channels in a G/M/s/s system. Beneš [5], also provided an approximation for the covariance function of the number of busy channels in an M/M/s/s system. Another paper dealing with the approximations for covariance function of the number of busy channels in an M/M/s/s system is Descloux [23]. Approximations for Erlang's loss formula and its derivatives have been given by Jagerman [50] by truncation of a complex series. In these papers, related mostly to teletraffic theory, the technique used is analytical and manipulative. For other papers belonging to this class, readers are referred to Saaty [93], Cooper [20], Holtzman [46], and references cited by Holtzman.

Many of the exact queueing results are given as transform expressions that are difficult to invert. Numerical inversion of Laplace transforms is a convenient technique when such results are needed. Some of the initial papers on this technique are Gaver [31], Weeks [100], Dubner and Abate [24], Chiu, Chen, and Huang [18], and Stehfest [96]. Nance, Bhat, and Claybrook [79] have applied the different methods presented in the above papers to invert the transform of the busy-period distribution of an M/G/1/N type queue occurring in a time-sharing system. Abate, Dubner, and Weinberg [1] have applied the inversion method to the transform of the waiting-time distribution for a mass-storage device. It must be pointed out, though, that in the process of numerical inversion of transforms it is desirable to experiment with more than one technique, since their performance is highly dependent on the original function.

A recent inversion technique, given by Knepley and Fischer [57], makes the time parameter discrete and approximates a Laplace transform by an infinite series. Recursive relations then provide the needed numerical results. Al-Khayyal and Gross [2] approximate and bound the root of the functional equation associated with the GI/M/s queue to give bounds and approximations for steady-state measures of effectiveness and probabilities. Another approach based on transforms has been given by Kotiah [61] for Markovian systems. (These procedures are classified under numerical schemes, since the approximation is made on the results of analysis. Nevertheless, it is appropriate to mention that the outcome of the procedure is an approximation at the process level.)

Approximation results are also given in the form of limit and convergence theorems. A typical form of a limit theorem is to describe the behavior of a certain process as one of the system parameters approaches a specific limiting value. Convergence in queueing theory has received some attention in the last decade (see, for example, the survey paper by Iglehart [48]); however, not all such theorems are meant to be used as approximations. Köllerström [60] shows that the waiting time for the G/G/s system, under some general conditions, converges to a negative exponential as $\rho \to 1$, and then reformulates the result as an approximation with error bounds. Tomko [99], for $\rho < 1$, gives an approximation to the waiting time $W(N)$, in the queue M/M/m/N, in terms of the waiting time W for M/M/m/∞, and he provides the rate of convergence. For $\rho = 1$, $W(N)$ is shown to converge to a uniform distribution as the

capacity $N \to \infty$, and for $\rho > 1$, $W(N)$ is shown to converge to a normal distribution. The accuracy for the approximation of each $W(N)$ by its corresponding asymptotic distribution is also estimated. Kyprianou [65] shows that the virtual waiting time conditional on its still being in the first busy period in M/G/1 and GI/M/1 is asymptotically, as $\rho \to 1$, gamma distributed with two degrees of freedom and mean 4m, where

$$m = \frac{\text{Var}[T] + \text{Var}[S]}{2(1/E[T] - 1/E[S])}.$$

Schassberger [95] approximates the G/G/1 queue by a sequence of queues in which the interarrival and service times for the n^{th} system are Erlangian mixtures which are convex combinations of Erlangian distributions. He also shows that the distribution function of the virtual waiting time in the n^{th} system converges weakly to that of the original system. Kennedy [51] proves a similar but more general result for the single-server queue.

In a way, the numerical analysis of queueing systems carried out in a series of papers by Neuts and Neuts and Klimko [81] can also be identified as an approximating technique. It is a system type of approximation, in that discrete phase-type distributions are used for interarrival and service times. In the same spirit, one could include papers that have appeared on other numerical aspects of queueing systems, such as the solution of Chapman-Kolmogorov equations for birth and death processes. We shall not elaborate on these topics here, since the emphasis in this paper is more toward identifying different aspects of approximations.

VALIDATION OF APPROXIMATIONS

Validation is an integral part of an approximation procedure. It is needed to support the applicability of the technique and the reliability of results. An applied scientist has to constantly evaluate the trade-off between the ease of application of a particular technique and the accuracy of the ensuing results. Therefore, we expect the validation procedure to relate in some way and provide a comparison between approximate and exact results. Generally, validation of approximations can be achieved through error analysis, experimentation, and simulation. The relative merits of these procedures are discussed in the following paragraphs.

In error analysis, the deviation from an exact result is estimated as a function of the system parameters. For example, if we approximate by truncating a series, any bound on the remainder of the series will bound the error. One of the error-analysis procedures is to show that the error converges to zero as one or more of the parameters take a limiting value (see, for example, Natvig [80]). If the result is of a limiting nature, then the rate of convergence may help provide an error estimate (Köllerström [60], Tomko [99]). For two-sided-inequality results the error is bounded by the length of the interval; however, one needs to compare the bounds with some exact results as well (Bloomfield and Cox [12], Marshall [75]). Apart from inequalities, numerical point approximation is the only approach through which error estimates are obtained.

Experimentation is the most common validation technique for approximations. The essential feature of this procedure is to compare the approximated and the exact results for some special cases; if the comparison is favorable, similar performance is expected, in general. Absence of support for this basis requires careful and exhaustive experimentation covering a wider range of parameters. Clearly, this approach can be used for any type of approximation. For example, it is used in Beneš [4], Heathcote and Winer [41], Holtzman [44], and Rider [88] to validate numerical approximations. Avi-Itzhak and Heyman [3], Bhat and Fischer [8], Kuczura [64], Leibowitz [67], and Marchal and Harris [73] have used this method in the context of system approximations. Gaver and Shedler [33], Heyman [43], and Reiser and Kobayashi [87]

have used it to validate diffusion-approximation techniques. The main disadvantage in the procedure, though, is the lack of certainty that the conclusions drawn from experimentation can be extrapolated into more general settings.

Simulation of stochastic system has become popular, due to its wide applicability, closeness to reality, and the ability to use statistical analysis techniques. It is the last property that makes simulation a seemingly dependable and appealing validation technique. As can be seen from Fishman [30], considerable work has been done on the statistical aspects of simulations. But a word of caution is that the analysis is all too often messy and heavily dependent on factors such as sample size. The general approach is to generate samples of the studied process and define estimates for the required measures of performance. If the process is of the regenerative type, then the classical statistical techniques can be used to obtain confidence intervals and percentiles (see, for example, Fishman [29], Iglehart [49], and Lavenberg and Slutz [66]). Otherwise, one has to deal with the usual problems arising in simulation, such as dependent samples, effect of initial state, and transient behavior. In either case the simulation model needs validation, and this is usually done through experimentation (see, Rosenshine and Chandra [91]). The use of simulation as a validation technique is common under system approximations (Chandy, Herzog, and Woo [16,17], Halfin [39], Moore [78], and Sauer and Chandy [94]) and process approximations (Halachmi and Franta [38], Heyman [43], Kobayashi [58,59], and Reiser and Kobayashi [87]. For the validation of numerical approximations, even though error analysis is easier, simulation may be used (see Descloux [23]). However, most of these authors have satisfied themselves by the relative size of the percentage difference between the simulated and approximate results. Very few of them have resorted to a statistical analysis of simulated results and provide information such as confidence bounds on their estimates. When one uses simulation for the validation of approximations, it is desirable to state the accuracy of the simulated results as well.

As discussed above, validation takes different forms that vary in their usefulness. We consider error analysis as the most reliable procedure. However, it is difficult to implement under system-approximation and diffusion-approximation techniques. Inequalities may not need validation if they are tight enough. Nevertheless, it should be noted that inequalities that are tight are hard to compute, and those that are simple to compute are not tight. Thus, a sensitivity analysis of inequalities over the rest of the parameters may be recommended (Bhat [7]). Experimentation and simulation are the more-common forms of validation techniques, but, while we use them, their limitations should be clearly understood.

FUTURE PROSPECTS

Given above is a broad picture of approximation techniques used in queueing theory. Existing work in the queueing literature has been included in one or the other category of approximation, considering the main thrust of the paper. It must be noted, however, that many times a combination of more than one technique may be needed for a complete solution.

The emergence of approximate results is directly related to the applicability of systems. Furthermore, except for the well-known approximations for the mean waiting time in G/G/1 and related systems, most of the simple and applicable results occur predominantly in application areas of queueing theory, such as telephone traffic and computer systems. There is a significant factor in this phenomenon to be noted by a researcher. Since approximate results are obtained for direct use in real-world problems, they should be easily computable. Therefore, it does not make sense, except as an intellectual exercise and a theoretical piece of research, to provide a better approximation which is much harder to compute than an available simpler approximation. Thus, all applicable approximate results need to be examined from an effort-benefit view point.

As indicated earlier, validation of approximate results has attracted considerable attention, specifically in the application areas. Nevertheless, not enough attention seems to have been paid to the quality of the validation technique itself. In the case of experimentation, more sensitivity analysis is needed. Wider use of statistical techniques related to point and interval estimation should be made when simulation is preferred.

Queueing-theory researchers have been criticized for studying systems that are not relevant to the real world. However, it seems to us that this criticism is largely due to the complexity of available results in the literature rather than due to the systems themselves. If one looks at some of the applied areas of queueing, one finds more complex systems than those in the general operations-research and applied-probability literature. The distinction is in the nature of analysis. The results found in the applied-area literature are applicable, though approximate. A large percentage of results found in the general literature is less useful, though rigorous. Therefore, if we want to keep queueing theory as an integral part of operations research and as a problem-solving tool in the general area of applied probability and mathematics, approximation techniques should be put to increasing use whenever necessary. The trend during the past decade is in this direction, and there is every reason to believe that this trend is going to continue further in the coming years, bringing more reliability and applicability for the techniques used.

ACKNOWLEDGMENT

The authors are grateful to the referees for pointing out some of the missing references and making constructive comments that have improved the presentation.

REFERENCES

[1] Abate, J., H. Dubner, and S. B. Weinberg, "Queueing Analysis for the IBM 2314 Disk Storage Facility," Journal of the Association for Computing Machinery *15*, 577-589 (1968).

[2] Al-Khayyal, F. A., and D. Gross, "On Approximating and Bounding GI/M/c Queues," in *TIMS Studies in Management Sciences 7*, pp. 233-245 (North Holland, Amsterdam, 1977).

[3] Avi-Itzhak, B., and D. P. Heyman, "Approximate Queueing Model for Multiprogramming Computer Systems," Operations Research *21*, 1212-1230 (1973).

[4] Beneš, V. E., "On Trunks with Negative Exponential Holding Times Serving a Renewal Process," Bell System Technical Journal *38*, 211-258 (1959).

[5] Beneš, V. E., "The Covariance Function of a Simple Trunk Group, With Applications to Traffic Measurements," Bell System Technical Journal *40*, 117-148 (1961).

[6] Bergman, R., and D. Stoyan, "On Exponential Bounds for the Waiting Time Distribution Function in GI/G/1," Journal of Applied Probability *13*, 411-417 (1976).

[7] Bhat, U. N., "Sensitivity Analysis of Performance Measures in Some Queueing Systems," Tech. Comment No. 30-74, Defense Communication Center, Reston, Virginia (Dec. 1974).

[8] Bhat, U. N., and M. J. Fischer, "Multichannel Queueing Systems with Heterogeneous Classes of Arrivals," Naval Research Logistics Quarterly *23*, 271-282 (1976).

[9] Bhat, U. N., and N. U. Prabhu, "A Probabilistic Model for Queues at Traffic Signals," Tech. Rep. IEOR 75007, Department of IE/OR, Southern Methodist Univeristy, Dallas (1975).

[10] Bhat, U. N., and Raju, S., "An Approximate Analysis of a Multi-server Finite Queue with Heterogeneous Customers," Tech. Rep. IEOR 77020, Department of IE/OR, Southern Methodist University, Dallas (1977).

[11] Bhat, U. N., A. C. Wheeler, and M. J. Fischer, "On the Existence of Limiting Distributions in Stochastic Systems with Secondary Inputs," Opsearch *11*, 81-89 (1974).

[12] Bloomfield, P., and D. R. Cox, "A Low Traffic Approximation for Queues," Journal of Applied Probability *9*, 832-840 (1972).

[13] Brosh, I., "Preemptive Priority Assignment in Multichannel Systems," Operations Research *17*, 526-535 (1969).

[14] Brumelle, S. L., "Some Inequalities for Parallel-Server Queues," Operations Research *19*, 402-413 (1971).

[15] Brumelle, S. L., "Bounds on the Wait in a GI/M/k Queue," Management Science *19*, 773-777 (1973).

[16] Chandy, K. M., N. Herzog, and L. Woo, "Approximate Analysis of General Queueing Networks," IBM Journal of Research and Development *19*, 43-49 (1975).

[17] Chandy, K. M., N. Herzog and L. Woo, "Parametric Analysis of Queueing Networks," IBM Journal of Research and Development *19*, 36-42 (1975).

[18] Chiu, R. F., C. F. Chen, and C. J. Huang, "A New Method for the Inverse Laplace Transformation Via the Fast Fourier Transform," *Twenty-Second Southwestern IEEE Conference Record,* Dallas, Texas, pp. 201-203, (1970).

[19] Conolly, B., "Generalized State Dependent Erlangian Queues: Speculations About Calculating Measures of Effectiveness," Journal of Applied Probability *12*, 358-363 (1975).

[20] Cooper, R. B., *Introduction to Queueing Theory* (MacMillan, New York, 1972).

[21] Cosmetatos, G. P., "Approximate Equilibrium Results for the Multi-Server Queue (GI/M/n)," Operations Research Quarterly *25*, 625-634 (1974).

[22] Cox, D. R., and H. D. Miller, *The Theory of Stochastic Processes* Wiley, New York, 1968).

[23] Descloux, A., "On the Accuracy of Loss Estimates," *Bell System Technical Journal 44*, 1139-1164 (1965).

[24] Dubner, H., and J. Abate, "Numerical Inversion of the Laplace Transforms by Relating Them to the Finite Fourier Cosine Transform," Journal of the Association for Computing Machinery *15*, 115-123 (1968).

[25] Feller, W., *An Introduction to Probability Theory and its Applications*, Vol. II (Wiley, New York, 1966).

[26] Fischer, M. J., "On Two Problems in Alternating Queues," Ph.D. Dissertation, School of Engineering and Applied Science, Southern Methodist University (1971).

[27] Fischer, M. J., "An Approximation to Queueing Systems with Interruptions," Defense Communications Engineering Center, Reston, Virginia (1976).

[28] Fischer, M. J., "Analysis and Design of Loop Service Systems a Diffusion Approximation," Operations Research *25*, 269-278 (1977).

[29] Fishman, G. S., "Estimation in Multiserver Queueing Simulations," Operations Research *22*, 72-78 (1974).

[30] Fishman, G. S., *Concept and Methods in Discrete Event Digital Simulation*, (Wiley, New York, 1973).

[31] Gaver, D. P., Jr., "Observing Stochastic Processes and Approximate Transform Inversion," Operations Research *14*, 444-459 (1966).

[32] Gaver, D. P., Jr., "Diffusion Approximations and Models for Certain Congestion Problems," Journal of Applied Probability *5*, 607-623 (1968).

[33] Gaver, D. P., Jr., and G. S. Shedler, "Processor Utilization in Multiprogramming Systems Via Diffusion Approximations," Operations Research *21*, 569-576 (1973).

[34] Ghosal, A., *Some Aspects of Queueing and Storage Systems*, Vol. 23 of Lecture Notes in Operations Research and Mathematical Systems (Springer-Verlag, New York, 1970).

[35] Granot, D., F. Granot, and A. Lemoine, "Approximations for Service Systems with Nonindependent Interarrival Times," Operations Research *23*, 162-166 (1975).

[36] Gross, D., "Sensitivity of Queueing Models to the Assumption of Exponentiality," Naval Research Logistics Quarterly *22*, 271-287 (1975).

[37] Gross, D., and C. M. Harris, *Fundamentals of Queueing Theory*, (Wiley, New York, 1974).

[38] Halachmi, B., and W. R. Franta, "A Diffusion Approximate Solution to the G/K/k Queueing System," Presented at ORSA/TIMS Meeting (Nov. 18, 1975).

[39] Halfin, S., "An Approximate Method for Calculating Delays for a Family of Cyclic-Type Queues," Bell System Technical Journal *54*, 1733-1754 (1975).

[40] Harrison, J. M., "The Heavy Traffic Approximation for Single Server Queues in Series," *Journal of Applied Probability 10*, 613-639 (1973).

[41] Heathcote, C. R., and P. Winer, "An Approximation for the Moments of Waiting Time," Operations Research *17*, 175-186 (1969).

[42] Heyman, D. P., "An Approximation for the Busy Period of the M/G/1 Queue Using a Diffusion Model," *Journal of Applied Probability 11*, 159-169 (1974).

[43] Heyman, D. P., "A Diffusion Model Approximation for the GI/G/1 Queue in Heavy Traffic," Bell System Technical Journal *54*, 1637-1646 (1975).

[44] Holtzman, J. M., "Bounds for A Dynamic-Priority Queue," Operations Research *10*, 461-468 (1971).

[45] Holtzman, J. M., "Analysis of Dependence Effects in Telephone Trunking Networks," Bell System Technical Journal *50*, 2647-2662 (1971).

[46] Holtzman, J. M., "Some Compatible Approximations in TeleTraffic Theory," Paper presented at ORSA/TIMS Meeting (Nov. 17, 1975).

[47] Iglehart, D., "Limiting Diffusion Approximations for the Many Server Queue and the Repairman Problem," Journal of Applied Probability *2*, 429-441,1965.

[48] Iglehart, D., "Weak Convergence in Queueing Theory," Advances in Applied Probability *5*, 570-594 (1973).

[49] Iglehart, D., "Simulating Stable Stochastic Systems, V: Comparison of Ratio Estimators," Naval Research Logistics Quarterly, *22*, 553-565 (1975).

[50] Jagerman, D. L., "Some Properties of the Erlang Loss Function," *Bell System Technical Journal 53*, 525-551 (1974).

[51] Kennedy, D. P., "The Continuity of the Single Server Queue," Journal of Applied Probability *9*, 370-381 (1972).

[52] Kimura, M., "Diffusion Models in Population Genetics," Journal of Applied Probability *1*, 177-232 (1964).

[53] Kingman, J. F. C., "Some Inequalities for the Queue GI/G/1," Biometrika *49*, 315-324 (1962).

[54] Kingman, J. F. C., "The Heavy Traffic Approximation in the Theory of Queues," in *Proceedings of a Symposium on Congestion Theory*, W. L. Smith and W. E. Wilkinson, eds. (University of North Carolina Press, Chapel Hill, 1965).

[55] Kingman, J. F. C., "Inequalities in the Theory of Queues," Journal of the Royal Statistical Society *B 32*, 102-110 (1970).

[56] Kleinrock, L. *Queueing Systems Vol. II: Computer Applications*, Chapter 2 (Wiley, New York, 1976).

[57] Knepley, J. E., and M. J. Fischer, "A Numerical Solution for Some Computational Problems Occurring in Queueing Theory, " in *TIMS Studies in Management Sciences 7*, pp. 271-285 (North Holland, Amsterdam, 1977).

[58] Kobayashi, H., "Application of the Diffusion Approximation to Queueing Networks I: Equilibrium Queue Distributions," Journal of the Association for Computing Machinery *21*, 316-328 (1974).

[59] Kobayashi, H., "Application of the Diffusion Approximation to Queueing Networks II: Nonequilibrium Distributions and Applications to Computer Modeling," Journal of the Association for Computing Machinery *21*, 459-469 (1974).

[60] Kollerstrom, J., "Heavy Traffic Theory for Queues with Several Servers I," Journal of Applied Probability, *11*, 544-552 (1974).

[61] Kotiah, T. C. T., "Approximations for the Transient Behavior of Some Queues," No. 40, SIU Papers in Mathematics and Mathematical Science, Southern Illinois Univeristy, Edwardsville, Illinois (1976).

[62] Kotiah, T. C. T., "On a Linear Programming Technique for the Steady State Behavior of Some Queueing Systems," Operations Research *25*, 289-303 (1977).

[63] Kotiah, T. C. T., J. W. Thompson, and W. A O'N. Waugh, "Use of Erlangian Distributions for Single-Server Queueing Systems," Journal of Applied Probability *6*, 584-593 (1969).

[64] Kuczura, A., "The Interrupted Poisson Process as an Overflow Process," Bell System Technical Journal *52*, 437-448 (1973),

[65] Kyprianou, E. K., "The Quasi-Stationary Distributions of Queues in Heavy Traffic," Journal of Applied Probability *9*, 821-831 (1972).

[66] Lavenberg, S. S., and D. R. Slutz, "Introduction to Regenerative Simulation," IBM Journal of Research and Development *19*, 458-462 (1975).

[67] Liebowitz, M. A., "An Approximate Method for Treating a Class of Multiqueue Problems," IBM Journal of Research and Development *5*, 204-209 (1961).

[68] Luchak, G., "The Solution of the Single Channel Queueing Equation Characterized by a Time Dependent Poisson Distributed Arrival Rate and a General Class of Holding Times," Operations Research *4*, 711-732 (1956).

[69] Maaloe, E., "Approximation Formulae for Estimation of Waiting-Time in Multiple-Channel Queueing System," Management Science *19*, Applied Series, 703-710 (1973).

[70] Marchal, W. G., "Some Simple Bounds and Approximations in Queueing," Tech. Mem. T-294, Institute of Management Science and Engineering, The George Washington University (Jan. 1974).

[71] Marchal, W. G., "An Approximate Formula for Waiting Time in Single Server Queues," AIEE Transactions *8*, 473-474 (1976).

[72] Marchal, W. G., "Some Simpler Bounds on Mean Queueing Times," private communication (1976).

[73] Marchal, W. G., and C. M. Harris, "A Modified Erlang Approach to Approximating GI/G/1 Queues," Journal of Applied Probability *13*, 118-126 (1976).

[74] Marshall, K. T., "Some Inequalities in Queueing," Operations Research *16*, 651-665 (1968).

[75] Marshall, K. T., "Bounds for Some Generalizations of the GI/G/1 Queue," Operations Research *16*, 841-848 (1968).

[76] Marshall, K. T., and R. W. Wolff, "Customer Average and Time Average Queue Lengths and Waiting Times," Journal of Applied Probability *8*, 535-542 (1971).

[77] McNeil, D. R., "Diffusion Limits for Congestion Models," Journal of Applied Probability *10*, 368-376 (1973).

[78] Moore, S. C. "Approximating the Behavior of Nonstationary Single-Server Queues," Operations Research *23*, 1011-1032 (1975).

[79] Nance, R., U. N. Bhat, and B. G. Claybrook, "Busy Period Analysis of a Time Sharing System: Transform Inversion," Journal of the Association for Computing Machinery *19*, 453-463 (1972).

