SENSITIVITY ANALYSIS OF QUEUING MODELS

By

ALI MOHAMED A. ALI

Bachelor of Science
Alexandria University
Alexandria, U.A.R.
1961

Master of Science
Oklahoma State University
Stillwater, Oklahoma
1965

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May, 1968

SENSITIVITY ANALYSIS OF

QUEUING MODELS

Thesis Approved:

James E Shamblin
Thesis Adviser

W. J. Bentley

Earl J. Ferguson

Rollo E. Venn

David E Rea

N Durham
Dean of the Graduate College

PREFACE

Mangement problems commonly arise from business systems in which people, machines, or materials form waiting lines for some type of servicing or processing. Waiting line (queuing) systems create a major class of decision-making problems in the management of business activities. Queuing system problems occur whenever a flow of arriving traffic consisting of elements (people or things) establishes a variable demand for service at facilities of limited service capacity. Waiting time, or delay (the time lapse between arrival and service to elements) varies inversely with the level of service capacity; i.e., the number of service stations (servers) and the rate of servicing maintained by each server. Management's objective in these problems is to select a "best" of alternative system operating schemes; that is, one which maintains an economic balance between waiting time and service capacity. Meeting this objective for any queuing system requires a practical and effective analytical means of solution; i.e., of predicting delays produced at specific arrival and service capacity levels. As one would expect, difficulty in determining such a means of solution increases with the scope and complexity of various properties which distinguish different types of waiting line systems.

Analytical optimization and sensitivity analysis for some situations in queuing theory is the primary objective of this dissertation. Basic classification of queuing theory is given here as a secondary objective. The concepts for the subject research developed through discussions with Dr. Shamblin, Associate Professor of Industrial Engineering. A review of the literature did not reveal any endeavors in the specific area for the sensitivity analysis in queuing theory.

My interest in queuing theory began in 1965 as a student of Dr. James E. Shamblin at Oklahoma State University. Interest in the area continued to grow through my association with him. The research resulting in this dissertation was supported by a scholarship from Ein-Shams University, U.A.R. Indebtedness is acknowledged to U.A.R. Government for the years of financial support it provided.

The members of my Advisory Committee: Professors W. J. Bentley, J. E. Shamblin, E. J. Ferguson, D. E. Bee, and R. E. Venn, deserve special credit for guiding my doctoral program and this investigation. A debt of gratitude is acknowledged to Dr. P. E. Torgersen, for serving on my Advisory Committee, and Dr. W. J. Fabrycky, for his contribution in my doctoral program, during their stay at Oklahoma State University. Thanks is due each of them for their inspiration and encouragement.

I also express my gratitude to Miss Velda Davis for her neat typing of my treatise.

TABLE OF CONTENTS

Chapter | Page

LIST OF TABLES

LIST OF FIGURES

Figure		Page

CHAPTER I

INTRODUCTION

Queuing Theory

Queuing theory is a branch of applied mathematics utilizing concepts from the field of stochastic processes. It has been developed in an attempt to predict fluctuating demands from observational data and to enable an enterprise to provide an adequate service for its customers with tolerable waiting. However, the theory also basically improves understanding of a queuing situation, enabling better control. The theory provides one with predictions about waiting times, the number waiting at any time, the length of busy periods and so forth. These predictions help the manager of the enterprise anticipate situations and take appropriate measures to alleviate congestion. In addition, it makes both the manager and customer aware of a constant need for new ideas for simplifying the complications of industrial situations.

Such problems are characterized by a variable rate of arrival of some kind of unit requiring some kind of service and by a variable rate of completing the service required. Broad statements of policy covering such situations are typically without specific meaning. For example, "Our policy is to provide the best possible service to customers. We always employ enough sales clerks to eliminate any possibility of customers having to wait for the attention of a salesgirl." This statement is quite vague and subject to a variety of interpretations. With queuing theory, either analytical or simulated, policy can be quantified and everyone, thus, brought into precise agreement as to what the statement of policy means.

Rational solutions of the problem of how many servicing stations to provide requires minimizing the total costs of keeping calling units waiting for service plus the costs of providing service stations.

Framing a Queuing Problem

To solve queuing problems, management must take six sets of estimates or forecasts. Some lean heavily on past data; some are matters of policy. The six sets are:

1. Frequency distribution of service calls.
2. Costs of waiting.
3. Distribution of service times.
4. Cost of maintaining service stations.
5. Relations of service stations to demand.
6. Priority rule [9].

Defining Alternative Solutions

The types of alternative courses of action which may require economic evaluation can be classified as follows:

1. Changing the number of serving stations.
2. Reorganizing the service stations so that the servicing time is changed.
3. Changing queue discipline.
4. Changing service policy.

Solution to a Queuing Problem

The solution to a queuing problem involves a set of specific values of the decision variables which minimize the sum of the cost associated with the properties of a specific queuing system.

Optimal Queuing Policy

The set of decision rules which minimizes these costs are referred to as optimal queuing policy. Optimal decision policies are obtained by the use of models.

Models

A model is a mathematical representation of the queuing system's properties and interrelationships. In order to construct a model, the associated properties and resulting interrelationships must be specified or assumed.