[80] Natvig, D., "On the Transient State Probabilities for a Queueing Model Where Potential Customers are Discouraged by Queue Length," Journal of Applied Probability *11*, 345-354 (1974).

[81] Neuts, M. F., "The Single Server Queue in Discrete Time — Numerical Analysis," Naval Research Logistics Quarterly, Part I: *20*, 297-302 (1973); Part II (with E. M. Klimko): *20*, 305-320 (1973); Part III: (with E. M. Klimko): *20*, 557-568 (1973).

[82] Newell, G. F., "Approximation Methods for Queues with Application to the Fixed Cycle Traffic Light," SIAM Review *7*, 223-239 (1965).

[83] Newell, G. F., "Queues with Time-Dependent Arrrival Rates I-III," Journal of Applied Probability *5*, 436-451, 579-606 (1968).

[84] Newell, G. F., *Applications of Queueing Theory*, Chapter 6 (Chapman and Hall, London, 1971).

[85] Newell, G. F., "Approximate Behavior of Tandem Queues," ITTE Special Report, University of California-Berkeley (1975).

[86] Nozaki, S. A., and S. M. Ross, "Approximations in Multi-Server Poisson Queues," ORC 76-10, University of California-Berkeley (April 1976).

[87] Reiser, M., and H. Kobayashi, "Accuracy of the Diffusion Approximation for some Queueing Systems," IBM Journal of Research and Development *18*, 110-124 (1974).

[88] Rider, K. L., "A Simple Approximation to the Average Queue Size in the Time-Dependent M/M/1 Queue," Journal of the Association for Computing Machinery *23*, 361-367 (1976).

[89] Rolski, T., and D. Stoyan, "On the Comparison of Waiting Times in GI/G/1 Queues," Operations Research *24*, 197-200 (1976).

[90] Rosenshine, M., "Approximating the M/E(k,n)/1 Queue by a Birth-death Process," paper presented at ORSA/TIMS Meeting, Miami (November 1976).

[91] Rosenshine, M., and M. J. Chandra, "Approximate Solutions for Some Two State Tandem Queues, Part 1: Individual Arrivals at the Second State," Operations Research *23*, 1155-1166 (1975).

[92] Ross, S. M., "Bounds on the Delay Distribution in GI/G/1 Queues," Journal of Applied Probability *11*, 417-421 (1974).

[93] Saaty, T. L., *Elements of Queueing Theory*, p. 159 and p. 162 (McGraw Hill, New York, 1961).

[94] Sauer, C. H. and K. M. Chandy, "Approximate Analysis of Central Server Models," IBM Journal of Research and Development *19*, 301-313 (1975).

[95] Schassberger, R., "On the Waiting Time in the Queueing System GI/G/1," Annals of Mathematical Statistics *41*, 182-187 (1970).

[96] Stehfest, H., "Algorithm 368: Numerical Inversion of Laplace Transforms," Communications of the ACM *13*, 47-49 (Jan. 1970).

[97] Stoyan, D., "Bounds and Approximations in Queueing Through Monotonicity and Continuity," Operations Research *25*, 851-863 (1977).

[98] Suzuki, T., and Y. Yoshida, "Inequalities for Many-Server Queue and Other Queues," Journal of the Operations and Research Society of Japan *13*, 59-77 (1970).

[99] Tomko, J., "The Rate of Convergence in Limit Theorems for Service Systems with Finite Queue Facility," Journal of Applied Probability *9*, 87-102 (1972).

[100] Weeks, W. T., "Numerical Inversion of Laplace Transforms Using Laguerre Functions," Journal of the Association for Computing Machinery *13*, 419-429 (1966).

[101] Wilkinson, R. I., "Theories for Toll Traffic Engineering in the U.S.A.," Bell System Technical Journal *35*, 421-514 (1956).

[102] Willemain, T. R., "Approximate Analysis of a Hierarchical Queueing Network," Operations Research *22*, 522-544 (1974).

[103] Wishart, D. M. G., "A Queueing System with Service-Time Distribution of Mixed Chi-Squared Type," Operations Research *8*, 174-179 (1959).

[104] Yu, O. S., "Stochastic Bounds for Heterogeneous-Server Queues with Erlang Service Times," Journal of Applied Probability *11*, 785-796 (1974).

DISTRIBUTION OF SAMPLE CORRELATION COEFFICIENTS*

Khursheed Alam

Clemson University
Clemson, South Carolina

ABSTRACT

Let $(Y, X_1, \ldots, X_K)$ be a random vector distributed according to a multivariate normal distribution, where $X_1, \ldots, X_K$ are considered as predictor variables and y is the predictand. Let r_i and R_i denote the population and sample correlation coefficients, respectively, between Y and X_i. The population correlation coefficient r_i is a measure of the predictive power of X_i. The author has derived the joint distribution of $R_1, \ldots, R_K$ and its asymptotic property. The given result is useful in the problem of selecting the most important predictor variable corresponding to the largest absolute value of r_i.

1. INTRODUCTION

The problem of selecting a variable or several variables from a set of predictor variables $\{X_i\}$ occurs frequently in the design of experiments. The correlation between a predictor variable X_i and the predictand Y measures the "leverage" of X_i upon Y. If X_i and Y are jointly distributed according to the standard bivariate normal distribution, with correlation coefficient r_i, then the conditional distribution of Y given X_i is normal $N(r_i X_i, 1 - r_i^2)$. The larger the absolute value of r_i, the smaller is the variance of the conditional distribution, and therefore the higher is the predictive power of X_i. Thus, the predictor variable corresponding to the largest value of r_i^2 may be considered as the most important (best) predictor variable.

The problem of selecting one or more of the predictor variables which have larger correlations with the predictand than the rest of the variables arises in the test of accuracy of a weapon system. The accuracy may be described by the radial distance between the target and the point of impact of a projectile released by the weapon. There are a number of contributory factors in missing the target. Let $X_1, X_2, \ldots$ denote the variables measuring the effects of the contributory factors. These variables are positively correlated with Y and among themselves. If the correlation between Y and X_1, say, is much larger than the correlations between Y and X_2, etc., then the factor associated with X_1 may be considered as a major contributory factor in missing the target which should be looked into for better control of the projectile. Generally, it is expensive to measure the variables, since the measurements invole the destruction of the projectile. Therefore, it is desirable to estimate the correlations between the variables from a sample of observations. The results of this paper, which deals with the joint distribution of the sample correlation coefficients, would be useful in screening a set of predictor variables for the selection of one or more of them on the basis of their correlations with the predictand.

*The author's work was supported by the Office of Naval Research under Contract N00014-75-0451.

The problem of selecting the predictor variable associated with the largest correlation with the predictand has been considered recently by Ramberg [3]. Rizvi and Solomon [4] and Alam, Rizvi, and Solomon [1] have also considered the problem of selecting, from $p \geqslant 2$ multivariate normal populations, the population with the largest multiple correlation between a single variate, classified as the predictand, and the remaining variates.

Let the random vector $(Y, X_1, \ldots, X_K)$ be distributed according to a multivariate normal distribution. Suppose that a sample of n observations is taken from the given distribution. Let r_i and R_i denote the population and sample correlation coefficients between Y and X_i, respectively. Let $r_i^* = r_i(1 - r_i^2)^{-1/2}$ and $R_i^* = R_i(1 - R_i^2)^{-1/2}$. In the following section we derive the distribution of $\mathbf{R}^* = (R_1^*, \ldots, R_K^*)$. The distribution of $\mathbf{R}^*$ is expressed in terms of the joint distribution of three independent random variables. The asymptotic distribution of $\mathbf{R}^*$ for large n is also given.

2. DISTRIBUTION of $\mathbf{R}^*$

Let $\overset{d}{\approx}$ mean "distributed as". Without loss of generality we can assume that the variables $Y, X_1, \ldots, X_K$ are standardized, that is, they are distributed with mean 0 and variance 1. Let r_{ij} denote the correlation coefficient between X_i and X_j, and let $\Sigma = (r_{ij})$ and $\hat{\Sigma} = (r_{ij}^2)$. Let $Y_t, X_{1t}, \ldots X_{Kt})$ denote the t-th observation in the sample, and let

$$\bar{X}_i = \frac{1}{n}\sum_{t=1}^{n} X_{it}, \quad \bar{Y} = \frac{1}{n}\sum_{t=1}^{n} Y_t, \quad S^2 = \sum_{t=1}^{n} (y_t - \bar{Y})^2.$$

Then

$$V_i = \left[\sum_{t=1}^{n} (Y_t - \bar{Y}) X_{it}\right] / S \tag{2.1}$$

and

$$W_i = \sum_{t=1}^{n} (X_{it} - X_i)^2 - V_i^2. \tag{2.2}$$

From the theory of linear regression analysis it is seen that $W_i \overset{d}{\approx} (1 - r_i^2)\, \chi^2_{n-2}$, chi-square with $n-2$ degrees of freedom, independent of V_i and $\mathbf{Y} = (Y_1, \ldots, Y_n)'$, $V_i \approx N(r_i S, 1 - r_i^2)$, and cov $(V_i, V_j) = r_{ij} - r_i r_j$, conditionally given $\mathbf{Y}$. Let $\lambda_{ij} = r_{ij} - r_i r_j$ and $\Omega = (\lambda_{ij})$. It is also seen that W_i can be represented as the sum of squares of $(n-2)$ orthogonal linear functions of $X_{i1}, \ldots, X_{in}$. That is,

$$W_i \overset{d}{\approx} \sum_{t=1}^{n-2} Z_{it}^2 \tag{2.3}$$

where $Z_t = (Z_{1t}, \ldots, Z_{kt})'$ are identically and independently distributed as $N(\mathbf{O}, \Omega)$, independent of $V_1, \ldots, V_K$ and S.

Let $\mathbf{T} = (T_1, \ldots, T_K)'$ be a random vector distributed as $N(\mathbf{O}, \Omega)$, independent of S and $\mathbf{W} = (W_1, \ldots, W_K)'$. Then

$$\begin{aligned} R_i^* &= V_i\, (W_i)^{-1/2} \\ &\overset{d}{\approx} (T_i + r_i S)\, W_i^{-1/2}. \end{aligned} \tag{2.4}$$

Therefore,

THEOREM 2.1: The joint distribution of the sample correlation coefficients between the predictand and the predictor variables of a multivariate normal distribution is given by (2.4), where $\mathbf{T} \overset{d}{\approx} N(O, \Omega)$, $S^2 \overset{d}{\approx} \chi^2_{n-1}$, and the distribution of $\mathbf{W}$ is given by (2.3). Moreover, S, $\mathbf{T}$ and $\mathbf{W}$ are jointly independent.

For large n, $\mathbf{W}$ is asymptotically distributed as $N[(n-2)\mathbf{f}, 2(n-2)\hat{\Omega}]$, by Theorem 4.2.4 of Anderson [2], where $\mathbf{f} = (1 - r_1^2, \ldots, 1 - r_K^2)'$ and $\hat{\Omega} = (\lambda_{ij}^2)$. Therefore,

COROLLARY 2.1: The asymptotic distribution of $\mathbf{R}^*$ is given by (2.4), where $\mathbf{T} \overset{d}{\approx} N(O, \Omega)$, $S^2 \overset{d}{\approx} \chi^2_{n-1}$, $W \overset{d}{\approx} N[(n-2)f, 2(n-2)\hat{\Omega}]$, and S, $\mathbf{T}$, and $\mathbf{W}$ are jointly independent.

The following corollary gives the asymptotic distribution of $\sqrt{n}\ (\mathbf{R}^* - \mathbf{r}^*)$, which is derived from (2.4) but follows also from the central limit theorem or Theorem 4.2.4 of Anderson.

Let $\gamma_{ij} = (r_{ij} - r_i r_j)\ (1 - r_i^2)^{-1/2}\ (1 - r_j^2)^{-1/2}$ and $\Gamma = (\gamma_{ij}^2)$. From (2.4) we have for large n

$$\sqrt{n}\ (R_i^* - r_i^*) = T_i(1 - r_i^2)^{-1/2} + \sqrt{n}\ r_i^* \left\{ \left[\frac{(1 - r_i^2)S^2}{W_i} \right]^{1/2} - 1 \right\} + O_p(n^{-1/2}) \tag{2.5}$$

$$= T_i(1 - r_i^2)^{-1/2} + \frac{r_i^*}{\sqrt{2}}\ (A - B_i) + O_p(n^{-1/2}),$$

where $A \overset{d}{\approx} N(O, 1)$, $\mathbf{B} = (B_1, \ldots, B_K)' \overset{d}{\approx} N(O, \Gamma)$. Moreover $\mathbf{T}$, A and B are jointly independent. Therefore, $\sqrt{n}\ (\mathbf{R}^* - \mathbf{r}^*)$ is asymptotically distributed as $N(\mathbf{O}, C)$, where

$$C_{ii} = 1 + r_i^{*2} \tag{2.6}$$

$$C_{ij} = \gamma_{ij} + \frac{1}{2}\ r_i^*\ r_j^*\ (1 + \gamma_{ij}^2).$$

Therefore

COROLLARY 2.2: For large n, $\sqrt{n}\ (\mathbf{R}^* - \mathbf{r}^*)$ is asymptotically distributed as $N(\mathbf{O}, C)$, where C is given by (2.6).

It is interesting to consider the following special cases: (1) $r_i = 0$, $r_{ij} = 0$, $i \neq j$ for all i and j; that is, the variables $Y, X_1, \ldots, X_K$ are jointly independent. We have $C = I$ and $\sqrt{n}\ (\mathbf{R}^* - \mathbf{r}^*) \overset{d}{\approx} N(O, I)$, asymptotically. (2) $r_i = 0$, $r_{ij} = \rho$, $i \neq j$ for all i and j; that is, the predictor variables $X_1, \ldots, X_K$ are equi-correlated and independent of Y. We have $C_{ii} = 1$, $C_{ij} = \rho$, $i \neq j$. (3) $r_i = \rho$, $r_{ij} = 0$, $i \neq j$ for all i and j; that is, the predictor variables are jointly independent and equi-correlated with Y. We have

$$C_{ii} = (1 - \rho^2)^{-1}$$

and

$$C_{ij} = \frac{\rho^2\ (2\rho^2 - 1)}{2(1 - \rho^2)^3}.$$

Consider the problem of selecting the best predictor variable. A standard procedure is to select the variable from the predictor variables corresponding to the largest value of the squared

correlation coefficients $R_1^2, \ldots, R_K^2$ or, equivalently, $R_1^{*2}, \ldots, R_K^{*2}$. By Corollary 2.2 the probability of a correct selection can be derived for large n from the multivariate normal distribution function. In special case (1) we have that $n \max (R_1^{*2}, \ldots, R_K^{*2})$ is distributed as the largest order statistic in a sample of K observation from χ_1^2 — chi-square with 1 degree of freedom. This result can be used also to test the significance of the correlation between the selected predictor variable and the predictand. Similar results are obtained for cases (2) and (3).

Let $\mathbf{r}^\circ$ be a given value of $\mathbf{r}^*$. From Corollary 2.2 we have that $n(\mathbf{R}^* - \mathbf{r}^\circ)' C^{-1} (\mathbf{R}^* - \mathbf{r}^\circ)$ is asymptotically distributed as $\chi^2_{K,\delta'}$, noncentral chi-square with K degrees of freedom and noncentrality parameter $\delta = n(\mathbf{r}^* - \mathbf{r}^\circ)' C^{-1}(\mathbf{r}^* - \mathbf{r}^\circ)$. Let $\hat{C} = C(\mathbf{R}^*, (R_{ij}))$ denote the estimate of C where R_{ij} are the sample correlation coefficients between X_i and X_j. $\hat{C}$ converges in probability to C as $n \to \infty$. Therefore, the statistic $\hat{\delta} = n(\mathbf{R}^* - \mathbf{r}^\circ)' \hat{C}^{-1}(\mathbf{r}^* - \mathbf{r}^\circ)$ can be used to test the hypothesis that $\mathbf{r}^* = \mathbf{r}^\circ$.

REFERENCES

[1] Alam, K., M. H. Rizvi, and H. Solomon, "Selection of Largerst Multiple Correlation Coefficients: Exact Sample Size Case," *Annals of Statistics 4*, 614-620 (1976).

[2] Anderson, T. W., *An Introduction to Multivariate Analysis*, (Wiley, New York, 1958).

[3] Ramberg, J. S., "Selecting the Best Predictor Variate," *Communications in Statistics A6*, 1133-1147 (1977).

[4] Rizvi, M. H., and H. Solomon, "Selection of Largest Multiple Correlation Coefficients: Asymptotic Case," Journal of the American Statistical Association, *68*, 184-188 (1973).

OPTIMAL PROJECT COMPRESSION WITH DUE-DATED EVENTS*

S. E. Elmaghraby and P. S. Pulat

North Carolina State University at Raleigh
Raleigh, North Carolina

ABSTRACT

The paper proposes an algorithm for the determination of the solution of the activities to be shortened and the amount by which they are to be shortened in order to minimize the total cost of project completion. This cost involves a linear penalty for tardienss of a set of key events and a linear cost of activity compression from its normal duration. The procedure is a generalization of the work of Fulkerson.

INTRODUCTION

This paper deals with the problem of optimal project "compression"—or early finish—assuming linear costs of shortening individual activities as well as linear penalties for tardiness of a subset of events. It proposes a model and an algorithm for its solution.

Assumed given is the project network $G = (N,A)$, where N is the set of nodes (or events), $N = \{1, 2, \ldots, n\}$, and A is the set of arrows (or activities). The network is acyclic, and we assume that each arrow leads from a small-numbered node to a higher-numbered one. For more background on such network representation of projects and the relevant terminology, see Ref. [2], Chapters 1 and 2. An activity may be designated either by its end nodes i and j or by its generic designation $e \in A$. Its duration is denoted by y_e, where $0 \leqslant l_e \leqslant y_e \leqslant u_e < +\infty$. It is assumed that u_e represents the "normal" duration of the activity, that is, its lowest-cost duration before any shortening is undertaken. As y_e is shortened away from u_e a cost is accumulated at a rate $a_e > 0$; that is, the cost of activity e when it is accomplished in duration y_e is given by

$$c_e = b_e - a_e y_e, \quad \text{for } l_e \leqslant y_e \leqslant u_e \text{ and } a_e, b_e, c_e > 0,$$

where b_e is the intercept of the line c_e with the cost axis. The impetus to shorten any activity (and thus incur additional costs) stems form the fact that a subset $K (\subseteq N)$ of the nodes, the so-called *key events*, have specified due dates $\{d_k; k \in K\}$ and penalties $p_k > 0$ incurred per unit time of tardiness. We presume that the last node n carries a due date; otherwise we ignore the subnetwork after the largest due-dated node. Let t_j denote the time of realization of node $j \in N$. Evidently, $\{t_j\}$ are dependent on the activity durations. If we put $y_e = u_e$ for all $e \in A$, and it turns out that all realization times of all key events are no larger than their respective due dates, then that must be the cheapest possible realization of the project.

*This research was partially supported by ARO Contract DAAG29-76-G-0204.

The problem of project compression arises when some, or all, of the key events are tardy. Then it is desired to determine the subsets of activities whose duration are to be shortened, and the amount of that shortening, in order to incur the smallest total cost (= cost of shortening plus cost of tardiness)..

Mathematically, the problem may be stated as

$$\text{(1.1)} \qquad \text{Minimize } z = \sum_{(ij)} c_{ij} + \sum_{k \in K} p_k v_k - \sum_{(ij)} (b_{ij} - a_{ij} y_{ij}) + \sum_{k \in K} p_k v_k$$

subject to

		Dual Variables
(1.2)	$t_i - t_j + y_{ij} \leqslant 0$, all $(ij) \in A$;	f_{ij}
(1.3)	$-t_1 + t_k - v_k \leqslant d_k$, all $k \in K$;	λ_k
(1.4)	$y_{ij} \leqslant u_{ij}$, all $(ij) \in A$;	g_{ij}
(1.5)	$-y_{ij} \leqslant -l_{ij}$, all $(ij) \in A$;	h_{ij}

The first set of constraints of (1.2) consists of the standard "earliest realization time" constraints, which express the condition that node j cannot be realized except after all (immediately) preceding nodes have been realized and all connecting activities have been completed. The second set of contraints of (1.3) is derived from the fact that $(t_k - t_1) - d_k$ may be a positive number (betraying a tardy event) or a negative number (betraying an early event), hence it can be written as $t_k - t_1 - d_k = v_k - w_k$, where v_k and w_k are $\geqslant 0$. Rearranging the variables and dropping the slack variable w_k, we get (1.3). Note that the variable $v_k \geqslant 0$ measures the tardiness of node $k \in K$, hence it is "costed" at the rate p_k in (1.1).

The problem specified by (1) is an LP with $3A + K$ constraints in $A + N$ original variables. (Note that we are using the same symbol to denote the set and its rank.) A frontal attack on this LP using regular simplex iterations would miss capitalizing on its special structure. Such capitalization is the subject of this paper.

For the moment it is appropriate to remark that a simpler version of the LP of (1), in which no due dates were specified (whence the constraints (1.3) were absent), was treated by Fulkerson [3]. Our approach relies heavily on that development, and may be viewed as a generalization of it. See also Ref. [5], pp. 165-169, for an out-of-kilter algorithm to the solution of that simpler version.

THEORY OF APPROACH

We rewrite the objective of (1.1) as

$$\text{(2)} \qquad \text{Maximize } \sum_{(ij)} a_{ij} y_{ij} - \sum_{k \in K} p_k v_k,$$

in which we ignore the constant term $\sum_{(ij)} b_{ij}$ and reverse signs. Let f_{ij}, λ_k, g_{ij}, and h_{ij} be the dual variables corresponding to constraints (1.2) to (1.5). Utilizing the objective in (2), the dual LP may be stated as

$$\text{(3.1)} \qquad \text{Minimize } \sum_{(ij)} u_{ij} g_{ij} - \sum_{(ij)} l_{ij} h_{ij} + \sum_{k \in K} d_k \lambda_k$$

subject to

$$\sum_{j \in A(1)} f_{1j} - \sum_{k \in K} \lambda_k, \tag{3.2}$$

where $A(1)$ is the set of nodes immediately following node 1 and connected to it,

$$-\sum_{j \in A(i)} f_{ij} + \sum_{j \in B(i)} f_{ji} = \begin{cases} \lambda_k & \text{if } i = k, \\ 0 & \text{if } i \notin K,\ i \neq 1, \end{cases} \tag{3.3}$$

where $B(i)$ is the set of nodes immediately preceding node i and connected to it,

$$f_{ij} + g_{ij} - h_{ij} = a_{ij}, \quad \text{all } (ij) \in A, \tag{3.4}$$

$$\lambda_k \leqslant p_k, \quad \text{all } k \in K, \tag{3.5}$$

$$f_{ij},\ g_{ij},\ h_{ij} \geqslant 0 \text{ for all } (ij) \in A, \text{ and } \lambda_k \geqslant 0 \text{ for all } k \in K. \tag{3.6}$$

Equality appears in (3.2) to (3.4) because the (primal) t_i variables were not constrained in sign. Remarking that g_{ij} and h_{ij} cannot simultaneously be positive in an optimal solution, since u_{ij} is assumed larger than l_{ij} (the trivial case in which $l_{ij} = u_{ij}$ fixes the magnitude of y_{ij} at their common value, and it is no longer a variable), we express these two variables in terms of f_{ij} as follows

$$g_{ij} = \max(0,\ a_{ij} - f_{ij}),$$

$$h_{ij} = \max(0,\ f_{ij} - a_{ij}).$$

To eliminate the "max" operator in these two expressions, divide the total "flow" f_{ij} into two parts, f_{ij}^1 and f_{ij}^2, such that $f_{ij} = f_{ij}^1 + f_{ij}^2$, $0 \leqslant f_{ij}^1 \leqslant a_{ij}$, and $f_{ij}^2 = f_{ij} - a_{ij} \geqslant 0$. Whence $g_{ij} = a_{ij} - f_{ij}^1 \geqslant 0$ and $h_{ij} = f_{ij}^2$. Substituting these new relations into (3) and reversing the sign of the objective function, we obtain finally the dual LP,

$$\text{Maximize } \sum_{(ij)} u_{ij} f_{ij}^1 + \sum_{(ij)} l_{ij} f_{ij}^2 - \sum_{k \in K} d_k \lambda_k \tag{4.1}$$

subject to

$$\sum_{j \in A(1)} (f_{ij}^1 + f_{ij}^2) = \sum_k \lambda_k, \tag{4.2}$$

$$-\sum_{j \in A(i)} (f_{ij}^1 + f_{ij}^2) + \sum_{j \in B(i)} (f_{ji}^1 + f_{ji}^2) = \begin{cases} \lambda_k & \text{if } i = k \in K, \\ 0 & \text{if } i \notin K,\ i \neq 1, \end{cases} \tag{4.3}$$

$$0 \leqslant f_{ij}^1 \leqslant a_{ij},$$

$$0 \leqslant f_{ij}^2, \text{ with } f_{ij}^2 = 0 \text{ if } f_{ij}^1 < a_{ij} \text{ or } f_{ij}^1 = a_{ij} \text{ but } y_{ij} > l_{ij}, \tag{4.4}$$

$$0 \leqslant \lambda_k \leqslant p_k, \quad \text{for all } k \in K. \tag{4.5}$$

Assume hereafter that $t_1 = 0$; whence constraints (1.3) would now read as follows: $t_k - v_k \leqslant d_k$, $k \in K$.

Complementary slackness conditions on the optimal solution of the primal and dual LPs of (1) and (3) reveal that if $t_k \geqslant d_k$, $k \in K$, then λ_k may be > 0, while if $t_k < d_k$, then $\lambda_k = 0$. Define $s_{ij}^1 = t_j - t_i - u_{ij}$ and $s_{ij}^2 = t_j - t_i - l_{ij}$, which represent the slack in the times of realization constraints of (1.2) corresponding to the upper and lower bounds on the value of the duration y_{ij}, respectively.

The proposed iterative procedure rests on the fact that an intermediate step corresponds to a particular primal feasible solution, and that cost minimization is equivalent to flow maximization in a specific subnetwork, namely, the so-called *critical subnetwork*. Demonstrating this

equivalence also gives the clue to the general procedure to be followed. The arguments are typical primal-dual arguments throughout.

An initial feasible solution to the primal LP of (1) is easily obtained by setting $y_{ij} = u_{ij}$ for all $(ij) \in A$, $t_1 = 0$, and $t_j = \max_{i \in B(j)} (t_i + u_{ij})$. There shall exist at least one "critical path" to each node $k \in K$, and the collection of the arcs that define these paths constitutes the *critical subnetwork*, denoted by P. For each arc in P we have $s_{ij}^1 = 0$; consequently, the dual variables f_{ij}^1 and g_{ij} may have values > 0, for all $(ij) \in P$, and the dual variables $\{\lambda_k\}$ may have values > 0 for $k \in \bar{K}$, where $\bar{K} \subseteq K$ is the subset of those $k \in K$ for which $t_k > d_k$. (Note that, since we assumed $u_{ij} > l_{ij}\ \forall (ij)$, $s_{ij}^2 > 0 \rightarrow h_{ij} = 0$; hence by definition $f_{ij}^2 = 0$.) The objective function of the critical subnetwork P is

$$\text{Maximize} \sum_{(ij) \in P} u_{ij} f_{ij}^1 - \sum_{k \in \bar{K}} d_k \lambda_k. \tag{5}$$

Consider node $k \in \bar{K}$, and let the paths in P leading to it be designated by $\pi_q(k)$, $q = 1, 2, \ldots$. Let $f_{ij}^1(k)$ be the portion of flow in arc $(ij) \in P$ that goes to node k. Then the objective function (5) may be rewritten as

$$\sum_{k \in \bar{K}} \left[\sum_{q} \sum_{(ij) \in \pi_q(k)} u_{ij} f_{ij}^1(k) - d_k \lambda_k \right].$$

By conservation of flow along the path $\pi_q(k)$, it must be true that the portion of $f_{ij}^1(k)$ destined for key event k is constant along the path, say $f^1(q,k)$. Furthermore, $\sum_{(ij) \in \pi_q(k)} u_{ij} = t_k$, by the definition of the critical path to node k. Finally, $\sum_q f^1(q,k) = \lambda_k$, by (3.3). All of which reduces the objective in (5) to $\sum_{k \in \bar{K}} \lambda_k (t_k - d_k)$, which, since $v_k = t_k - d_k > 0$ for $k \in \bar{K}$, finally reduces to

$$\text{Maximize} \sum_{k \in \bar{K}} \lambda_k v_k. \tag{6}$$

For any given durations $\{y_{ij}\}$, the critical subnetwork is fixed and v_k is a constant that measures the tardiness of node $k \in \bar{K}$. Consequently, we can maximize (6) by maximizing the values of λ_k, which, in turn, implies our maximizing the flows in P. We achieve this [4] by selecting the most tardy event(s) and maximizing the flow to it, while respecting constraints (3.3) and (3.5), then the second most tardy event(s), and so forth until all nodes in $\bar{K}$ have been considered. This completes one cycle of iteration, since now the optimum of the restricted dual LP of (3) is in hand.

Focus is now shifted to the primal LP of (1), where the durations of activities are shortened in such a way as to maintain the complementary slackness conditions for optimality. An activity $(ij) \in A$ may be shortened if the following two conditions are satisfied simultaneously:

(i) $f_{ij}^1 = a_{ij}$ (no activity is shortened if its flow f_{ij}^1 is $< a_{ij}$),

(ii) if $(ij) \in \pi_q(k)$ for some k with $t_k = d_k$, then $\lambda_k = 0$ (no activity is shortened if it lies on a CP to some key event that is on time but whose associated λ is > 0).

An activity $(i_o j_o)$ satisfying these two conditions is called an *eligible activity*, which may be shortened until one of the following eventualities occurs: (1) Some $s_{ij}^1 = 0$ for $(ij) \in A - P$,

whence the critical subnetwork P will be augmented by a new path containing activity (ij); (2) $s^2_{i_o j_o} = 0$; i.e., the activity has been shortened to its lower bound $l_{i_o j_o}$; (3) the due date d_m for some node $m \in \bar{K}$ is met; i.e., $t_m = d_m$ for some $m \in \bar{K}$. Such compression is carried out for all eligible activities in P. When this phase is completed, at least one more slack variable s^1_{ij}, s^2_{ij} or v_k is driven to zero. In the case of some s^1_{ij} or s^2_{ij} becoming equal to zero, the corresponding dual activity (f^1_{ij} or f^2_{ij}, respectively) is added to the restricted dual problem, implying the augmentation of the restricted dual LP by that activity. Optimization of the newly augmented restricted dual proceeds as described above. In the case of some v_k being reduced to zero, an intermediate step is needed to ensure that its λ_k is also reduced to zero if at all possible. This is accomplished by rerouting the flow into node k to other nodes, as explained in the algorithm below. The termination of the process is realized when, for all $k \in K$, either of the following two conditions is satisfied:

(i) $t_k \leqslant d_k$,

(ii) $t_k > d_k$ and $\lambda_k = p_k$.

At that point the optimum is in hand.

ALGORITHM: The statement of the algorithm will be accompanied by some explanatory remarks to render it more accessible. (See also the flow chart, Figure 1.) Recall that A represents the set of arcs, $B(j)$ the set of nodes immediately before node j and connected to it with one arc, $\bar{K}$ the set of *tardy* due-dated key events, and P the subnetwork composed of all arcs on the critical paths to the nodes in the set $\bar{K}$, the so-called *critical subnetwork.*

STEP 0. *Initialization*: Set all $y_{ij} = u_{ij}$, $f^1_{ij} = 0 = f^2_{ij}$, $\forall\ (ij) \in A$; $t_1 = 0$, $\lambda_k = 0$, $\forall\ k \in K$. Determine all node realization times $t_j = \max_{i \in B(j)} (t_i + u_{ij})$ $j = 2, 3, \ldots, n$ and $(ij) \in A$

This step determines the normal times of realization of the nodes of the network prior to any compression.

STEP 1. *Determine the Critical Subnetwork P.* Determine $v_k = \max(0, t_k - d_k)$ for all $k \in K$. Let K denote the set of tardy nodes, $\bar{K} = \{k: v_k > 0\}$, and denote the difference set $K - K$ by K°; i.e., $K^\circ = \{k: v_k = 0\}$. If $K^\circ = K$, or if at any stage of iteration $\lambda_k = p_k$ for each $k \in \bar{K}$, stop; the optimum is in hand. Otherwise, let M denote the set of tardy nodes that have been examined. At the outset, $M = \{\phi\}$, the null set. Node k will be added to the set M in Step 2 if that node is the currently most tardy one. Construct the critical subnetwork P as the union of all arcs leading to nodes in $\bar{K}$.

STEP 2. *Determine the Most Tardy Key Event* : Let node k be the node such that $v_k = \max v_m$, $m \in \bar{K} - M$. If k is not unique, choose any one arbitrarily. Add k to the subset M.

This step identifies the most tardy key event not yet considered. The set M accumulates the tardy events that have been considered.

STEP 3. *Maximize Flow to Most Tardy Node*: The rationale behind the approach of augmenting the flow is that we are desirous of *first* diverting any previously passed flow to some other (more tardy) key event which may have a cheaper penalty than the currently investigated key event. This is done in the hope that any compression performed for the current key event may also reduce the delay in other key events and thus achieve considerable economy.

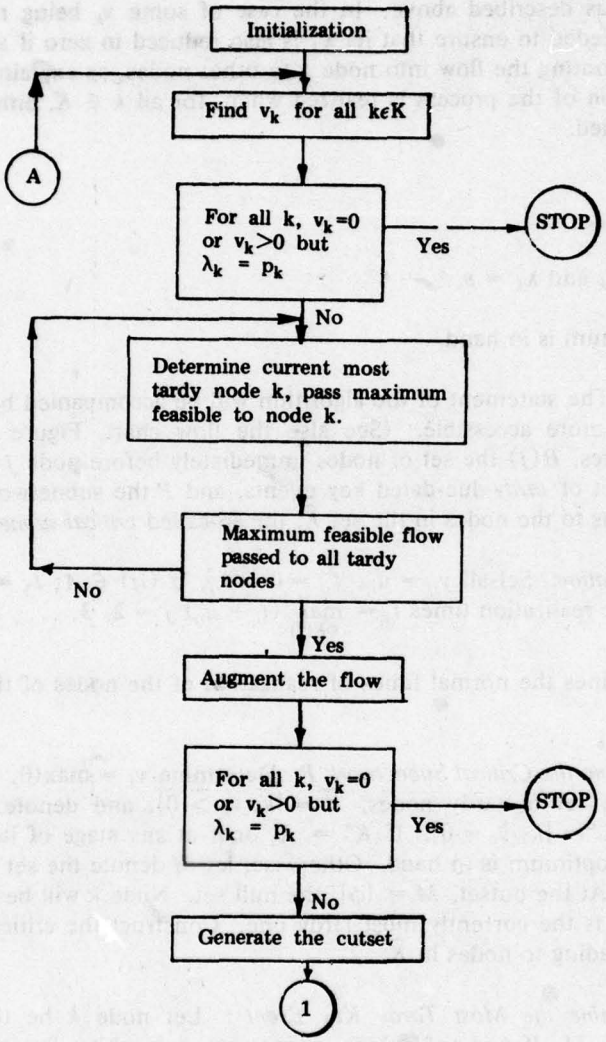


FIGURE 1. Flow chart of the algorithm

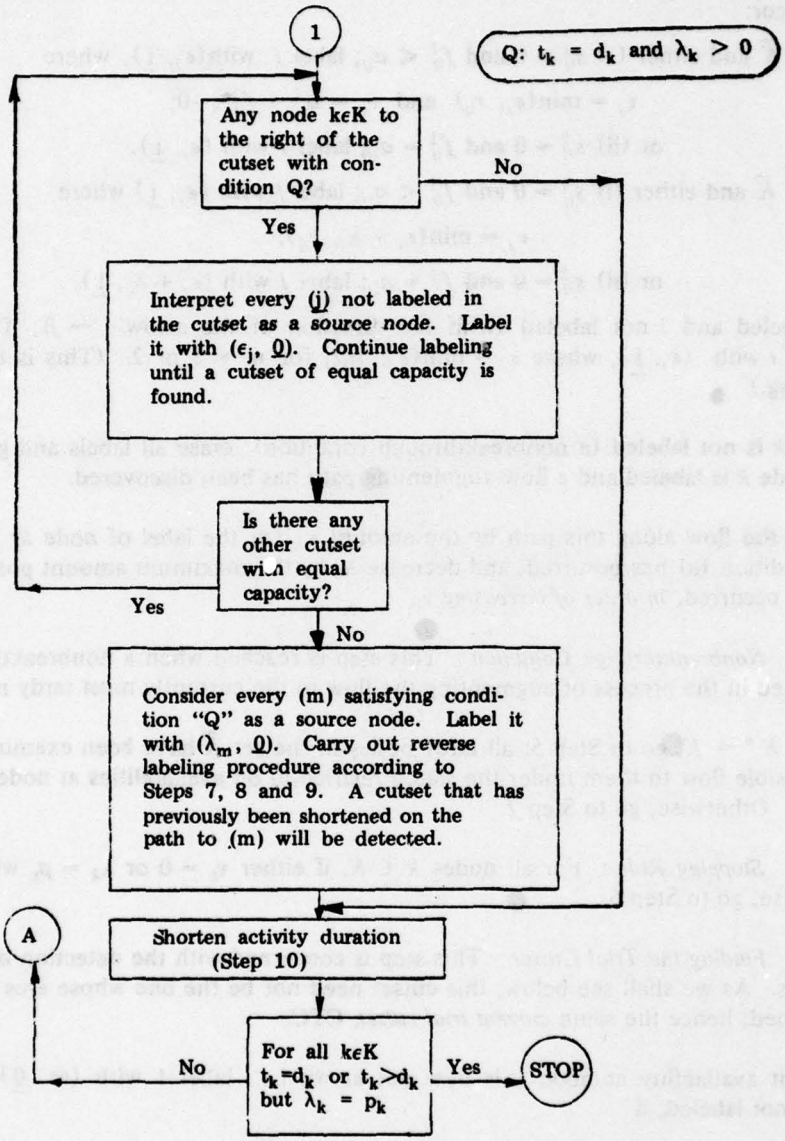


FIGURE 1 (Continued). Flow chart of the algorithm

Assume the limit availability at node 1 to be equal to $p_k - \lambda_k > 0$. Let $\Pi(k) = \{e: e \in \pi_q(k)\}$; i.e., $\Pi(k)$ is the subset of arcs that lie on the critical path(s) to node $(\in \bar{K})$. Label node 1 with $(p_k - \lambda_k, 0)$. (Recall that initially $\lambda_k = 0$.) In general, for any labeled node i and any node j not labeled, such that $(ij) \in \Pi(k)$, one of the following two conditions will occur:

(a) $i \notin \bar{K}$ and either (i) $s_{ij}^1 = 0$ and $f_{ij}^1 < a_{ij}$; label j with $(\epsilon_j, \underline{i})$, where

$$\epsilon_j = \min(\epsilon_i, r_{ij}) \quad \text{and} \quad r_{ij} = a_{ij} - f_{ij}^1 > 0;$$

or (ii) $s_{ij}^2 = 0$ and $f_{ij}^1 = a_{ij}$; label j with $(\epsilon_i, \underline{i})$.

(b) $i \in \bar{K}$ and either (i) $s_{ij}^1 = 0$ and $f_{ij}^1 < a_{ij}$; label j with $(\epsilon_j, \underline{i})$ where

$$\epsilon_j = \min(\epsilon_i + \lambda_i, r_{ij});$$

or (ii) $s_{ij}^2 = 0$ and $f_i^1 = a_{ij}$; label j with $(\epsilon_i + \lambda_i, \underline{i})$.

For any j labeled and i not labeled (with the direction of the arrow $i \rightarrow j$), if $s_{ij}^w = 0$ and $f_{ij}^w > 0$, label i with $(\epsilon_i, \underline{j})$, where $\epsilon_i = \min(\epsilon_j, f_{ij}^w)$, for $w = 1$ or 2. (This is the so-called reverse labeling.)

If node k is not labeled (a nonbreakthrough condition), erase all labels and go to Step 4. Otherwise, node k is labeled and a flow-augmenting path has been discovered.

Increase the flow along this path by the amount ϵ_k (= the label of node k) in the usual manner if condition (a) has occurred, and decrease λ_i by the maximum amount possible if condition (b) has occurred, *in order of increasing* v_i.

STEP 4. *Nonbreakthrough Condition* : This step is reached when a nonbreakthrough condition is reached in the process of augmenting the flow to the currently most tardy note $k \in \bar{K}$.

If $M \cup K^\circ = K$, go to Step 5; all tardy nodes in the set $\bar{K}$ have been examined, and the maximum feasible flow to them under the stated restriction on availabilities at node $\underline{1}$ has been accomplished. Otherwise, go to Step 2.

STEP 5. *Stopping Rule* : For all nodes $k \in K$, if either $v_k = 0$ or $\lambda_k = p_k$ when $v_k > 0$, stop. Otherwise, go to Step 6.

STEP 6. *Finding the Trial Cutset* : This step is concerned with the detection of a cutset to the key events. As we shall see below, this cutset need not be the one whose arcs will eventually be shortened; hence the same *current trial cutset*, CTC.

The limit availability at node 1 is now put at ∞; i.e., label 1 with $(\infty, \underline{0})$. For any i labeled and j not labeled, if

(a) $s_{ij}^1 = 0$ and $f_{ij}^1 < a_{ij}$ or $s_{ij}^2 = 0$ and $f_{ij}^1 = a_{ij}$, label j with $(\epsilon_j, \underline{i})$, where $\epsilon_j = \min(\epsilon_i, r_{ij})$; $r_{ij} = a_{ij} - f_{ij}^1$ in the former eventuality, and $r_{ij} = \infty$ in the latter.

(b) $f_{ij}^1 = a_{ij}$ and either $s_{ij}^1 = 0$ or $(s_{ij}^1 < 0$ and $s_{ij}^2 > 0)$, then j cannot be labeled from i. Arc (ij) is a member of the CTC.

Successive application of this step will eventually result in the detection of the complete CTC.

STEP 7. *Testing the CTC* : The CTC determined in Step 6 is checked for "feasibility," which refers to the satisfaction of the complementary slackness conditions for optimality. Verify the following condition:

(7) Q: {There exists $k \in K$, k not labeled, subject to $t_k = d_k$ and $\lambda_k > 0$}

There are two possibilities :

(a) Conditions Q is *false*. Then the CTC is a *feasible* cutset. Go to Step 10.

(b) Condtion Q is *true*. Then the CTC is an *infeasible* cutset under current flow conditions. We must determine either: (1) another cutset to the right of the CTC that is of equal capacity but for which condition Q is false, hence it is feasible; of (2) modify the flow so that condition Q is false for the CTC. These two objectives are accomplished by the following subroutines.

Consider every unlabeled node $j \in [P \cap \text{CTC}]$ as a *source node* with label $(\epsilon_j, \underline{0})$. For each such node, continue labeling as in Step 3 until one of three condtions occurs:

(i) A tardy node $k \in \bar{K}$ is labeled with ϵ_j from j. Erase the the ϵ labels generated by this node j only. Continue with another node $j \in [P \cap \text{CTC}]$.

(ii) A tardy node $k \in \bar{K}$ is labeled with $\epsilon_k > \epsilon_j$, say $\epsilon_k = \epsilon_j + \delta$, $\delta > 0$. (The additional flow β will emanate from some node k in the set $\bar{K}$ with $\lambda_k > 0$; see Step 3, condition (b).) Augment the flow into node k by δ, relabel node j with $(\epsilon_j, \underline{j})$, and repeat the labeling process.

(iii) No tardy node $k \in \bar{K}$ is labeled from j. Retain the newly generated ϵ_j labels. Continue with another node $j \in [P \cap \text{CTC}]$.

The outcome of this phase is a set of newly labeled nodes (with ϵ-labels) — which set may be empty if all nodes $j \in [P \cap \text{CTC}]$ result in breakthrough to some key node $k \in \bar{K}$ — and the originally labeled nodes to the left of the CTC. The new CTC is now defined by the set of arcs that separate the labeled nodes (both old and new) from the unlabeled ones. If the new CTC shares at least one arc with the old CTC, then the new CTC is infeasible; go to Step 8. However, if the new CTC is different from the old, go to the start of Step 7 with the new CTC.

Step 8. *The Resolution of Infeasibility* : The algorithm arrives at this step when the CTC is infeasible, because there exist activities which are the only current cheapest ones to be shortened (to decrease the tardiness of some key event(s)), but some of these activities are on a path to at least one key event that satisfies condition Q. For these key events, we either reduce their λ's to zero, or a *previously shortened* activity (or set of activities) must be detected and *lengthened* by an amount equal to the shortening of the activities on the CTC, thus leaving unchanged the realization times of these key event.

Define the set

$$S_1 \triangleq \{k: v_k = 0,\ \lambda_k > 0 \text{ and no } j \in A(k) \text{ is in } P \text{ with positive flow into it}\},$$

where $A(k)$ is the set of nodes immediately following node k and connected to it with an arc. Note that, according to the construction in Step 7, $S_1 \neq \phi$ at the outset, though it may become

empty later on. (If $S_1 = \phi$, go to Step 10.) In the subnetwork of *unlabled nodes*, label each $k \in S_1$ as a *source* node with $(\lambda_k, \underline{0})$. Continue labeling until one of the following, (a) or (b), occurs.

(a) A node $k' \in \bar{K}$ is labeled. Augment the flow into node k' in the usual manner (see Step 3). If λ_k is reduced to zero for some $k \in S_1$, let $S_1 = S_1 - \{k\}$. If S_1 becomes empty, go to Step 7. Otherwise, label node k with $(\lambda_k^{new}, \underline{0})$ where λ_k^{new} is the reduced value of λ_k, and continue labeling until a cutset is detected. Eventually, either no node $k \in S_1$ satisfies condtion Q (hence the set S_1 is empty), in which case return to the start of Step 7, or case (b) occurs.

(b) Node i cannot be labeled from node j. If arc $ij \in P$, j labeled and i not labeled, then ij is a *candidate* for joining the CTC. (The manner in which these candidate arcs are handled is treated in Step 9 below.) If arc $ji \in P$, j labeled and i not labeled, then ji joins the CTC.

At the end of this step, all possible *candidate arcs* to join the CTC are in hand. If there are such candidates, go to Step 9. Otherwise, if there are no candidates, a new CTC has been determined that is to the right of all nodes satisfying condition Q, hence it is feasible; go to Step 10.

STEP 9. In this step the feasibility of the *candidate* activities is checked. If the candidate activities are to the left of all nodes satisfying condition Q that lie on paths containing these candidate activites, then they are *feasible*, in the sense that they can be lengthened. Otherwise, the algorithm continues its search for the leftmost activities to be lengthened. We acheive this by declaring the unlabeled nodes incident on the candidate activities as *temporary source nodes*, one at a time, and labeling them with T. For any j labeled with T, i not labeled, and

(i) $ij \in P$, $s_i = 0$ and $f_{ij}^w > 0$; $w = 1, 2$, label i with T.

(ii) $ji \in P$, $s_{ji}^1 = 0$ and $f_{ji}^1 < a_{ji}$ or $s_{ji}^2 = 0$, label i with T.

continue labeling until either

(a) A node in the CTC is labeled; erase all T labels;
or

(b) Nonbreakthrough to any node in CTC occurs. A new set of *candidate* activities is in hand. Change the labels of those nodes that are to the right of (i.e. occur after) candidate activities from T labels to P labels. Go to the start of Step 9. In Step 10, nodes that have P labels are considered as labeled nodes.

This step terminates when labeling of all nodes T incident to the candidate activities result in (a). A feasible cutset is in hand. Go to Step 10.

STEP 10. *Shortening Activity Durations* : This step is reached when either the CTC is declared feasible, or the reverse labeling of Steps 8 and 9 has terminated. In either case we have a cutset C. Let $C_1 \triangleq$ {set of labeled nodes} and $C_2 \triangleq$ {set of nodes that are not labeled}. The forward arcs in C are to be shortened and the reverse arcs are to be lengthened.

(a) For all forward arcs of C, if

$$y_{ij} \geqslant u_{ij} \text{ and } s_{ij}^1 > 0, \text{ set } \delta_{ij} = s_{ij}^1;$$

$$y_{ij} \leqslant u_{ij} \text{ and } s_{ij}^1 \leqslant 0, \text{ set } \delta_{ij} = s_{ij}^2.$$

(b) For all reverse arcs of C, if

$$y_{ij} < u_{ij}, \text{ set } \delta_{ij} = -s_{ij}^1;$$
$$y_{ij} = u_{ij}, \text{ set } \delta_{ij} = \infty.$$

(c) For nodes $k \in \bar{K} \cap C_2$, determine $v_k = t_k - d_k$.

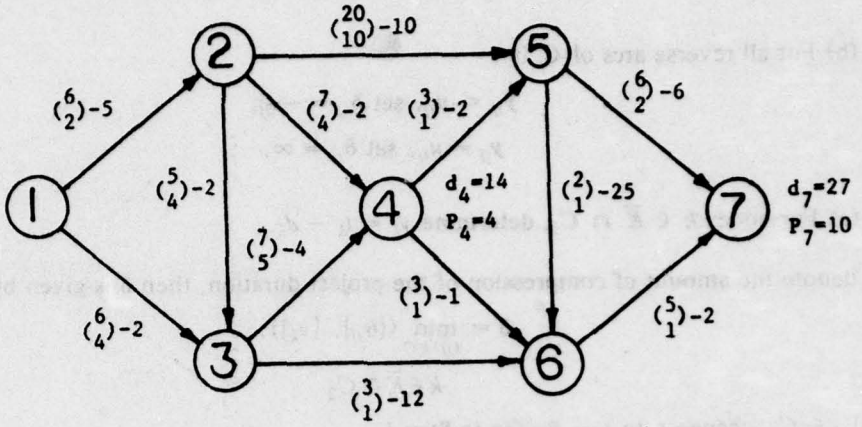
Let δ denote the amount of compression of the project duration, then δ is given by

$$\delta = \min_{(ij)\in C} (\{\delta_{ij}\}, \{v_k\}).$$
$$k\in\bar{K}\cap C_2$$

For all $j \in C_2$, change t_j to $t_j - \delta$. Go to Step 1.

Steps 7 and 8 are concerned with the detection of a feasible cutset, if it exists. For, indeed, if condition Q of (7) is satisfied and one undertakes the shortening of the duration of the CTC, then one would violate the complementary slackness optimality conditions, since now some key event k will be such that $t_k < d_k$ *and* $\lambda_K > 0$. If a feasible cutset is detected, it is compressed. Otherwise, *Steps 9* and *10* purport to discover the *last shortened* cutset (which is also the most expensive) and *lengthen* those arcs in it that would offset the shortening of the other arcs, so that, for those nodes satisfying Condition Q, their t_k's remain equal to their d_k's, and the complementary slackness conditions are preserved. The whole procedure terminates either when $t_k = d_k$ for all $k \in K$ or for some key events $t_k > d_k$ but $\lambda_k = p_k$, which indicates that it is more expensive to compress any cutset in the critical subnetwork P than to advance the completion time of these events, at which time iteration is halted.

EXAMPLE: Consider the network shown in Figure 2. There are two key events: $K = \{\underline{4}, \underline{7},\}$ with $d_4 = 14$ and $p_4 = 4$; $d_7 = 27$ and $p_7 = 10$. For the sake of clarity of the figures, we subsequently refrain from explicitly showing u_{ij} and l_{ij} on each arc (ij); the arc duration can be easily deduced from $t_j - t_i$. Initialization, followed by Steps 1 and 2, results in the critical subnetwork shown in heavy lines in Figure 3. Since $v_4 = 4$ and $v_7 = 6$, node $\underline{7}$ is the most tardy node. Step 3 is initiated with $(10, \underline{0})$ at node $\underline{1}$, since $p_7 = 10$ and $\lambda_7 = 0$. The first "go around" of Step 3 terminates with the flows shown in Figure 3, with no breakthrough to node $\underline{7}$. It is reinitiated with $k = 4$ and labeling $(4, \underline{0})$ at node $\underline{1}$, to result in the second go around, with no breakthrough to node $\underline{4}$ and the flows as shown in Figure 4. Since now $M = K$, the conditions of Steps 4 and 5 lead to Step 7, and it is easy to see that the cutset is feasible (since neither v_4 nor v_7 equals zero), and $C_1 = \{(2,3), (6,7)\} \subset P$, as shown in Figure 5. Here, $C_1 = \{1,2,5,6\}$, $C_2 = \{3,4,7\}$, and $\delta_1 = 1$. Step 9 shortens the activities in C by 1 unit, resulting in the node realization times shown in Figure 6. Note that *both* tardy nodes have advanced in time by δ: node $\underline{4}$ now is realized at time 17 and node $\underline{6}$ at time 32. Both are still tardy. This compression step has cost $2+2 = 4$ units but saved $4+10 = 14$ units, an obvious economic advantage. Moreover, the second arc between nodes $\underline{2}$ and $\underline{3}$ is introduced to carry the flow $f_{2,3}^2$, since $s_{2,3}^2 = 0$ and $f_{2,3}^1 = 2 = a_{2,3}$. Iteration is returned to Step 1, with node $\underline{7}$ as the most tardy node. Node $\underline{1}$ is now labeled $(8, \underline{0})$, since $p_7 - \lambda_7 = 8$. Flow maximization results in saturating arc (1,2), and the nonbreakthrough condition is quickly reached. The next most tardy node is $\underline{4}$, but no additional flow can be passed to it; therefore, the minimal cost is in hand, since $M = \{4,7\}$; it is $C_2 = \{(1,2)\} \subset P$ (see Figure 7). It is easy to see that condition Q is not satisfied, and that $\delta = 3(= -v_4)$. Shortening the project duration by 3 units results in the node time realization shown in Figure 8. Now node $\underline{7}$ is the only tardy node, and it is obvious



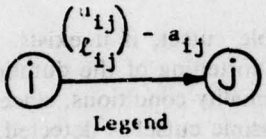
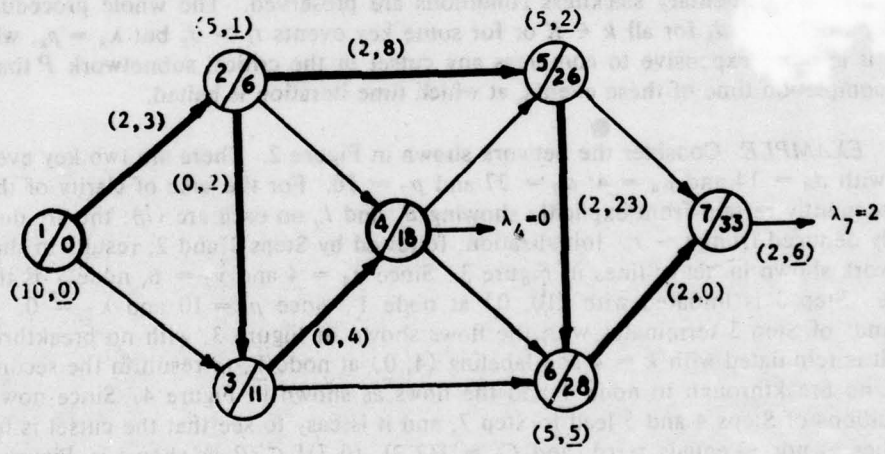
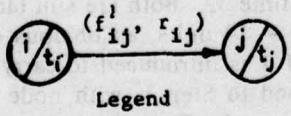


FIGURE 2. Statement of network problem





$(f^1_{ij} \leq a_{ij};\ r_{ij} = a_{ij} - f^1_{ij})$

FIGURE 3. Labeling for flow augmentation to node 7

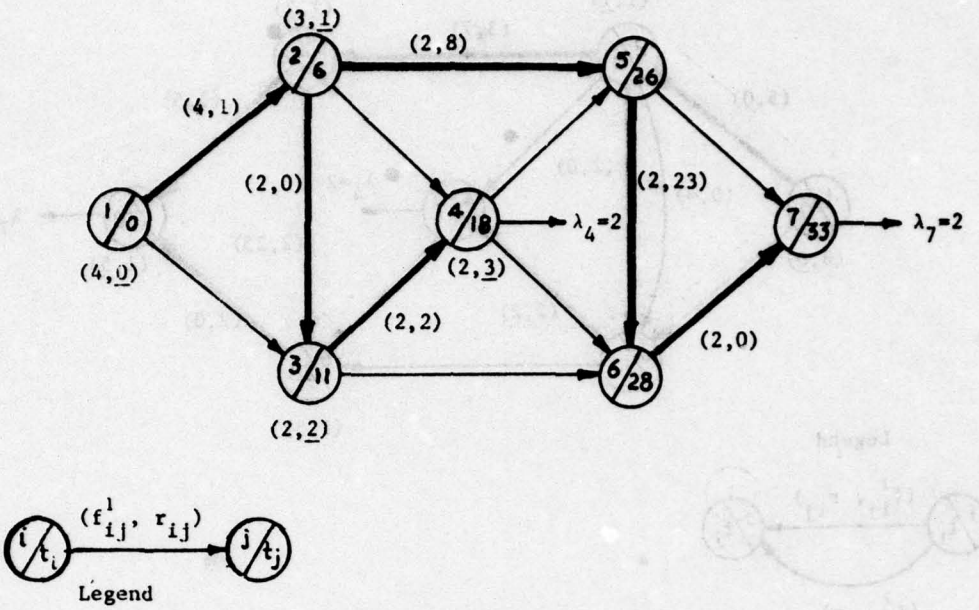


FIGURE 4. Labeling for flow augmentation to node 4

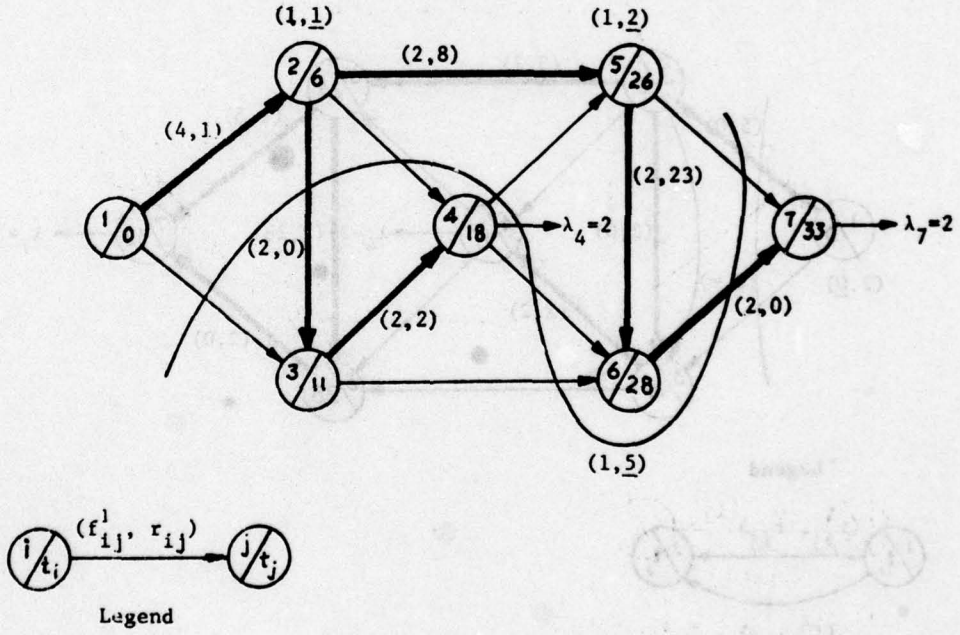


FIGURE 5. Resultant cutset

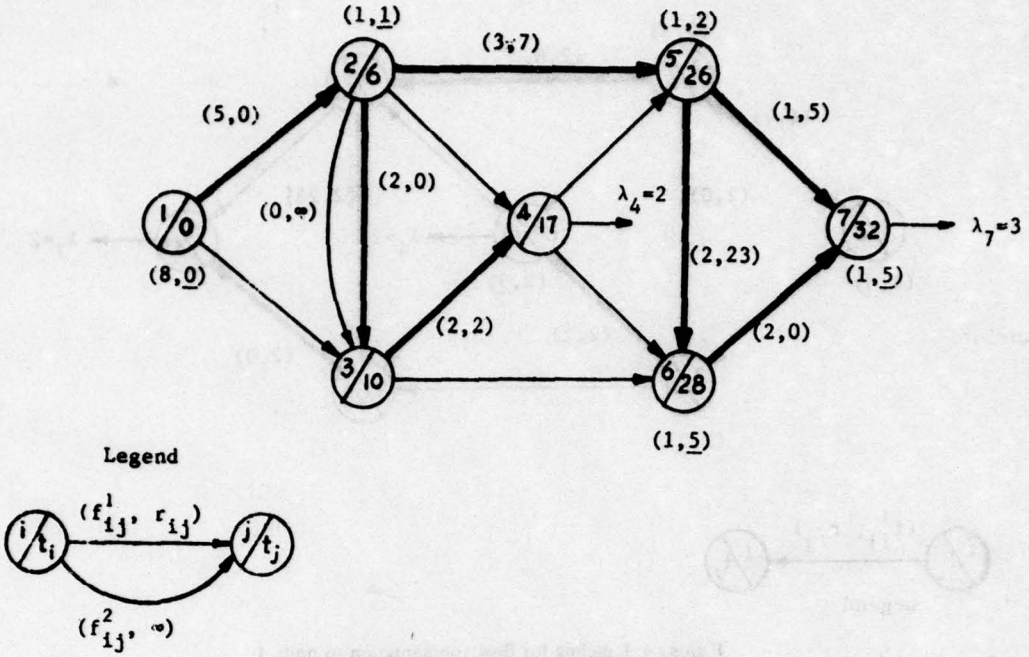


FIGURE 6. Result of shortening activities (2,3) and (6,7) by 1

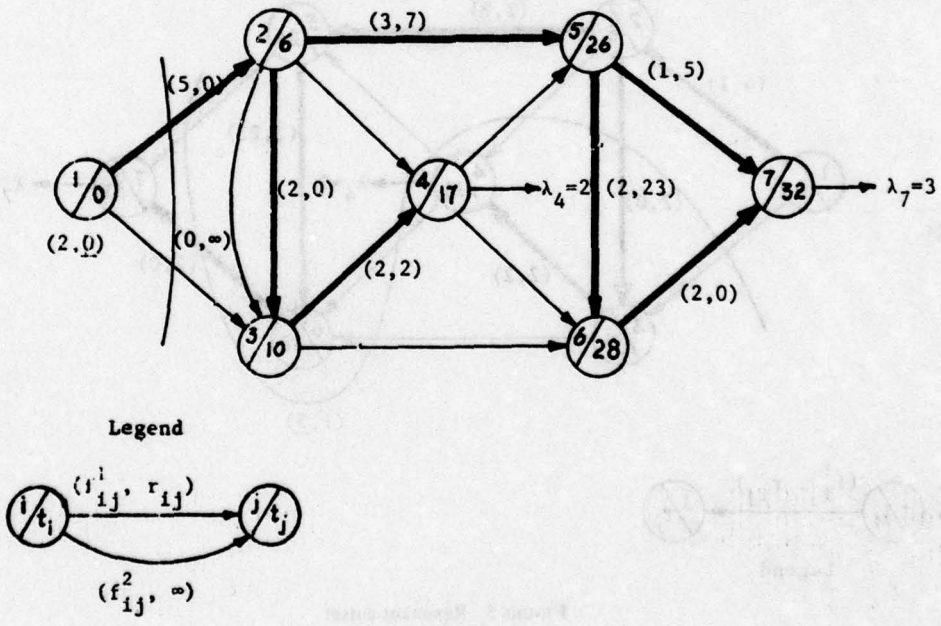


FIGURE 7.

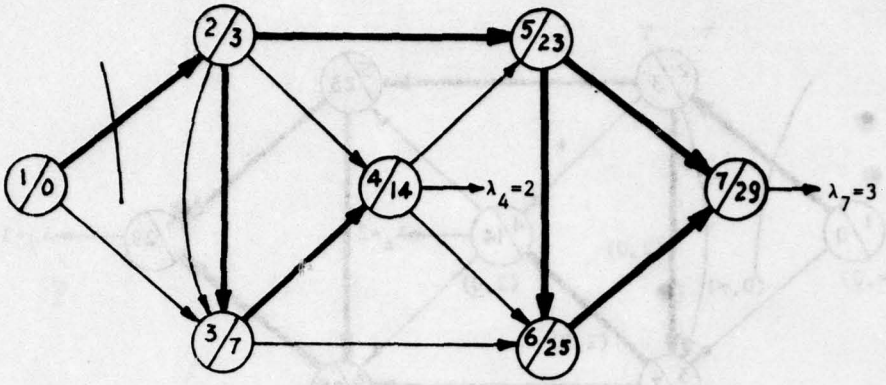


FIGURE 8.

that no additional flow can be sent to it, since arc (1,2) is saturated. Therefore, the cutset $C_3 = \{(1,2)\} \subset P$. Unfortunately, Condition Q of (7) is true for node 4, since $t_4 = 14 = d_4$ and $\lambda_4 = 2$. Consequently, activity (1,2) cannot be shortened, and Step 7(b) is initiated.

Node 2 is labeled $(\epsilon_2, 0)$ (see Figure 9) leading to labeling of nodes 5, 6, and 7, a breakthrough (Step 7(i)). All ϵ labels are then erased. Since node 2 is the only labeled node in the CTC, we conclude that it is infeasible, and go to Step 8.

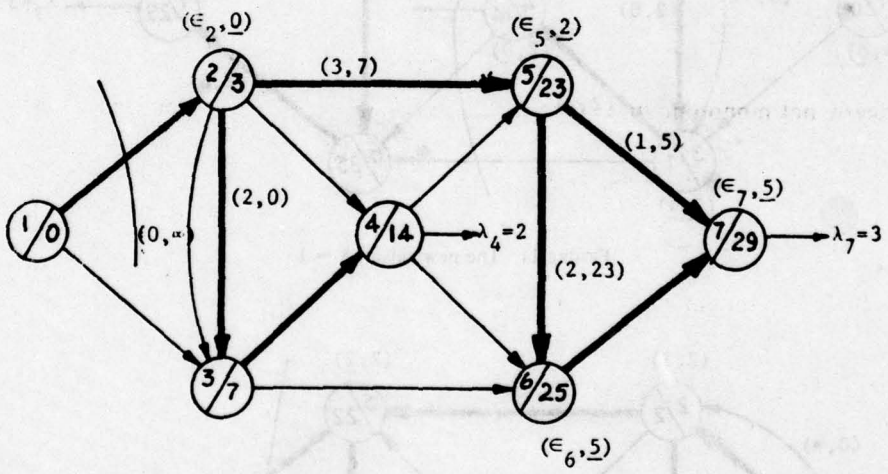


FIGURE 9. Labeling form $(\epsilon_1, 0)$ of node 2. Breakthrough

Step 8 specifies labeling of node 3 with (2, 4) (see Figure 10). Node 2 cannot be labeled from 3 because $f^2_{2,3}$ while $s^2_{2,3} = 0$; hence no other node can be labeled, and Step 8(b) is initiated. Since arc (2,3) ϵ P and node 3 is labeled but node 2 is not, arc (2,3) is now a candidate for joining the CTC. Node 4 is the only node in the set S_1 and it is labeled, so arc (2,3) is the only such candidate. Step 9 requires labeling 2 with T (shown also in Figure 10), which is in the CTC indicating that arc (2,3) is to the left of all nodes satisfying Condition Q, hence the T-label is erased (Step 9(a)). The cutset in hand is feasible. The set of labeled nodes is $C_1 = \{1,3,4\}$; the set of unlabeled nodes is $C_2 = \{2,5,6,7\}$, with the cutset $C_4 = \{(1,2), (2,3)\} \subset P$, as shown in Figure 11. It is easy to deduce that $\delta = 1$, and the subsequent compression step shortens activity (1,2) but *simultaneously lengthens* activity (2,3) by one unit. The realization time of node 4 is left unaffected at 14, whereas the time of realization of node 7 is decreased

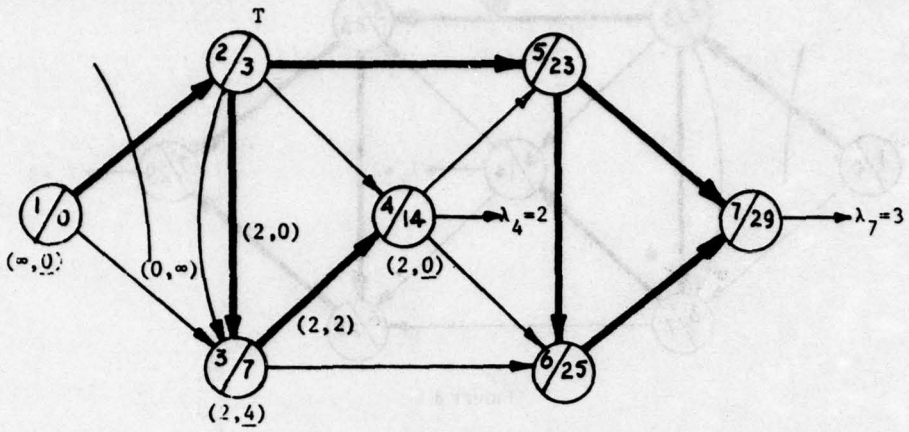


FIGURE 10. Labeling for (2, 0) of node 4

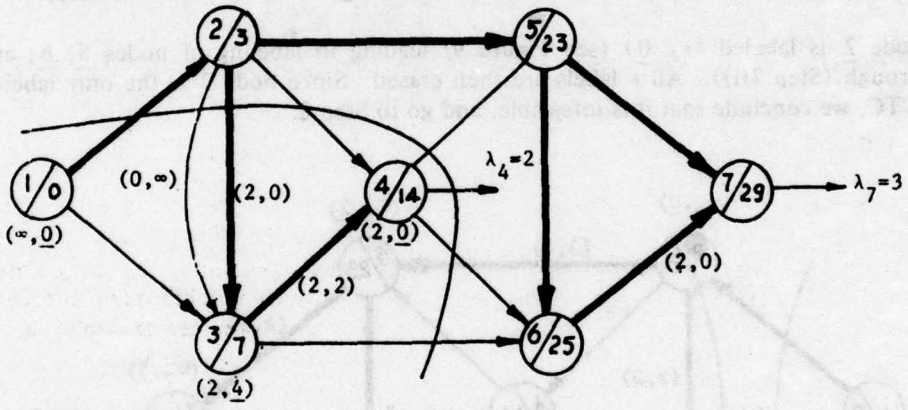


FIGURE 11. The new cutset; $\delta = 1$

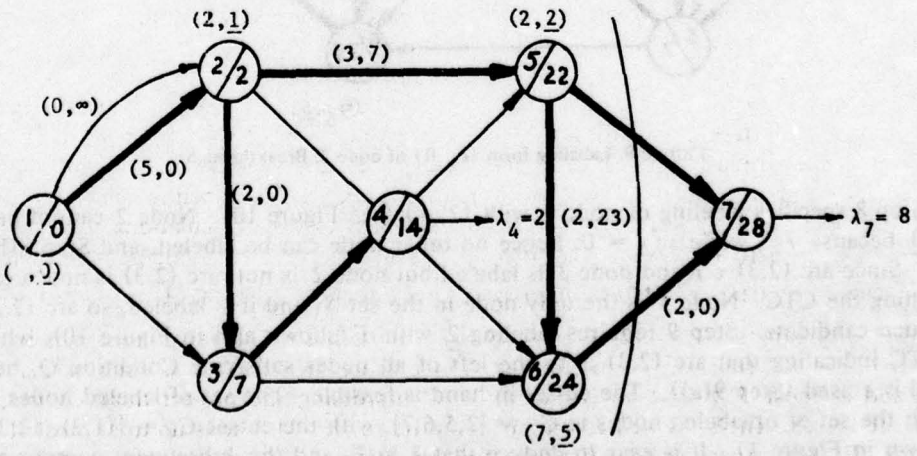


FIGURE 12. Reduced durations and new labeling

by 1 unit. The situation now is as shown in Figure 12. Note the second arc of infinite "capacity" between nodes 1 and 2, introduced because now $\delta^1_{1,2} = 0$. Since node 7 is still the only tardy node, the iterations are initiated with node 1, labeled with (7, 0), since $p_7 - \lambda_7 = 7$. Flow maximization saturates arc (5,7), resulting finally in the cutset $C_5 = \{(5,7), (6,7)\}$. Compressing the duration of the project by $\delta = 1 = \min\{d_7 - t_7;\ t_7 - l_{5,7};\ t_7 - l_{6,7}\}$, node 7 reaches its due date and iteration is halted. Figure 13 gives the optimal solution. The cost of compression is 30, and the gross savings amount to 76; hence, the net gain is 46 units.

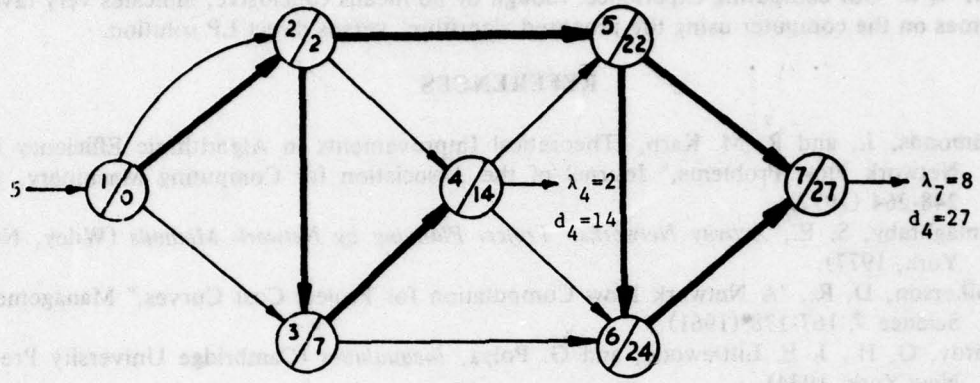


FIGURE 13. Optimal solution

CONCLUDING REMARKS

For the sake of brevity in exposition we chose a small sample project of only 12 activities and seven nodes. Consequently, it did not exhibit all the possibilities that may arise in the course of iteration. (For instance, this example did not permit a breakthrough to a tardy node from a node in the set S_1.) In a separate report [6] we present a larger example, in which all the branches of the algorithm are taken, together with the computer code and some experimental results. For the moment, it suffices to indicate that our preliminary trials with networks of up to 10 nodes and 18 arcs and only *two* key events were solved, and they consumed between 0.4 and 0.9 seconds on the IBM 370/175.

We have advocated a network flow algorithm, in lieu of a frontal attack on the LP of (1), in order to capitalize on the special structure of that LP. Therefore, it would be of interest to determine an upper bound on the number of computations required.

The complexity of the proposed algorithm can be gleaned from the following calculation. We assume that all durations and due dates are integers. It has been shown [1] that to acheive flow maximization requires $0(A^2n)$ computations. There are at most $K \leqslant n$ tardy nodes, each of which may generate a CTC which, in turn, would initiate new labelings in search of a feasible cutset. The resulting search would require at most $0(A^2n)$ calculations, repeated at most n-2 times (since there are at most that many CTCs in a network of n nodes). Thus, to locate a feasible cutset or declare a CTC infeasible would require at most $\bar{K} \times 0(A^2n) + \frac{n}{2}(n-2) \times 0(A^2n)$ calculations. If infeasibility of the CTC is established, reverse labeling is initiated for each node $k \in K^*$ satisfying Condition Q. Since there are at most K such nodes, this step would add at most $K \times 0(A^2n)$ calculations. Compressing the project duration involves the arcs in the feasible cutset as well as the nodes in the set $\bar{K}$. This adds up to at most $A + K$ calculations. Finally, assuming that the problem was originally stated in integers, each δ is $\geqslant 1$,

and there may be at most $\sum_{k \in \bar{K}} v_k$ compressions, i.e., repetitions of the whole procedure. Totalling up the individual steps, we obtain

$$\left[\bar{K} \times 0(A^2 n) + \frac{n}{2}(n-2) \times 0(A^2 n) + K \times 0(A^2 n) + (A + K)\right] \sum_{k \in K} v_k$$
$$\leqslant \left(\sum_{k \in K} v_k\right) \times 0(A^2 n^3),$$

since $K \leqslant n$. Our computing experience, though by no means conclusive, indicates very favorable times on the computer using the proposed algorithm, versus direct LP solution.

REFERENCES

[1] Edmonds, J., and R. M. Karp, "Theoretical Improvements in Algorithmic Efficiency for Network Flow Problems," Journal of the Association for Computing Machinery, *19*, 248-264 (1972).

[2] Elmaghraby, S. E., *Activity Networks: Project Planning by Network Methods* (Wiley, New York, 1977).

[3] Fulkerson, D. R., "A Network Flow Computation for Project Cost Curves," Management Science *7*, 167-178 (1961).

[4] Hardy, G. H., J. E. Littlewood, and G. Polya, *Inequalitites* (Cambridge University Press, New York, 1934).

[5] Lawler, E. L., *Combinatorial Optimization: Networks and Matroids* (Holt, Rinehart and Winston, New York, 1976).

[6] Pulat, P. S., "Illustrative Example and Program Documentation for Optimal Project Compression with Due-Dated Events," OR Report No. 138, North Carolina State University, Raleigh, North Carolina (October, 1978).

A NEW ANALYSIS OF A LOT-SIZE MODEL WITH PARTIAL BACKLOGGING

David Rosenberg

School of Business Administration
Old Dominion University
Norfolk, Virginia

ABSTRACT

We reformulate the cost equation for the lot-size model with partial backlogging. The formulation is in terms of "fictitious demand rate," a new inventory decision variable that simplifies the analysis. Using decomposition by projection, we obtain an optimal solution in a straightforward manner. The form of the solution sheds additional light on the behavior of the model. Some of these insights are elucidated by numerical examples.

INTRODUCTION

Most works in inventory theory on infinite-time-horizon lot-size models have been concerned with the extreme cases, wherein all demand occurring during stockout is backlogged or not backlogged (lost sales). The abundance of such models contrasts sharply with the scarcity of inventory models that consider the hybrid situation of partial backlogging. A realistic application of the partial-backlogging concept is in the demand for spare parts. Intuitively, it seems reasonable to assume that, during stockout, critical needs will be satisfied from other sources and less urgent needs will be met by backordered items.

The scant inventory literature on infinite-time-horizon lot-size models with partial backlogging contains only one such model for which an optimal solution has been obtained ([1], pp. 256-259). The purpose of this paper is to show that the analysis of this solved model and the mathematical form of the resulting optimal solution can be greatly simplified. This simplification results from reformulating the model in terms of a new inventory decision variable, which we term the "fictitious demand rate" (FDR). The decision variable FDR is a crucial modeling concept in the analytical study of monopoly price — inventory models [2]. An important byproduct of the simplified analysis of partial backlogging is a new economic interpretation of the circumstances under which this operating doctrine is optimal.

The remaining sections of the paper review the optimization procedure of Montgomery, Bazaraa, and Keswani (MBK) [1], develop the revised version of the infinite-time-horizon lot-size model with partial backlogging, establish an optimal policy, discuss and interpret the policy, and give numerical illustrations.

BACKGROUND

The following variable-cost-rate model for a lot-size inventory system with partial backlogging is given in ([1] p. 256):

(1a) $$K(Q, S) = \frac{AD}{Q + (1-b)S} + \frac{IC(Q - bS)^2}{2[Q + (1-b)S]} + \frac{\pi SD}{Q + (1-b)S} + \frac{\bar{\pi} bS^2}{2[Q + (1-b)S]} + \frac{\pi_0(1-b)SD}{Q + (1-b)S},$$

where

D = demand rate (DR), $D > 0$,
Q = order quantity, $Q > 0$,
C = unit cost, $C > 0$,
I = carrying rate as a per cent of unit cost, $I > 0$,
A = cost per order, $A > 0$,
S = stockout demand per inventory review cycle, $S > 0$,
π = unit-shortage penalty, $\pi \geqslant 0$,
$\bar{\pi}$ = backorder unit-shortage cost rate, $\bar{\pi} \geqslant 0$,
π_0 = unit profit, $\pi_0 \geqslant 0$,
b = fraction of stockout-demand backordered, $0 < b < 1$.

The nonconvexity of (1a) motivated MBK to create the *ad hoc* optimization procedure which we proceed to outline. By use of the nonsingular transformation

$$U = Q + S(1-b)$$
$$V = Q - bS$$

model (1a) became

(1b) $$K(U, V) = \frac{AD}{U} + \frac{ICV^2}{2U} + \frac{\pi D(U-V)}{U} + \frac{\bar{\pi}(U-V)^2}{2U} + \frac{\pi_0 D(1-b)(U-V)}{U}$$

Observing the nonconvexity of (1b), MBK applied the projection concept. To facilitate the use of projection, the transformation

$$\beta = \frac{V}{U}$$

was applied to (1b), which became

$$Y(U, \beta) = \frac{AD}{U} + \pi D + \pi_0 D(1-b)(1-\beta) + U\left[\frac{\bar{\pi} b}{2}(1-\beta)^2 + \frac{IC}{2}\beta^2\right].$$

Then decomposition by projection was applied sequentially in the following manner:

$$Y(U^*, \beta^*) = \min_{\beta \leqslant 1} \min_{U} Y(U, \beta) = \min_{\beta \leqslant 1, U} Y(U, \beta).$$

The outer minimization, unfortunately, was not straightforward. This optimization difficulty is circumvented by our proposed reformulation which yields a sequence of strictly convex subproblems.

REFORMULATION

Model (1a) is based on the assumption of a uniform demand rate, which implies that the following formula for inventory review cycle length T holds:

$$T = [Q + (1-b)S]/D.$$

For expository convenience we rewrite model (1a) in terms of T,

(1c) $$K(Q, S) = \frac{A}{T} + \frac{IC(Q\ bS)^2}{2DT} + \frac{\pi S}{T} + \frac{\bar{\pi} bS^2}{2DT} + \frac{\pi_0(1-b)S}{T}.$$

The analysis will be greatly simplified if we introduce the new inventory decision variable, fictitious demand rate X, which we define by means of the following transformation:

$$X = (Q - bS)/T.$$

Substituting X in the appropriate places in model (1c), we obtain the following reformulation:

$$\text{(2a)} \qquad C(X, T) = \frac{A}{T} + \frac{ICTX^2}{2D} + \pi(D - X) + \frac{\bar{\pi}bT(D - X)^2}{2D} + \pi_0(1 - b)(D - X).$$

In Figure 1 we give the geometry of partial backlogging and indicate the key decision variables the reader has encountered up to this point. From a geometric standpoint we are modeling the stockout level as the product of the inventory review cycle length T and the term $(DR - FDR)$, which is the difference between the actual demand rate and an artificial demand rate.

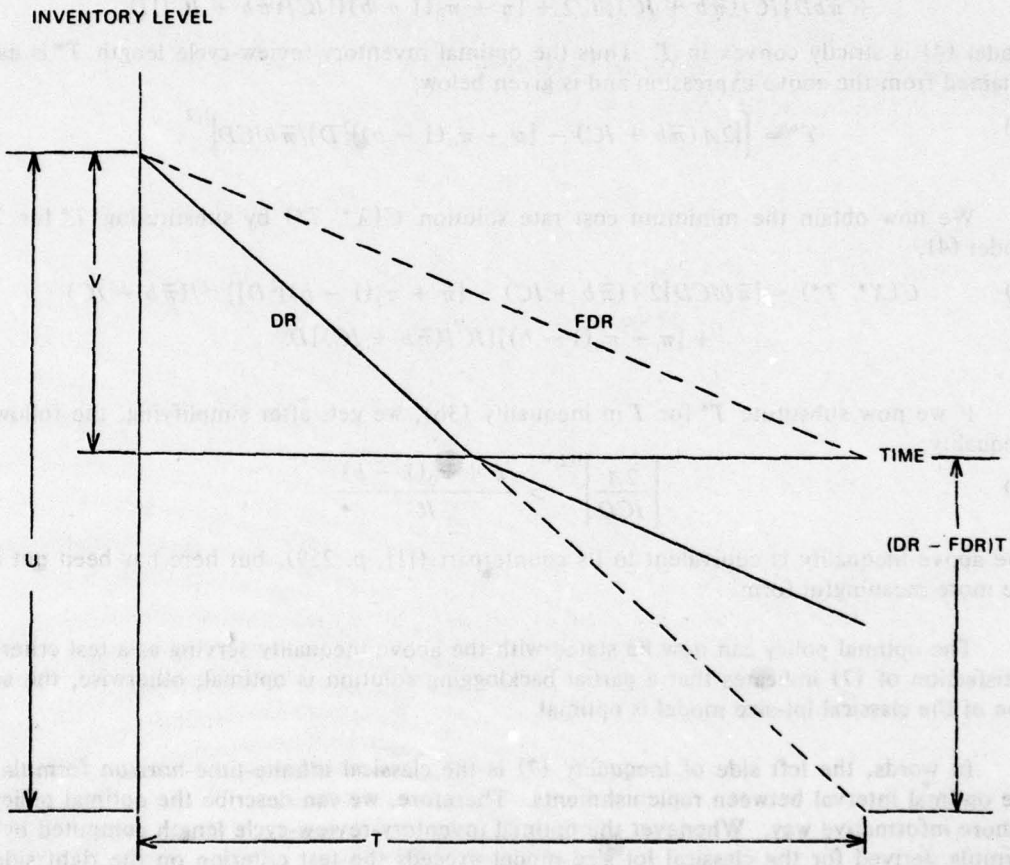


FIGURE 1. The inventory fluctuations in a lot-size model with Partial Backlogging

Preparatory to seeking an optimal solution, we write model (2a) in the following equivalent form,

$$\text{(2b)} \qquad C(X,T) = X\left\{\frac{(IC + \bar{\pi}b)T}{2D}X - [\bar{\pi}bT + \pi + \pi_0(1 - b)]\right\} + \frac{A}{T} + \frac{\bar{\pi}bDT}{2} + [\pi + \pi_0(1 - o)]D.$$

We shall assume that model (2) is twice continuously differentiable, since this validates the subsequent use of the projection technique for obtaining an optimal solution.

OPTIMAL SOLUTION

We obtain an optimal solution by applying decomposition by projection in the following stepwise manner:

$$C(X^*, T^*) = \min_T \min_X C(X, T) = \min_{X,T} C(X, T).$$

For fixed positive T, (2b) is strictly convex in X. Hence, the optimal fictitious demand rate $X^*(T)$ is easily obtained from model (2b),

$$X^*(T) = [\bar{\pi}bT + \pi + \pi_0(1-b)]D/[(IC + \bar{\pi}b)T], \tag{3a}$$

$$\text{if } T > [\pi + \pi_0(1-b)]/IC. \tag{3b}$$

We now substitute $X^*(T)$ into model (2b), which yields the following expression in T:

$$C[X^*(T), T] = \{A - D[\pi + \pi_0(1-b)]^2/2(\bar{\pi}b + IC)\}T^{-1} + \bar{\pi}bD[IC/(\bar{\pi}b + IC)]T/2 + [\pi + \pi_0(1-b)][IC/(\bar{\pi}b + IC)]D. \tag{4}$$

Model (4) is strictly convex in T. Thus the optimal inventory-review-cycle length T^* is easily obtained from the above expression and is given below:

$$T^* = \left\{[2A(\bar{\pi}b + IC) - [\pi + \pi_0(1-b)]^2 D\}/\bar{\pi}bICD\right\}^{1/2}. \tag{5}$$

We now obtain the minimum cost rate solution $C(X^*, T^*)$ by substituting T^* for T in model (4),

$$C(X^*, T^*) = \{\bar{\pi}bICD[2A(\bar{\pi}b + IC) - [\pi + \pi_0(1-b)]^2 D]\}^{1/2}/(\bar{\pi}b + IC) + [\pi + \pi_0(1-b)]\{IC/(\bar{\pi}b + IC)\}D. \tag{6}$$

If we now substitute T^* for T in inequality (3b), we get, after simplifying, the following inequality:

$$\left[\frac{2A}{ICD}\right]^{1/2} > \frac{\pi + \pi_0(1-b)}{IC} \tag{7}$$

The above inequality is equivalent to its counterpart ([1], p. 259), but here has been put into the more meaningful form.

The optimal policy can now be stated with the above inequality serving as a test criterion. Satisfaction of (7) indicates that a partial backlogging solution is optimal; otherwise, the solution of the classical lot-size model is optimal.

In words, the left side of inequality (7) is the classical infinite-time-horizon formula for the optimal interval between replenishments. Therefore, we can describe the optimal policy in a more informative way. Whenever the optimal inventory-review-cycle length computed by the formula derived for the classical lot size model exceeds the test criterion on the right side of (7), partial backlogging is optimal. Otherwise, the classical lot-size operating doctrine should be employed.

DISCUSSION

By examining in detail the test criterion on the right side of inequality (7) we will gain insight concerning the circumstances under which the optimal policy specifies partial backlogging. We begin by observing that the numerator of the test criterion ratio contains parameters

that appear nowhere else in the inequality. This implies that, for any given value of the left side, under suitable conditions, namely, the value of the numerator of the right side, the right side can exceed the left side.

The terms comprising the numerator have an economic characterization. They represent penalty costs incurred when a demanded item is not inventoried. Thus the economic content of inequality (7) is that, whenever the penalty for not inventorying demanded items is relatively small, then partial backlogging is a viable operating doctrine. In the numerical examples given below, this economic result is illustrated. We will demonstrate the fact that the value of the right side of (7) can be varied to the extent that the optimal operating doctrine will be changed by variation in the unit profit.

NUMERICAL EXAMPLES

We now introduce numerical cases to elucidate the concepts. Both cases have the following data in common:

$$D = 250, \quad C = 10, \quad I = 0.2,$$
$$A = 10, \quad \pi = 0.2, \quad \bar{\pi} = 0.1, \quad b = 0.5,$$

CASE 1, $\pi_0 = 0.2$:

$$[2(10)/0.2(10)(250)]^{1/2} = 0.2 > \frac{0.2 + 0.2(1 - 0.5)}{0.2(10)} = 0.15. \tag{7}$$

The above computation implies that the partial backlogging operating doctrine is optimal. The optimal replenishment interval is

$$T^* = \left\{[2(10)[0.1(0.5) + 0.2(10)] - [0.2 + 0.2(1 - 0.5)]^2(250)\}/0.1(0.5)(0.2)(10)(250)\right\}^{1/2} \tag{5}$$
$$= 0.86.$$

CASE 2, $\pi_0 = 2$:

$$0.2 < \frac{0.2 + 2(1 - 0.5)}{0.2(10)} = 0.6. \tag{7}$$

The above computation implies that the classical lot-size operating doctrine is optimal. The optimal replenishment interval is obviously 0.2.

REFERENCES

1. Montgomery, D.C., M.S. Bazaraa, and A.K. Keswani, "Inventory Models with a Mixture of Backorders and Lost Sales," Naval Research Logistics Quarterly, *20*, 255-63 (1973).
2. Rosenberg, D. "Monopoly Inventory Models," Doctoral Dissertation, New York University, New York, N.Y. (1977).

OPTIMAL BETTING STRATEGIES FOR FAVORABLE GAMES*

Eduardo J. Subelman

Department of System Science
University of California
Los Angeles, California

ABSTRACT

We examine the problem of a gambler interested in maximizing the expected value of a convex utility function of his fortune after n plays of a game. We allow any probability distribution to rule the outcome of each play, and this distribution may change from play to play according to a Markov process. We present results regarding the existence of an optimal policy and its structural dependence on the gambler's fortune. The well-known results of Bellman and Kalaba for exponential and logarithmic utility functions and coin-tossing games are generalized. We also examine the situation of general state spaces and show that the same structural results hold.

1. INTRODUCTION

Consider the following scenario: A gambler has a fortune x. He knows that the mechanism against which he is playing is in some state $i \in \{0, 1, 2, \ldots\}$. He can place any positive wager y not exceeding his fortune ($0 \leqslant y \leqslant x$). The mechanism will then be set in motion, and when it stops his fortune will be $x + \underset{\sim}{R}_i y$, where $\underset{\sim}{R}_i$ is a random variable with known distribution G_i. After this, the state of the gambling mechanism changes to j, with known probability P_{ij}, independent of $\underset{\sim}{R}_i$. The gambler is informed of the new state of the mechanism and is allowed to place a new wager, the process repeating itself a predetermined number of times, n.

We consider gamblers who are interested in maximizing the expected value of a utility function V of their fortune at the end of n plays. The problem is to determine the structure of the optimal wager.

Previous authors [1,2,4,7,9,10] have examined similar problems. They have, however, restricted themselves to coin tossing games (i.e., $\underset{\sim}{R}_i$ can only take the values ± 1) and utility functions that are either of the power type or logarithmic type. The notion of different states for the gambling mechanism has not, to our knowledge, been presented in the literature. The present paper will allow $\underset{\sim}{R}_i$ to be a general random variable, with the only conditions being $P[\underset{\sim}{R}_i < -1] = 0$ (i.e., the loss is limited to the wager) and, for some real number θ, $E\underset{\sim}{R}_i \leqslant \theta < \infty$.

*Partially supported by the U.S. Army Research Office under Grant DAAG29-77-0040 with the University of California. Reproduction in whole or in part is permitted for any purpose of the United States Government.

The utility function V will be allowed to be any function that is increasing, concave, and right continuous at the origin.

The past twenty years have been a rebirth of interest in gambling theory, which, after all, spawned probability theory (see Epstein [5]). Dubins and Savage's work [4] considers unfavorable gambling situations and the optimal strategies for these games. They prove, among other things, that if the gambler's goal is to reach a fortune of N before going broke and he is allowed even money wagers with probability $P < 1/2$ of success, then bold play (bet min $\{N - x, x\}$, where x is his fortune) is optimal. Ross [10] also considers coin-tossing games with $P < 1/2$ and shows that playing timidly (i.e., betting the minimum) maximizes the expected playing time until the gambler goes broke.

The first modern day author to consider coin tossing games with $P > 1/2$ is Kelly [7]. He shows that if the gambler wishes to maximize $\lim_{n\to\infty} \frac{1}{n} \log S_n$, where S_n is the gambler's fortune after n plays, and he is restricted to betting a constant proportion of his fortune at each toss of a coin, then he should bet the proportion $(2p - 1)$. Bellman and Kalaba [1], using Kelly's model, show that betting the proportion $(2p - 1)$ maximizes $E \log S_n$ for all n.

Breiman [2] generalizes Kelly's model and proves that, for games with $P > 1/2$, the gambling system which maximizes $E \log S_n$ asymptotically minimizes the expected time until $S_n \geqslant x$ (for sufficiently large x) and asymptotically maximizes $P[S_n > y]$ (for sufficiently large n, and every y). Brieman's work is the basis for a number of papers which advocate maximizing $E \log S_n$.

Thorp [12] gives a good overview of the various games available today which are favorable to the gambler. Pasternack [9] considers coin-tossing games but discounts money over time, and he finds conditions under which the gambler is better off by making risk-free investments.

In Section 2 of this paper we examine the general model of a gambler and determine the existence of an optimal policy as well as the structural properties of the optimal expected utility.

In Section 3, we extend Bellman and Kalaba's [1] results to our more general model.

Section 4 examines the dependence of the optimal wager on the gambler's fortune.

Section 5 indicates how the results can be extended to more-general state spaces.

Throughout this paper we use the term "increasing" to mean "nondecreasing" and use the explicit "strictly increasing" to mean exactly that.

2. FORMULATION OF MODEL

Let $V_n[x, i, \pi]$ be the expected utility when the player begins a sequence of n plays with a fortune of x, the (known) state of the gambling system is i, and strategy π is followed.

Define $V_n[x, i] = \sup_{\pi} V_n[x, i, \pi]$.

Following the usual dynamic programming formalism, set

$$V_n[x, i, y] = \sum_j P_{ij}\, EV_{n-1}\,[x + \underset{\sim}{R}_i,\ y,\ j]$$

and we obtain the functional equation

$$V_n[x, i] = \sup_{0 \leqslant y \leqslant x} V_n[x, i, y].$$

(The integrability of $V_n[x, i]$ will follow from Theorem 2.1 below.)

We now proceed to establish the properties of $V_n[x, i]$. We will require the following:

LEMMA 2.1: Let $f(x, y)$ be concave on $\{(x, y) \mid x \geqslant 0,\ 0 \leqslant y \leqslant x\}$. Let

$$g(x) = \sup_{0 \leqslant y \leqslant x} f(x, y). \tag{1}$$

Then

(a) $g(x)$ is concave in x, hence continuous in x;

(b) if $f(x, y)$ is increasing in x, then $g(x)$ is increasing in x;

(c) if $f(x, y)$ is differentiable in both x and y and for every $x > 0$ the maximizer of the right-hand side of (1), which we denote by $y(x)$, has $y(x) > 0$, then $g(x)$ is differentiable, with

$$g'(x) = f'_x[x, y(x)] + f'_y[x, y(x)]. \tag{2}$$

PROOF: The proofs of (a) and (b) are routine. To establish (c) it suffices to show that for every $\bar{x} > 0$ there exists only one α having

$$g(x) - \alpha x \leqslant g(\bar{x}) - \alpha \bar{x} \quad \text{for all } x \geqslant 0. \tag{3}$$

Let α satisfy (3) for fixed $\bar{x}$, and let $\bar{y} = y(\bar{x})$ and $\bar{u} = \bar{x} - \bar{y}$. For every $y \geqslant 0$ we get

$$\begin{aligned} f(\bar{u} + y, y) - \alpha(\bar{u} + y) &\leqslant g(\bar{u} + y) - \alpha(\bar{u} + y) \\ &\leqslant g(\bar{x}) - \alpha(\bar{u} + \bar{y}) = f(\bar{u} + \bar{y}, \bar{y}) - \alpha(\bar{u} + \bar{y}). \end{aligned}$$

Thus the differentiable function in y, $f(\bar{u} + y, y) - \alpha(\bar{u} + y)$, attains its maximum over $\{y \mid y \geqslant 0\}$ at $\bar{y}$. This assures that its derivative at $\bar{y}$ is zero, which proves (2).

The properties of $V_n[x, i]$ now follow by induction from Lemma 2.1 and the observation that, with $\psi = \max\{1, 1 + \theta\}$, $V_n(x, i) \leqslant V(x\psi^n)$.

THEOREM 2.1: For all i, $V_n[x, i]$ is increasing, concave, and continuous in x and increasing in n. In particular, this assures that, for every n, x, and i there exists an optimal wager $y_n(x, i)$ having $V_n[x, i, y_n(x, i)] = V_n[x, i]$. Also, if $V[x]$ is strictly increasing and differentiable, so is $V_n[x, i]$ for all n and i.

From this theorem, it follows that we can rewrite the optimality equation as

$$V_n[x, i] = \max_{0 \leqslant y \leqslant x} V_{n-1}[x, i, y]$$

$$V_0[x, i] = V[x].$$

Let $y_n[x, i]$ be the optimal wager, as determined from this equation. If it is not unique, we will allow $y_n(x, i)$ to be either the smallest or the largest of such values, although we will assume consistency in this choice.

From the properties of $V_n[x, i]$, one can easily establish

PROPOSITION 2.1: If $E\underset{\sim}{R}_i \leqslant 0$, then $y_n(x, i) = 0$.

PROOF: By Jensen's inequality

$$V_n[x, i] \leqslant \max_{0 \leqslant y \leqslant x} \sum_j P_{ij} V_{n-1}[x + yE\underset{\sim}{R}_i, j],$$

and the result follows, since $V_{n-1}[x, j]$ is increasing in x and $V_n[x, i] \geqslant \sum_j p_{ij} V_{n-1}[x, j]$.

This result, of course, demonstrates the reason for calling concave utility functions "risk averse." We will in the sequel assume, unless otherwise stated, that $E\underset{\sim}{R}_i > 0$ and consider this our definition of favorable games.

By imposing some structure on the random variables $\underset{\sim}{R}_i$, one can obtain monotonicity results for the dependence of $V_n(x, i)$ on i. Call the random variable $\underset{\sim}{R}_i$ *stochastically increasing* in i if $P(\underset{\sim}{R}_i > r)$ is increasing in i for all r. It is well known (e.g., [8]) that in this case $Ef(\underset{\sim}{R}_i)$ is increasing in i for every monotone function f. Also, call a stochastic matrix $P = [P_{ij}]$ *monotone* (e.g., [3,6]) if, for every k, $\sum_{j=k} P_{ij}$ is increasing in i. In this case, the rows of the matrix are the distributions of a stochastically increasing sequence of random variables. So $\sum_j P_{ij} a_j$ is increasing for every monotone sequence $\{a_j\}$. In particular, this implies that $\sum_j P_{ij} a_{ij}$ is increasing in i whenever $\{a_{ij}\}$ is increasing in both i and j.

From the above, one can prove by induction

PROPOSITION 2.2: Assume that $\underset{\sim}{R}_i$ is stochastically increasing in i and the matrix of transition probabilities is stochastically monotone. Then $V_n[x, i]$ is increasing in i.

The two simplest cases of stochastically monotone matrices are matrices having all rows coincide (corresponding to independent choice of the gamble at each stage) and the identity matrix (corresponding to the same game played repeatedly).

It is easy to find examples that show that this monotonicity result no longer holds if one merely assumes that the gambles are ordered stochastically and does not impose special conditions on the transition matrix.

Let $V[x] = x$. The 3 gambles available are coin tosses with probabilities of winning 0.6, 0.7 and 0.8 respectively. The transition matrix is

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

Clearly, all conditions except the monotonicity of $\mathbf{P}$ are satisfied, yet one can verify

$$V_n[x, 2] = (0.7)^n 2^n x,$$

$$V_n[x, 3] = 0.8\ (0.6)^{n-1} 2^n x,$$

and, for all $n \geqslant 2$, $V_n[x, 3] < V_n[x, 2]$.

3. SOME SPECIAL UTILITY FUNCTIONS

Among the many functions that fit our requirements for a utility function, some have received great attention in the literature. Among them are the logarithmic and power function, for which we can find simple forms for the $V_n[x, i]$ and $y_n[x, i]$.

THEOREM 3.1: If

(a) $V[x] = \log x$

or

(b) $V[x] = x^\beta/\beta$, $\beta \leq 1$, $\beta \neq 0$,

then we have respectively

(a) $V_n[x, i] = V_n[1, i] + \log x$

or

(b) $V_n[x, i] = V_n[1, i]\, x^\beta$,

and in both cases the optimal wager is given by $y_n(x, i) = \alpha(i)x$, where $\alpha(i)$ is determined by solving the problem

(a) $\max_{0 \leq \alpha \leq 1} E \log (1 + \alpha R_i)$

or

(b) $\max_{0 \leq \alpha \leq 1} E(1 + \alpha R_i)^\beta$, if $\beta > 0$,

$\min_{0 \leq \alpha \leq 1} E(1 + \alpha R_i)^\beta$, if $\beta < 0$,

and does not depend on n or x.

The proof in all cases is a simple induction on n.

These results generalize those of [1].

These utility functions have constant proportional risk aversion. It is known (e.g., [11]) that in this case the decision to accept or reject a proportional gamble is independent of the gambler's wealth. This explains why $y_n(x, i)$ is a scalar multiple of x, with the scalar being independent of n.

4. DEPENDENCE OF THE OPTIMAL STRATEGY ON THE FORTUNE

We now return to the general setting of Section 2 and investigate the dependence of the optimal wager $y_n(x, i)$ on the gambler's fortune x. From the results of the previous section, it is appealing to generalize and postulate that either

(a) $y_n(x, i)$ is increasing in x

or

(b) $\frac{1}{x} y_n(x, i)$ is monotone in x (increasing or decreasing).

Both of these properties do not necessarily hold, as shown by the example

$$V[x] = \begin{cases} x - x^2/2 & \text{for } 0 \leqslant x \leqslant 0.8, \\ 0.32 + 0.2x & \text{for } 0.8 < x. \end{cases}$$

The gamble is a coin toss: $P[\underset{\sim}{R} = 1] = 0.6$, $P[\underset{\sim}{R} = -1] = 0.4$. One can easily show that

$$y_1(x) = \begin{cases} x, & 0 \leqslant x \leqslant 1/6, \\ 0.2(1-x), & 4/6 \leqslant x \leqslant 3/4, \\ x - 0.7, & 3/4 \leqslant x, \end{cases}$$

which is clearly not monotone in x. One can also verify that $\frac{1}{x} y_1(x)$ is not monotone.

One can then ask if there exist simple functions of x and $y_n(x, i)$ which are monotone in x. The search for such functions is greatly assisted by Topkis' results on the minimization of submodular functions on a lattice [13].

Let $\underset{\sim}{R}_i$ be a random variable with $P[\underset{\sim}{R}_i \geqslant -1] = 1$. Define $\rho_i = \sup\{r > 0 | P[0 < \underset{\sim}{R}_i < r] = 0\}$. Note that $\rho_i > 0$ or $\rho_i = -\infty$.

THEOREM 4.1: Assume the state of the gambling system is i. Assume $\rho_i > 0$. Then $x + \rho_i y_n(x, i)$ is increasing in x.

PROOF: Recall

$$V_n[x, i] = \max_{0 \leqslant y \leqslant x} \sum_j P_{ij} EV_{n-1}[x + \underset{\sim}{R}_i\ y,\ j].$$

By making the substitution $v = x + \rho_i y$ we have

$$V_n[x, i] = -\min_v \left\{ -\sum_j P_{ij} EV_{n-1}\left[x\left(1 - \frac{\underset{\sim}{R}_i}{\rho_i}\right) + \frac{\underset{\sim}{R}_i}{\rho_i} v,\ j \right] \right\} \tag{4a}$$

$$\text{subject to } x \leqslant v \leqslant x(1 + \rho_i). \tag{4b}$$

Clearly this constraint set is a sublattice. Now consider

$$-V_{n-1}\left[x\left(1 - \frac{R}{\rho_i}\right) + \frac{R}{\rho_i} v,\ j \right].$$

For $R \leqslant 0$, $1 - \frac{R}{\rho_i} \geqslant 0$ and $\frac{R}{\rho_i} \leqslant 0$.

For $R > 0$, $\frac{R}{\rho_i} \geqslant 0$, and one need only consider $R \geqslant \rho_i$ (by definition of ρ_i), in which case $\frac{R}{\rho_i} \geqslant 1$. In both cases $-V_{n-1}\left[x\left(1 - \frac{R}{\rho_i}\right) + \frac{R}{\rho_i} v, j \right]$ is submodular. Since mixtures of submodular functions are also submodular, we find that

$$-\sum_j P_{ij} EV_{n-1}\left[x\left(1 - \frac{\underset{\sim}{R}_i}{\rho_i}\right) + \frac{\underset{\sim}{R}_i}{\rho_i} v,\ j \right]$$

is submodular on the sublattice of (4b). By Theorem 6.2 of [13] the value $v_n(x, i)$ that optimizes (4) is increasing in x. This completes the proof.

We note that if the optimizing value of y is not unique, then consistently choosing either the smallest or the largest such value will guarantee monotonicity.

The decision variable we have considered so far is $y_n(x, i)$, the amount to bet. An alternative is to decide how much not to bet. An example, to be presented below, shows that $u_n(x, i) = x - y_n(x, i)$ is not monotone in x. However, in the same spirit of Theorem 4.1 can establish the following result.

Let $\underset{\sim}{R}_i$ be a random variable with $P[\underset{\sim}{R}_i \geqslant -1] = 1$. Define $\xi_i = |\inf\{r > 0 | P[r < \underset{\sim}{R}_i < 0\}|$. Then we have $\xi_i \in (0, 1]$ or $\xi_i = \infty$.

THEOREM 4.2: Assume $\xi_i \in (0, 1]$. Then $x - \xi_i y_n(x, i)$ is increasing in x.

Proof: If we make the change $u = x - \xi_i y$ and noting that we need not consider those values R such that $-\xi_i < R < 0$, the proof is similar to that of Theorem 4.1.

The monotone functions of x defined above are tight, in the sense that if $\rho_i > 0$ ($\xi_i \leqslant 1$) and $\rho > \rho_i (\xi > \xi_i)$, then there exist problems for which $x + \rho y_n(x, i)$ $[x - \xi y_n(x, i)]$ is not increasing in x, as the following example shows:

$$V[x] = \begin{cases} 2x & x \leqslant 4, \\ 6 + \frac{1}{2}x, & 4 < x. \end{cases}$$

There is only one stage, and $P[\underset{\sim}{R} = 1] = \frac{1}{2}$ and $P\left[\underset{\sim}{R} = -\frac{1}{2}\right] = \frac{1}{2}$. Thus, $\rho_i = 1$ and $\xi_i = \frac{1}{2}$.

The optimal strategy can be easily computed

$$y_n(x, i) = \begin{cases} x, & 0 \leqslant x \leqslant 2, \\ 4 - x, & 2 \leqslant x \leqslant 4, \\ 2x - 8, & 4 \leqslant x \leqslant 8, \\ x, & 8 \leqslant x. \end{cases}$$

For $\rho > 1$ the function $x + \rho y_1(x, i)$ is clearly strictly increasing on [0,2] and strictly decreasing on (2,4]. Similarly, for $\xi > 1/2$, the function $x - \xi y_n(x, i)$ is strictly increasing on [0,4] and strictly decreasing on (4,8].

It is also interesting to note that the above results do not depend on the allowable values of x and y. Thus, even if we are constrained to integer fortunes and integer wagers, the results will still hold, since they are based on the isotonicity results.

The combination of Theorems 4.1 and 4.2 yields

COROLLARY 4.1: Assume $\rho_i > 0$ and $\xi_i \in (0, 1]$. Then, for $\epsilon > 0$

$$\frac{\epsilon}{\xi_i} \geqslant y_n(x + \epsilon, i) - y_n(x, i) \geqslant -\frac{\epsilon}{\rho_i}.$$

The importance of this result lies in the reduction of computational effort: If the solution for a fortune x is known, we can bound the possible optimal wager values for a player with fortune $x + \epsilon$.

In particular, for coin-tossing games ($\rho = \xi = 1$) we have

COROLLARY 4.2: For coin-tossing games

(a) $x + y_n(x, i)$ is increasing in x: "The more you have the more you strive for";

(b) $x - y_n(x, i)$ is increasing in x: "The richer you are the more you save";

(c) $|y_n(x + \epsilon, i) - y_n(x, i)| \leqslant \epsilon$.

From Corollary 4.1 one can also deduce immediately

COROLLARY 4.2: Assume $\rho_i > 0$, $\xi_i \in (0, 1]$. Then $y_n(x, i)$ is uniformly continuous in x.

EXTENSIONS TO MORE-GENERAL STATE SPACES

All of the preceding results can easily be extended to more-general state spaces, and, in fact, it is not necessary to require that the transition rule between states be independent of the outcome of the gamble. However, in this latter situation, Proposition 2.1 no longer holds, as the following example shows.

The first game is a coin toss with probability 0.4 of winning. Depending on whether we win or lose this toss, we will win or lose all successive games. The utility function is $V(x) = x$. One can easily see that, for $n = 2$, the optimal strategy is: bet x. If you win, bet $2x$, if you lose, bet nothing. Thus, although in the first toss the probability of winning is 0.4, one should wager all one's fortune.

REFERENCES

[1] Bellman, R., and R. Kalaba, "On the Role of Dynamic Programming in Statistical Communication Theory," IRE Transactions on Information Theory *IT-3*, 197-203 (1957).

[2] Breiman, L., "Optimal Gambling Systems for Favorable Games," *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. I, pp. 65-78 (University of California Press, Berkeley and Los Angeles, 1961).

[3] Daley, D. J., "Stochastically Monotone Markov Chains," Zeitschrift fur Wahrscheinlichkeits theorie und verwandte Gebiete *10*, 305-317.

[4] Dubins, L., and L. Savage, *How to Gamble if You Must* (McGraw-Hill, New York, 1965).

[5] Epstein, R., *The Theory of Gambling and Statistical Logic*, (Academic Press, New York, 1967).

[6] Keilson, J., and A. Kester, "Monotone Matrices and Monotone Markov Processes," Technical Report No. 79, Department of Statistics, Stanford University, Stanford, California (1974).

[7] Kelly, J. L., Jr., "A New Interpretation of Information Rate," Bell System Technical Journal *35*, 917-926, (1956).

[8] Lehman, E. L., "Ordered Families of Distributions," Annals of Mathematical Statistics *26*, 399-419, (1955).

[9] Pasternack, B., "Optimal Gambling and Investment Systems Under Discounting and Disbursement," ORC 74-1, Operations Research Center, University of California, Berkeley, California (1974).

[10] Ross, S., "Dynamic Programming and Gambling Models," Advances in Applied Probability *6*, 593-606, (1974).

[11] Rothblum, U. G., "Multivariate Constant Risk Posture," Journal of Economic Theory *10*, 309-332 (1975).

[12] Thorpe, E., "Optimal Gambling Systems for Favorable Games," Review of the International Statistical Institute *37*, 273-293 (1969).

[13] Topkis, D. M., "Minimizing a Submodular Function on a Lattice," Operations Research *26*, 305-321 (1978).

SOME SIMPLE VICTORY-PREDICTION CONDITIONS FOR LANCHESTER-TYPE COMBAT BETWEEN TWO HOMOGENEOUS FORCES WITH SUPPORTING FIRES*

James G. Taylor

Department of Operations Research
Naval Postgraduate School
Monterey, California

ABSTRACT

This paper develops new "simple" victory-prediction conditions for a linear Lanchester-type model of combat between two homogeneous forces with superimposed effects of supporting fires not subject to attrition. These simple victory-prediction conditions involve only the initial conditions of battle and certain assumptions about the nature of temporal variations in the attrition-rate coefficients. They are developed for a fixed-force-ratio-breakpoint battle by studying the force-ratio equation for the linear combat model. An important consideration is shown to be required for developing such simple victory-prediction conditions: victory is not guaranteed in a fixed-force-ratio-breakpoint battle even when the force ratio is always changing to the advantage of one of the combatants. One must specify additional conditions to hold for the cumulative fire effectivenesses of the primary weapon systems in order to develop correct victory-prediction conditions. The inadequacy of previous victory-prediction results is explained by examining (for the linear combat model without the supporting fires) new "exact" victory-prediction conditions, which show that even the range of possible battle outcomes may be significantly different for variable-coefficient and constant-coefficients models.

1. INTRODUCTION

Even though combat between two military forces is a complex random process (see Note 1 on p. 65 of Taylor and Brown [21]), as a consequence of F.W. Lanchester's [10] pioneering 1914 work, from about the end of World War II military operations analysts have used simplified deterministic differential-equation models to develop insights into the dynamics of combat [1-3, 25-27]. Today, Lanchester-type complex system models, which rely on modern digital-computer technology for their implementation (see, for example, Bonder and Honig [3]), have been developed for various levels of combat, from combat between battalion-sized units [4] to theater-level operations [5,7] ([20] for further references). Nevertheless, a simple combat model may yield an understanding of important relations that are difficult to perceive in a more complex model, and such insights can provide valuable guidance for higher-resolution computerized investigations (see Bonder and Farrell [2] and Weiss [27]). In this paper we will develop new victory-prediction conditions for several such simplified Lanchester-type models of

*This research was partially supported by the Office of Naval Research as part of the Foundation Research Program at the Naval Postgraduate School and partially by the U.S. Army Research Office, Durham, North Carolina with R&D Project No. 1L161102BH57-05 Math (funded under MIPR No. ARO 22-77).

combat between two homogeneous forces of primary weapon systems (such as infantry) [27] with superimposed effects of supporting fires [27] not subject to attrition (see Figure 1 of Taylor and Parry [24]) in order to obtain some insights into the dynamics of combat (among others, the tradeoff between quality and quantity of weapon systems). Such results are not only important in their own right but are also useful in the quantitative analysis of tactics (see, for example, [13, 15]).

It is important for the military operations analyst to have a clear understanding of how force-level and weapon-system-performance factors interact to determine the outcome of battle. In this paper we show that there are two types of battle-outcome-prediction conditions, "simple" ones and "exact" ones. In his well-known survey on the Lanchester theory of combat, Dolansky [6] suggested the development of outcome-predicting relations without solving in detail as one of several problems for future research. The work at hand is a step towards this problem's reslution (see also Taylor [16] and Taylor and Comstock [23]). Furthermore, work by Bonder and Farrell [2] and Taylor [14, 21] shows that in general the analytical solution by infinite series to variable-coefficient Lanchester-type equations is so complicated that it provides by itself little information about battle outcome. We show that a new consideration is required for developing simple battle-outcome-prediction conditions from the force-ratio equation. We then use this new approach to develop new simple victory-prediction conditions for a linear model of Lanchester-type combat with supporting fires. These results extend earlier work by Bach, Dolansky, and Stubbs [1] and Taylor and Parry [24]. Bach et al. [1] considered the constant-coefficient version of the linear model considered here. In [24] we studied the Riccati equation satisfied by the force ratio for this model with variable coefficients. One of our major results was the development of conditions on the battle's initial state that we thought were sufficient to guarantee victory in a fixed-force-ratio-breakpoint battle.

We show by means of counterexample that our earlier approach [24] of developing conditions that guarantee that the force ratio's rate of change always has the same sign (positive or negative but never zero) is inadequate to determine correct victory-prediction conditions. We then develop new, simple victory-prediction conditions for the linear model by showing that a fixed-force-ratio "breakpoint" is actually reached by the course of battle. These results are particularly significant [16,17,24] because they show that developing victory-prediction conditions from the force-ratio equation is much more difficult than we had initially thought (especially for time-dependent attrition-rate coefficients). By examining new "exact" force-annihilation-prediction conditions for the variable-coefficient linear model without supporting fires, we show that even the range of battle outcomes may be different for constant-coefficient and variable-coefficient models: there is a range of values for the initial force ratio (i.e. more than a single value) such that neither side will ever be annihilated if and only if the cumulative fire effectiveness of each side's weapon system is bounded. Finally, the significance of the model with supporting fires for understanding the dynamics of combat is discussed.

2. COMBAT MODELED BY VARIABLE-COEFFICIENT LANCHESTER-TYPE EQUATIONS OF MODERN WARFARE WITH SUPPORTING FIRES

We consider the following Lanchester-type equations with (nonnegative) time-dependent attrition-rate coefficients:

$$(1)\qquad \begin{cases} dx/dt = -a(t)y - \beta(t)x, & \text{with } x(0) = x_0, \\ dy/dt = -b(t)x - \alpha(t)y, & \text{with } y(0) = y_0, \end{cases}$$

The equations (1) are valid only for $x,y > 0$. The first, for example, becomes $dx/dt = 0$ for $x = 0$. Two situations that have been hypothesized to yield the above equations are (a) "aimed-fire" combat between two homogeneous forces with "operational" losses [1] and (b)

"aimed-fire" combat between two homogeneous (primary) forces with superimposed effects of supporting fires not subject to attrition [24] (see Figure 1). The prediction of the attrition-rate coefficients from weapon-system performance data is discussed in [2] (see also [24]; further references are given in [20]). The model is further discussed in Taylor and Parry [24].

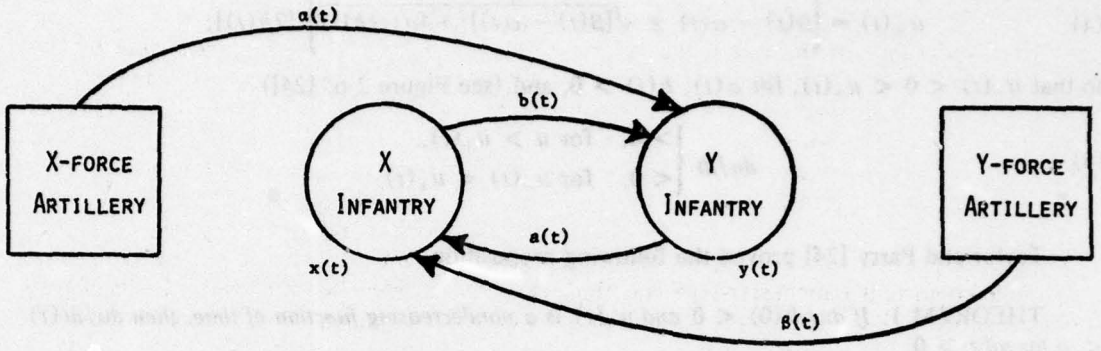


FIGURE 1. Combat between two homogeneous forces (infantry) with supporting weapons (artillery) not subject to attrition.

Taylor and Parry [24] noted that the force ratio, $u = x/y$, satisfies the Riccati equation

$$du/dt = b(t)u^2 + [\alpha(t) - \beta(t)]u - a(t), \text{ with } u(0) = u_0 = x_0/y_0, \tag{2}$$

and used this fact to develop much useful information about the behavior and implications of the model (1). Before proceeding further with analysis, we must specify battle-termination conditions for our model.

As Weiss [25] has emphasized, engagements that continue until one side is wiped out are rare. Although we are well aware that battle termination is a complex random process for which it is by no means certain that force levels are the only significant variables, we assume that combat ends when either of the two given breakpoint force ratios is reached. As pointed out by Taylor and Parry [24], the entire subject of modeling battle termination is a problem are in contemporary defense-planning studies. There is far from universal agreement on this topic (see Taylor [13] for further references). These breakpoint force ratios, denotes as u_X^f when X wins and u_Y^f when Y wins, satisfy $0 \leqslant u_Y^f < u_0 < u_X^f \leqslant +\infty$. See, for instance, Farrell and Freedman [8] for an example of the use of such battle-termination conditions in contemporary defense analysis. Corresponding to a fight to the finish, i.e. a battle until the annihilation of one side or the other, is the case in which $u_Y^f = 0$ and $u_X^f = +\infty$.

From the force-ratio equation (2) Taylor and Parry [24] developed a "local" condition of force superiority, e.g. "Y is winning" when

$$b(t)\, x^2(t) + [\alpha(t) - \beta(t)]\, x(t)\, y(t) < a(t)\, y^2(t), \tag{3}$$

and they sought to develop "global" conditions for winning a fixed-force-ratio-breakpoint battle (i.e. conditions sufficient to guarantee victory) by suitably strengthening hypotheses. Unfortunately, there was a flaw of a rather fundamental nature in our arguments [24] (and also subsequent ones [16]), as the counterexample given in the next section shows. Furthermore, subsequent research has shown that even the range of possible battle outcomes may be significantly different for models with variable (i.e. time-dependent) attrition-rate coefficients and those with constant ones (see Section 5). Thus, our difficulties were of a fundamental nature that one will encounter in general for variable-coefficient Lanchester-type combat models.

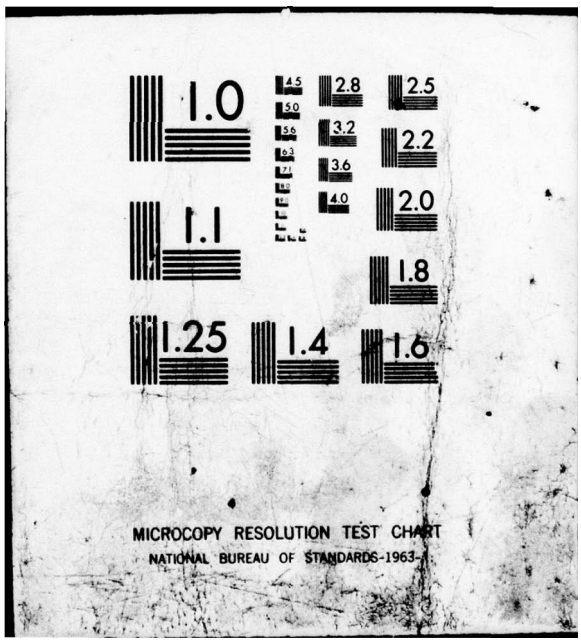
AD-A070 635 OFFICE OF NAVAL RESEARCH ARLINGTON VA F/G 15/5
NAVAL RESEARCH LOGISTICS QUARTERLY. VOLUME 26, NUMBER 2.(U)
JUN 79

UNCLASSIFIED

NL

3 OF 3
AD
A070635

END
DATE
FILMED
8-79
DDC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
6.3
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

Thus, the purpose of this paper is to give a new consideration that is required, in general, for victory prediction developed via the force-ratio equation and to use this new theoretical framework to develop correct victory-prediction conditions for the model (1). To this end, let $u_+(t)$ and $u_-(t)$ denote the roots of $b(t)u^2 + [\alpha(t) - \beta(t)]u - a(t) = 0$. It follows that, for $b(t) > 0$,

$$u_\pm(t) = \left\{\beta(t) - \alpha(t) \pm \sqrt{[\beta(t)-\alpha(t)]^2+4a(t)b(t)}\right\}/[2b(t)], \tag{4}$$

so that $u_-(t) < 0 < u_+(t)$, for $a(t)$, $b(t) > 0$, and (see Figure 2 of [24])

$$du/dt \begin{cases} > 0, & \text{for } u > u_+(t), \\ < 0, & \text{for } u_-(t) < u < u_+(t). \end{cases} \tag{5}$$

Taylor and Parry [24] proved the following proposition.

THEOREM 1: *If $du/dt(0) < 0$ and $u_+(t)$ is a nondecreasing function of time, then $du/dt(t) < 0$ for all $t > 0$.*

It is therefore of interest to know when $u_+(t)$ will be nondecreasing. Let

$$R(t) = a(t)/b(t), \quad \text{and } S(t) = [\beta(t) - \alpha(t)]\sqrt{a(t)b(t)}, \tag{6}$$

where $R(t)$ represents the relative fire effectiveness (Y to X) of the primary units, and $S(t)$ represents the net effectiveness of Y's supporting units, normalized by the "intensity" of combat between the primary units. Observing that $u_+(t) = \sqrt{R(t)}\,\{[S(t)/2] + \sqrt{[S(t)/2]^2+1}\}$, Taylor and Parry [24] also proved

THEOREM 2: *If $R(t)$ and $S(t)$ are both nondecreasing functions of time, then $u_+(t)$ is nondecreasing function of time.*

However, the author incorrectly concluded [24] that the hypotheses of Theorem 1, i.e., $du/dt(0) < 0$ and $u_+(t)$ being nondecreasing, were sufficient to guarantee victory for Y. The following counterexample, moreover, shows that these conditions only guarantee that Y cannot lose, not that he will win.

3. AN INSTRUCTIVE COUNTEREXAMPLE

In this section we give a counterexample which shows that $du/dt(0) < 0$ and $u_+(t)$ nondecreasing are not sufficient to guarantee victory for Y in a fixed-force-ratio-breakpoint battle. We consider the case in which the supporting fires are absent and the relative fire effectiveness of the primary weapon systems is constant, i.e., $\alpha(t) = \beta(t) = 0$, for all $t \geqslant 0$, $a(t) = k_a h(t)$, and $b(t) = k_b h(t)$. Then, as observed by Farrell [2], Taylor [12], and others (see Section 3 of Taylor and Brown [21]),

$$x(t) = x_0 \cosh \theta(t) - y_0\sqrt{\lambda_R}\sinh \theta(t),$$

where $\lambda_R = k_a/k_b$ and $\theta(t) = \sqrt{k_a k_b}\int_0^t h(s)\,ds$.

We now show that $x_0/y_0 < \sqrt{\lambda_R}$ does not always imply that the X force will be annihilated if $\lim_{t\to+\infty} \theta(t) = M < +\infty$. For example, consider a fire fight in which the combatants take

"aimed-fire" combat between two homogeneous (primary) forces with superimposed effects of supporting fires not subject to attrition [24] (see Figure 1). The prediction of the attrition-rate coefficients from weapon-system performance data is discussed in [2] (see also [24]; further references are given in [20]). The model is further discussed in Taylor and Parry [24].

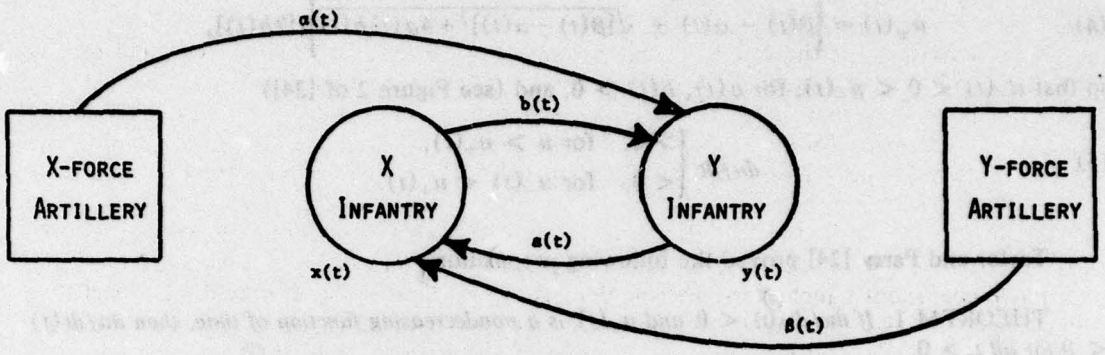


FIGURE 1. Combat between two homogeneous forces (infantry) with supporting weapons (artillery) not subject to attrition.

Taylor and Parry [24] noted that the force ratio, $u = x/y$, satisfies the Riccati equation

$$du/dt = b(t)u^2 + [\alpha(t) - \beta(t)]u - a(t), \text{ with } u(0) = u_0 = x_0/y_0, \tag{2}$$

and used this fact to develop much useful information about the behavior and implications of the model (1). Before proceeding further with analysis, we must specify battle-termination conditions for our model.

As Weiss [25] has emphasized, engagements that continue until one side is wiped out are rare. Although we are well aware that battle termination is a complex random process for which it is by no means certain that force levels are the only significant variables, we assume that combat ends when either of the two given breakpoint force ratios is reached. As pointed out by Taylor and Parry [24], the entire subject of modeling battle termination is a problem are in contemporary defense-planning studies. There is far from universal agreement on this topic (see Taylor [13] for further references). These breakpoint force ratios, denotes as u_X^f when X wins and u_Y^f when Y wins, satisfy $0 \leqslant u_Y^f < u_0 < u_X^f \leqslant +\infty$. See, for instance, Farrell and Freedman [8] for an example of the use of such battle-termination conditions in contemporary defense analysis. Corresponding to a fight to the finish, i.e. a battle until the annihilation of one side or the other, is the case in which $u_Y^f = 0$ and $u_X^f = +\infty$.

From the force-ratio equation (2) Taylor and Parry [24] developed a "local" condition of force superiority, e.g. "Y is winning" when

$$b(t)\, x^2(t) + [\alpha(t) - \beta(t)]\, x(t)\, y(t) < a(t)\, y^2(t), \tag{3}$$

and they sought to develop "global" conditions for winning a fixed-force-ratio-breakpoint battle (i.e. conditions sufficient to guarantee victory) by suitably strengthening hypotheses. Unfortunately, there was a flaw of a rather fundamental nature in our arguments [24] (and also subsequent ones [16]), as the counterexample given in the next section shows. Furthermore, subsequent research has shown that even the range of possible battle outcomes may be significantly different for models with variable (i.e. time-dependent) attrition-rate coefficients and those with constant ones (see Section 5). Thus, our difficulties were of a fundamental nature that one will encounter in general for variable-coefficient Lanchester-type combat models.

cover and continue to reduce their vulnerability, so that the fire effectiveness decays exponentially over time; i.e., $a(t) = k_a e^{-\gamma t}$, and $b(t) = k_b e^{-\gamma t}$. For this example, $M = \sqrt{k_a k_b}/\gamma$. Consequently, even when $x_0/y_0 < \sqrt{\lambda_R}$, we can always choose γ so that $\lim_{t\to+\infty} x(t) = x_0 \cosh M - y_0\sqrt{\lambda_R} \sinh M > 0$. We observe that $R(t)$ $(= a(t)/b(t) = \lambda_R)$ and $S(t)$ $(= 0)$ are both nonincreasing (so that $u_+(t)$ is nonincreasing), but yet $x_0/y_0 < \sqrt{\lambda_R}$ does not imply that X will be annihilated or even lose a fixed-force-ratio-breakpoint battle. The latter follows, since

$$\lim_{t\to+\infty} u(t) = \sqrt{\lambda_R}\,[(u_0 + \sqrt{\lambda_R})e^{-2M} + u_0 - \sqrt{\lambda_R}]/[(u_o + \sqrt{\lambda_R})e^{-2M} - (u_0 - \sqrt{\lambda_R})]$$

can be made to take on any value betwen 0 and u_0 by the appropriate choice of M (through our choice of γ).

Thus, this counterexample shows that $R(t)$ and $S(t)$ nonincreasing so that $u_+(t)$ is *nonincreasing*) and $du/dt(0) < 0$ only imply that $du/dt(t) < 0$ for all $t \geq 0$. These conditions do not imply that X will lose a fixed-force-ratio-breakpoint battle in finite time. However, for this example if we additionally assume that $\lim_{t\to+\infty} \theta(t) = +\infty$, then X must lose the battle. In the next section we show how this additional assumption is extended to the general model (1) to yield conditions that are sufficient to guarantee an X loss.

4. SIMPLE VICTORY-PREDICTION CONDITIONS FOR A FIXED-FORCE-RATIO-BREAKPOINT BATTLE

For convenience, we introduce the notation that $b(t) \in L(0,T)$ means that $\int_0^T b(t)\,dt$ exists (and is given by a finite quantity). We assume that, for $t \geq 0$, $a(t)$ and $b(t)$ are continuous, except for a finite number of points in time. It follows that $b(t) \notin L(0,+\infty)$ means that $\lim_{T\to+\infty} \int_0^T b(t)\,dt = +\infty$. We then have

THEOREM 3: *Assume that* (A1) $R(t)$ *and* $S(t)$ *are nondecreasing functions of time,* (A2) $b(t) \notin L(0, +\infty)$, *and* (A3) $R(t)$ *is not identically equal to zero. If* $b_0x_0^2 + \alpha_0x_0y_0 < a_0y_o^2 + \beta_0x_0y_0$, *then* X *will lose any fixed-force-ratio-breakpoint battle in finite time.*

PROOF: We denote $a(0)$ as a_0, etc. Then $R(t)$ and $S(t)$ nondecreasing implies that $u_+(t)$ is nondecreasing by Theorem 2. The initial-condition inequality $b_0\,x_0^2 + \alpha_0\,x_0\,y_0 < a_0\,y_0^2 + \beta_0\,x_0\,y_0$ implies that $du/dt(0) < 0$, so that Theorem 1 tells us that $du/dt(t) < 0$ for all $t \geq 0$. It remains to show that $u(t) \to u_f < u_0$ in finite time. The latter result may be proven by showing that $u(t) \leq u_0 - K_1 \int_{t_1}^t b(s)\,ds$, where $K_1 > 0$, since $\lim_{t\to+\infty} \int_0^t b(s)\,ds = +\infty$. There are now two cases to be considered: (C1) $S(t) < 0$ for all $t \geq 0$, and (C2) there exists $t_1 \geq 0$ such that $R(t_1) > 0$ and $S(t_1) \geq 0$.

CASE (C1): $S(t) < 0$ for all $t \geq 0$. We observe that it is impossible to have $a_0 = 0$. Hence, $R(t) > 0$ for all $t \geq 0$ and $du/dt = b(t)R(t)\,\{u^2/R(t) + [-S(t)/R^{1/2}(t)]u - 1\} \leq b(t)R_0\{u^2/R(t) + [-S(t)/R^{1/2}(t)]u - 1\} \leq b(t)R_0\{u_0^2/R_0 + [-S_0/R_0^{1/2}]u_0 - 1\} = [b(t)/b_0]du/dt(0)$. The first inequality follows from $R(t)$ being nonincreasing and $du/dt(t) < 0$ for all $t \geq 0$, while the second follows from $R(t)$ and $S(t)$ being nondecreasing and $u(t)$ nonincreasing. It follows that $u(t) \leq u_0 + (1/b_0)\,du/dt(0) \int_0^t b(s)\,ds$, and the theorem is proven in this case.

CASE (C2): There exists $t_1 \geq 0$ *such that* $R(t_1) > 0$ *and* $S(t_1) \geq 0$. We begin by observing that $u(t) \leq u_0 + \int_{t_1}^t (du/dt)\,dt$ for $t \geq t_1 \geq 0$. The minimum of du/dt considered

as a function of u occurs at $u^*(t) = R^{1/2}(t)\,S(t)/2$. Consequently, for $t \geqslant t_1$ we then have $du/dt(t,u=0) \leqslant du/dt(t,u) \leqslant du/dt(t,u^*)$ for $0 \leqslant u \leqslant u^*$, where $du/dt(t,u)$ denotes that we are considering du/dt to depend on the two indicated variables. Thus, for $t \geqslant t_1 \geqslant 0$ and $0 \leqslant u \leqslant R^{1/2}(t)\,S(t)/2$, we have $du/dt(t) \leqslant -a(t) \leqslant -b(t)\,R(t_1)$. Also, for $t \geqslant t_1 \leqslant 0$ and $0 \leqslant R^{1/2}(t)\,S(t)/2 \leqslant u < u_+(t)$, we have $du/dt = b(t)\,R(t)\,\{[u/R^{1/2}(t) - S(t)/2]^2 - [1 + (S(t)/2)^2]\} \leqslant b(t)\,R(t_1)\,\{[u/R^{1/2}(t) - S(t)/2]^2 - [1 + (S(t)/2)^2]\} \leqslant b(t)\,R(t_1)\,\{[u(t_1)/R^{1/2}(t_1) - S(t_1)/2]^2 - [1 + (S(t_1)/2)^2]\} = [b(t)/b(t_1)]\,du/dt(t_1)$. The first inequality on du/dt follows from $R(t)$ being nondecreasing and $du/dt < 0$. The second follows from $R(t)$ and $S(t)$ being nondecreasing, $u(t)$ nonincreasing, and the fact that $0 \leqslant u(t)/R^{1/2}(t) - S(t)/2$. Thus, we have shown that, for $t \geqslant t_1 \geqslant 0$,

$$du/dt(t) \leqslant \begin{cases} -b(t)\,R(t_1), & \text{for } 0 \leqslant u \leqslant R^{1/2}(t)S(t)/2, \\ -b(t)\,\{[-1/b(t_1)]\,du/dt(t_1)\}, & \text{for } 0 \leqslant R^{1/2}(t)S(t)/2 \leqslant u < u_+(t). \end{cases}$$

It follows that $u(t) \leqslant u_0 - K_1 \int_{t_1}^{t} b(s)\,ds$, where $K_1 =$ minimum $\{R(t_1), [-1/b(t_1)]\,du/dt(t_1)\} > 0$, and the theorem is proven in the second case. *Q.E.D.*

COMMENT 1: The victory-prediction inequality in Theorem 3 may be written as $u_0 < \sqrt{R_0}\,[(S_0 + \sqrt{S_0/2)^2 + 1}]$. Thus, instead of the six absolute quantities (i.e., two force levels and four attrition-rate coefficients), there are only three independent relative-capability parameters (one relative-primary-force-size parameter and two relative-fire-effectiveness parameters) involved in victory prediction: (1) the initial force ratio of primary systems, (2) the initial relative fire effectiveness of the primary weapon systems, and (3) the initial net fire effectiveness of the supporting weapons normalized by the intensity of combat between the primary weapon systems.

COMMENT 2: As we pointed out previously [23,24], when the supporting fires are always equally effective (i.e. $\alpha(t) = \beta(t)$), their effects "cancel out," and, in terms of the force ratio, the battle's outcome (although accelerated) is the same as though they were not present.

COMMENT 3: Let us introduce the "elapsed normalized battle time" (for combat between the primary weapon system) $\tau - \tau_0$, which is defined by

$$\tau - \tau_0 = \sqrt{a(t)b(t)}\;t = \int_0^t \sqrt{a(s)b(s)}\;ds\,, \tag{7}$$

where $\sqrt{a(t)b(t)} = (1/t)\int_0^t \sqrt{a(s)b(s)}\;ds$ denotes the average intensity of combat between the primary weapon systems. Then the assumptions that $R(t)$ is nondecreasing, $b(t) \notin L(0,+\infty)$, and $R(t)$ is not identically equal to zero yield that the elapsed normalized battle time (for combat between the primary weapon systems) grows without bound (i.e. $\tau - \tau_0 \to +\infty$ as $t \to +\infty$). The elapsed-normalized-battle-time parameter $= \int_0^t \sqrt{a(s)b(s)}\;ds$ was introduced by Taylor and Brown [22] in their study of variable-coefficient Lanchester-type equations of modern warefare (8). Its introduction sometimes significantly reduces the complexity of analytical results (see [22]).

In Theorem 3 the assumptions that (A2) $b(t) \notin L(0,+\infty)$ and (A3) $R(t)$ is not identically equal to zero insure that the battle will terminated in finite time. They mean physically that Y's fire effectiveness against X does not decay "too rapidly" over time, so that not only is the course of battle always moving toward a Y victory but, also, that victory is actually reached. We conjecture that, for such combat between two homogeneous forces, conditions such as (A2) and (A3) on the primary-weapon-system attrition-rate coefficients are *always* necessary to

insure that the battle will terminate in finite time. We observe that (A2) is automatically satisfied for constant attrition-rate coefficients. Moreover, for a fight to the finish with supporting fires absent, we know that the conjecture is true, as results in the next section show.

5. EXACT FORCE-ANNIHILATION-PREDICTION CONDITIONS FOR VARIABLE-COEFFICIENT LANCHESTER-TYPE EQUATIONS OF MODERN WARFARE

In this section we give results (see Theorem 4 below) that show how battle termination (i.e., the combat's reaching of the battle-termination conditions) is related to the boundedness of the primary systems' cumulative fire effectivenesses. Let us observe that (for the assumptions made below) a firer's (for example, an X unit's) cumulative fire effectiveness being bounded is equivalent to the integrability of his attrition-rate coefficient over the interval $[0, +\infty)$ (e.g., $b(t) \in L(0, +\infty)$). For a fight to the finish and Lanchester-type equations of modern warfare, with the supporting fires absent (i.e. $\alpha(t) = \beta(t) = 0$ for all $t \geqslant 0$), we show that if each primary system's cumulative fire effectiveness remains bounded, i.e., $a(t)$ and $b(t) \in L(0, +\infty)$, then neither side need ever be annihilated. (It follows by the Cauchy-Schwarz inequality for integrals that $a(t)$ and $b(t) \in L(0, +\infty)$ implies that $\lim_{t \to +\infty} \int_0^t \sqrt{a(s)b(s)}\, ds < +\infty$. Thus, the elapsed normalized battle time $\tau - \tau_0$, as given by (7), remains bounded.) This result explains what happened in the counterexample given in Section 3. Although Theorem 4 may be simply stated, its proof is fairly lengthy [18] and will not be included in this paper. Moreover, it is not developed from the force-ratio equation (2).

Accordingly, we consider combat modelled by

$$dx/dt = -a(t)y \qquad \text{and} \qquad dy/dt = -b(t)x\,, \tag{8}$$

where the battle begins at $t=0$, and we assume that $a(t)$ and $b(t)$ are defined, positive, and continuous for $t_0 < t < +\infty$, with $t_0 \leqslant 0$. We also assume that $a(t), b(t) \in L(t_0, T)$ for any finite T. We further take the attrition-rate coefficients $a(t)$ and $b(t)$ to be given in the form

$$a(t) = k_a g(t) \qquad \text{and} \qquad b(t) = k_b h(t)\,, \tag{9}$$

where k_a and k_b are positive constants. In other words, we assume that previous analysis has determined the attrition-rate coefficients to be of the above form (9). This general form has been suggested by the various specific attrition-rate functional forms that have appeared in the literature (e.g., see [14, 19, 21, 23]), and all the attrition-rate-coefficient examples known to this author are of this form. The reader is directed to Taylor [19] for a discussion of how the parameters k_a and k_b may in turn be related to weapon-system-capability and engagement-characteristics parameters.

We observe that $a(t)/b(t) = k_a/k_b$ in the special case in which $g(t) \equiv h(t)$. In other words, k_a and k_b are basically "scale factors" which are useful for the parametric study of battle outcomes. Motivated by the form of well-known constant-coefficient results (e.g. see [23]), we introduce the combat-intensity parameter λ_I and the relative-fire-effectiveness parameter λ_R defined by

$$\lambda_I = \sqrt{k_a k_b} \qquad \text{and} \qquad \lambda_R = k_a/k_b\,. \tag{10}$$

From our assumptions about $a(t)$ and $b(t)$, it follows that $a(t) \notin L(0, +\infty)$ means that $\lim_{t \to +\infty} \int_0^t a(s)\, ds = +\infty$.

We introduce the following hyperbolic-like general Lanchester functions [21, 22] (GLF) $C_X(t)$ and $S_X(t)$, which are the two linearly independent solutions to the *X-force-level equation*

(11) $$d^2x/dt^2 - \{[1/a(t)]da/dt\}dx/dt - a(t)b(t)x = 0 ,$$

with initial conditions

(12) $$C_X(t_0) = 1, \qquad S_X(t_0) = 0 ,$$
$$[1/a(t_0)]dC_X/dt(t_0) = 0, \qquad [1/a(t_0)]dS_X/dt(t_0) = 1/\sqrt{\lambda_R} ,$$

where t_0 denotes the largest finite time at which $a(t)$ or $b(t)$ ceases to be defined, positive, or continuous. We set $t_0 = 0$ if no such finite time exists. Analogous GLF for the corresponding Y force-level equations are similarly defined.

Let $F(Q) = [C_X(0) - QS_X(0)]/[QC_Y(0) - S_Y(0)]$. Then the following theorem [18] is an extension of Taylor and Comstock's [23] force-annihilation-prediction results.

THEOREM 4 (Taylor [18]): *The X force will be annihilated in finite time if and only if* $x_0/y_0 < \sqrt{\lambda_R}\, F(Q^*_{\max})$. *Neither side will be annihilated in finite time if and only if*

$$\sqrt{\lambda_R}\, F(Q^*_{\max}) \leqslant x_0/y_0 \leqslant \sqrt{\lambda_R}\, F(Q^*_{\min}) ,$$

where

$$\lim_{t \to +\infty} S_X(t)/C_X(t) = 1/Q^*_{\max}$$

and

$$\lim_{t \to +\infty} S_Y(t)/C_Y(t) = Q^*_{\min} .$$

We always have $Q^*_{\min} \leqslant Q^*_{\max}$. *Furthermore* $Q^*_{\min} < Q^*_{\max}$ *if and only if both* $a(t)$ *and* $b(t) \in L(0, +\infty)$.

We observe that for $t_0 = 0$ we have $F(Q) = 1/Q$. Some examples of the analytical determination of $Q^* = Q^*_{\max} = Q^*_{\min}$ are given in the paper by Taylor and Comstock [23], while examples of the prediction of force annihilation are given in Taylor and Brown [22].

Theorem 4 explains what was going on in the counterexample given in Section 3. For a fight to the finish, we find that $du/dt(t) < 0$ for all $t \geqslant 0$ when $x_0/y_0 < \sqrt{\lambda_R} = \sqrt{k_a/k_b}$. Since $C_X(t) = C_Y(t) = \cosh \theta(t)$, etc., it follows by Theorem 4 that neither side will be annihilated in finite time for $\sqrt{\lambda_R}\,(1 - e^{-2M})/(1 + e^{-2M}) \leqslant x_0/y_0 \leqslant \sqrt{\lambda_R}\,(1 + e^{-2M})/(1 - e^{-2M})$. Thus, there exist initial force ratios such that $du/dt(t) < 0$ always but yet X is never annihilated. Moreover, Theorem 4 tells us that this can happen for more than a single value of x_0/y_0 only when both $a(t)$ and $b(t) \in L(0, +\infty)$.

6. DISCUSSION

Although highly idealized, the model (1) is significant because of the insights provided into the dynamics of combat. We may consider (1) to model combat between two homogeneous forces (primary weapon systems) with superimposed effects of supporting fires. Lanchester [10] apparently believed in 1914 that the modern trend in warfare was toward greater concentration of forces (i.e., higher troop density) and formulated his now-classic model of combat (without supporting fires) in order to quantitatively justify the principle of concentration. It is significant to note (see [9]) that the actual trend in combat operations over the past two thousand years of military history has been towards greater dispersion of forces (i.e., lower troop density). Some figures for the last hundred years are shown in Table I (see Stewart [11]). Furthermore, the model (1) may be used to gain insights into whether or not it is "beneficial" to concentrate forces, i.e., whether or not a side should make its initial commitment of forces

TABLE 1. Increase in the Dispersion of Troops from the U.S. Civil War to World War II (from Stewart [11])

Item	Civil War	World War I	World War II
Area of 100,000 men (in square miles)	26.8	140	1727
Average frontage of 100,00 men (miles)	8.0	11	38.4
Average depth of 100,000 men (miles	3.3	13	45

as large as possible. Results show that, if the "intensity" of the supporting-fire combat exceeds that of the primary systems [i.e., $\alpha(t)\beta(t) > a(t)b(t)$], then the victor should not concentrate his forces. Actually, additional hypotheses are required. For simplicity, we have omitted them here. (See Taylor [17] for a detailed analysis of the decision to concentrate forces). Considering the past increases [9] in the fire effectiveness of supporting weapons relative to those for primary weapon systems (e.g., small arms), we would expect that, in general, $\alpha(t)\beta(t) > a(t)b(t)$ on the modern battlefield. Consequently, the victor should not concentrate his forces, according to the above. Thus, the model (1) yields a theoretical result that is in better agreement with the historical trend in military operations than is that yielded by Lanchester's original model without supporting fires (i.e., the victor should always concentrate forces).

A major contribution of this paper has been to show that in the development of victory-prediction conditions it must be proven that the battle-termination conditions will be actually reached. The counterexample in Section 3 showed that even though the force ratio is always becoming more favorable, for example, to Y, this condition does not guarantee that Y will win a fixed-force-ratio-breadpoint battle. This fact was not appreciated by us in our earlier work [24] and its subsequent extension [16]. In [16] we tried to develop general outcome-prediction conditions based on a comparison of the force ratio and the instantaneous force-change ratio (for cases of no replacements and withdrawals, the instantaneous casualty-exchange ratio). Unfortunately, this work contains the same type of flaw as that of Taylor and Parry [24] discussed in this paper: the conditions developed are sufficient to guarantee, for example, that the force ratio keeps on changing in Y's favor but not that he will win a fixed-force-ratio-breakpoint battle in finite time. Thus, the development of battle-outcome-prediction conditions is much more difficult for combat modelled by time-dependent attrition-rate coefficients than we had earlier thought. However, we showed that some of our earlier incorrect outcome-prediction results (i.e., those in Taylor and Parry [24]) could be corrected by the incorporation of additional simple attrition-rate-coefficient assumptions that yield that the elapsed normalized battle time (see Comment 2 above) grows without bound and consequently guarantee that Y will actually win. Unfortunately, our general outcome-prediction results involving the force ratio and the instantaneous force-change ratio [16]) cannot be so easily corrected.

In general, outcome-prediction conditions provide insights into the tradeoff between quality and quantity of weapon systems. In his classic paper Lanchester [10] assumed that the combatants' fire effectivenesses were constant over time and deduced his famous square law, which allows one to trade off quality versus quantity of weapon systems by means of the condition for equality of fighting strengths, $x_0/y_0 = \sqrt{a/b}$, where a and b denote constant attrition-rate coefficients. In this paper we have given both simple and exact outcome-prediction conditions which provide such tradeoff insights for Lanchester-type combat with supporting fires.

Our simple outcome-prediction conditions [see Theorem 3, which only involves (a) the initial force ratio of primary systems, (b) the initial relative fire effectiveness of the primary systems, and (c) the initial net fire effectiveness of the supporting weapons normalized by the intensity of combat between primary systems (see Comment 1 above)] are rather strong sufficient conditions. In other words, X may still lose when they are not satisfied. On the other hand, Theorem 4 gives necessary and sufficient conditions for force annihilation (and also nonannihilation of both sides). We have accordingly called these outcome-prediction conditions exact. So-called higher transcendental functions (e.g., GLF such as the LCS [21,22] functions), unfortunately, may be involved (e.g., for $t_0 < 0$ and $a(t)/b(t) \neq$ constant) in these exact force-annihilation-prediction conditions. By contrasting the complexity of these two types of conditions (i.e., the simple and the exact), we see the price in mathematical complexity that one has to pay for greater accuracy in outcome prediction.

REFERENCES

[1] Bach, R., L. Dolansky, and H. Stubbs, "Some Recent Contributions to the Lanchester Theory of Combat," Operations Research *10*, 314-326 (1962).

[2] Bonder, S., and R. Farrell, editors, "Development of Models for Defense Systems Planning," Report No. SRL 2147 TR 70-2 (U), Systems Research Laboratory, The University of Michigan, Ann Arbor, Michigan (September 1970).

[3] Bonder, S., and J. Honig, "An Analytic Model of Ground Combat: Design and Application," in *Proceedings U.S. Army Operations Research Symposium 10*, (1971), pp. 319-394.

[4] Bostwich, S., F. Brandi, C. Burnham, and J. Hurt, "The Interface Between DYNTACS-X and Bonder-IUA , in *Proceedings U.S. Army Operations Research Symposium 13*, (1974), pp. 494-502.

[5] Cordesman, A., editor, "Developments in Theater Level War Games," unpublished materials for C-5 Working Group of 35th Military Operations Research Symposium, 1975.

[6] Dolansky, L., "Present State of the Lanchester Theory of Combat," Operations Research *12*, 344-358 (1964).

[7] Farrell, R., "VECTOR 1 and BATTLE: Two Versions of a High-Resolution Ground and Air Theater Campaign Model," in *Military Strategy and Tactics*, R. Huber, L. Jones, and E. Reine, editors (Plenum Press, New York, 1975), pp. 233-241.

[8] Farrell, R., and R. Freedman, "Investigations of the Variation of Combat Model Prodictions with Terrain Line of Sight," Report No. AMSAA-1, FR75-1, Vector Research, Inc., Ann Arbor, Michigan (January 1975).

[9] Historical Evaluation and Research Organization, "The Fundamentals of Land Combat for Developing Computer Simulation Models of Ground and Air-Ground Warfare," unpublished seminar notes, Dunn Loring, Virginia (1976).

[10] Lanchester, F.W., "Aircraft in Warfare: The Dawn of the Fourth Arm—No. V., The Principle of Concentration," Engineering *98*, 422-423 (1914) (reprinted on pp. 2138-2148 of the *World of Mathematics, Vol. IV*, J. Newman editor, (Simon and Schuster, New York, 1956)).

[11] Stewart, W.G., "Interaction of Firepower, Mobility, and Dispersion," Military Review *40*, No. 3, 26-33 (1960).

[12] Taylor, J.G., "A Note on the Solution to Lanchester-Type Equations with Variable Coefficients," Operations Research *19*, 709-712 (1971).

[13] Taylor, J.G., "Survey on the Optimal Control of Lanchester-Type Attrition Processes," presented at the *Symposium on the State-of-the-Art of Mathematics in Combat Models*, June 1973 (also Tech. Report NPS55Tw74031, Naval Postgraduate School, Monterey, California, March 1974 (AD 778 630)).

[14] Taylor, J.G., "Solving Lanchester-Type Equations for 'Modern Warfare' with Variable Coefficients," Operations Research *22*, 756-770 (1974).

[15] Taylor, J.G., "On the Treatment of Force-Level Constraints in Time-Sequential Combat Problems," Naval Research Logistics Quarterly *22*, 617-650 (1975).

[16] Taylor, J.G., "On the Relationship Between the Force Ratio and the Instantaneous Casualty-Exchange Ratio for Some Lanchester-Type Models of Warfare," Naval Research Logistics Quarterly *23*, 345-352 (1976).

[17] Taylor, J.G., "Optimal Commitment of Forces in Some Lanchester-Type Combat Models," Operations Research *26*, 96-114 (1979).

[18] Taylor, J.G., "Prediction of Zero Points of Solutions to Lanchester-Type Differential Combat Equations for Modern Warfare," SIAM Journal on Applied Mathematics *36*, to appear in No. 3 in 1979.

[19] Taylor, J.G., "Recent Developments in the Lanchester Theory of Combat," in *Operational Research '78, Proceedings of the Eighth IFORS International Conference on Operational Research*, K.B. Haley, editor (North-Holland, Amsterdam, 1979), pp. 773-806.

[20] Taylor, J.G., "Attrition Modelling," in *Operationsanalystische Spiele in der Verteidigungsplanung*, R.K. Huber et al., editors , (Oldenbourg, München, 1979), pp. 139-189.

[21] Taylor, J.G., and G.G. Brown, "Canonical Methods in the Solution of Variable-Coefficient Lanchester-Type Equations of Modern Warfare," Operations Research *24*, 44-69 (1976).

[22] Taylor, J.G., and G.G. Brown, "Further Canonical Methods in the Solution of Variable-Coefficient Lanchester-Type Equations of Modern Warfare: A New Definition of Power Lanchester Functions," Tech. Report NPS55-77-27, Naval Postgraduate School, Monterey, California, June 1977 (AD A044 302).

[23] Taylor, J.G., and C. Comstock, "Force-Annihilation Conditions for Variable-Coefficient Lanchester-Type Equations of Modern Warfare," Naval Research Logistics Quarterly *24*, 349-371 (1977).

[24] Taylor, J.G., and S.H. Parry, "Force-Ratio Considerations for Some Lanchester-Type Models of Warfare," Operations Research *23*, 522-533 (1975).

[25] Weiss, H. K., "Requirements for a Theory of Combat," Memorandum Report No. 667, Ballistic Research Laboratories, Aberdeen Proving Ground, Maryland (April 1953).

[26] Weiss, H.K., "Lanchester-Type Models of Warfare," in *Proceedings of the First International Conference on Operational Research*, M. Davies, R.T. Eddison, and T. Page, editors (Operations Research Society of America, Baltimore, 1957), pp. 82-98.

[27] Weiss, H.K., "Some Differential Games of Tactical Interest and the Value of a Supporting Weapon System," Operations Research *7*, 180-196 (1959).

☆ U. S. GOVERNMENT PRINTING OFFICE: 1979 — 281-491/3

INFORMATION FOR CONTRIBUTORS

The NAVAL RESEARCH LOGISTICS QUARTERLY is devoted to the dissemination of scientific information in logistics and will publish research and expository papers, including those in certain areas of mathematics, statistics, and economics, relevant to the over-all effort to improve the efficiency and effectiveness of logistics operations.

Manuscripts and other items for publication should be sent to The Managing Editor, NAVAL RESEARCH LOGISTICS QUARTERLY, Office of Naval Research, Arlington, Va. 22217. Each manuscript which is considered to be suitable material for the QUARTERLY is sent to one or more referees.

Manuscripts submitted for publication should be typewritten, double-spaced, and the author should retain a copy. Refereeing may be expedited if an extra copy of the manuscript is submitted with the original.

A short abstract (not over 400 words) should accompany each manuscript. This will appear at the head of the published paper in the QUARTERLY.

There is no authorization for compensation to authors for papers which have been accepted for publication. Authors will receive 250 reprints of their published papers.

Readers are invited to submit to the Managing Editor items of general interest in the field of logistics, for possible publication in the NEWS AND MEMORANDA or NOTES sections of the QUARTERLY.