Three Phases of Queuing Analysis

(1) The determination of the properties (specification of assumptions) of a

queuing system.

(2) Mathematical model formulation and manipulation for an optimal solution.

(3) An analysis or evaluation of the solution. This analysis should evaluate the indifferent range and the insignificant limits of a variable which give insignificant effect on the optimal system cost.

It should be noted that the last phase will be defined and referred to as <u>Sensitivity Analysis of the Queuing Model</u> and is the major topic of the research reported in this treatise. Certain key questions and issues arise in the analysis of queuing problems which are of interest and significance both to the analyst and to decision maker in his work. In addition, management needs to have a general knowledge of the significance of the additional mathematical sophistication of the various models.

Thus, an essential step in any queuing analysis is the determination of how far one needs to go in using the variable in a certain model and when one should change the real situation to the optimal one. A thorough sensitivity analysis of the queuing model is necessary to help resolve this question.

An important, yet often overlooked, property of any decision model is the sensitivity to changes in parameter values. If one has constructed a model that appears to give reasonable, reliable results in trial applications and that is relatively economical to solve and if the variation in parameter values can be evaluated, a whole new dimension has been added to the decision-making process. With a model one may explore various possible values of this parameter and observe the effect of parameter change on dimensions and resulting economic outcome, either cost or profit. Thus, one could establish the sensitivity of costs and decisions to change parameter estimates.

There have been considerable research results reported in the technical literature concerning the assumptions of properties of various queuing systems and the formulation of the resulting model. It is the intent of this research to review these results and evolve analytical optimization for the developed model. In addition, the sensitivity analysis is given for each model. General and specific classifications of the queuing theory are given in Chapter II. Since the sensitivity is the major part of this research, the following definition should be carefully noted.

Sensitivity

Sensitivity analysis evaluates the responsiveness of management decisions to various factors associated with such decisions. Sensitivity analysis can be used to evaluate the responsiveness of a model to changes in various controllable factors (parameters), and to evaluate the responsiveness of a model to various properties from which the model is derived. It can be used to measure the responsiveness of a model to non-optimal or incorrect values of decision variables. This information can be used to facilitate the appraisal of alternative courses of action. These measures of responsiveness can be obtained by a measurement of the change of the output of a model of a system based on a controlled change to an input to that model of a system.

The whole concept of queuing analysis is based on a conceptualized mathematical model of the queuing system's properties. Thus, as in using any mathematical model to represent a physical system as an aid in the decision process, it is of interest to know:

(1) The sensitivity of the model to the use of "incorrect" (non-optimal) values of decision variables.

(2) The sensitivity of the model to the use of "incorrect" estimates of the input parameters.

(3) The sensitivity of the model to the use of "incorrect" properties of the system that define the model.

In regard to the sensitivity analysis of queuing models, the following questions are posed:

(1) What is the effect of an error in the

decision variable on the total system cost (TC)?

(2) What are the insignificant limits of the parameter (or the variable) which give insignificant effect on the total system cost?

(3) What is the indifferent range of the parameter (or the variable) which gives insignificant effect on the total system cost?

These questions arise in the industrial environment for several reasons; i.e., (a) a production control manager who can control the properties of the actual queuing system may want to know whether it is worthwhile to change the properties, (b) the optimal decision rules cannot be employed, or their employment may be more costly than some alternative rules, (c) the detailed estimation of the parameters may be too costly.

A measure of the sensitivity of a queuing system can be calculated in terms of the change in the total system cost of a queuing model. In this treatise, the only measure of sensitivity which will be utilized is the ratio of the difference between the actual and the optimal total system costs to the optimal total system costs. This sensitivity measure will be used in the calculation of the insignificant limits and the indifferent range of the parameter (or the variable) in use.

Literature Review

The original work in queuing theory was done by A. K. Erlang, a Danish telephone engineer. Erlang started his work in 1905 in an attempt to determine the effect of fluctuating service demand (arrivals) on the utilization of automatic dial equipment. It has been only the end of World War II that work on queuing models has been extended to other kinds of problems [1]. Since that time, the field of queuing theory has been a fruitful area of research by economists, mathematicians, statisticians, and computer manufacturers. In the technical literature relative to the topic of queuing theory, attempts have been made to define such terms as the "queuing problems", "queuing systems", "waiting line models", "waiting line systems", and "delay phenomena". It can be readily concluded that there does not presently exist any commonly accepted terms or definitions by the researchers in this area. Churchman [5] classifies problems involving queue into two different types depending on their structure. The first type of problem involves arrivals which are randomly spaced and/or service time of random duration. This class of problems includes situations requiring either determination of the optimal number of service facilities or the optimal arrival rate (or times of arrival) or both. The second type of queuing problem is not concerned with either controlling the times of arrivals or the number of facilities, but rather is concerned with the order or sequence in which service is provided to available units by a series of service points. Kaufmann [6] showed the equations of state that are used in the calculation of a number of different average dimensions. These equations were first presented about thirty-five years ago by the Danish engineer Erlang, in connection with the problem of telephone communications. These are generalized by means of so-called birth and death equations which allow one to describe a great many station cases or permanent systems. Lajos Takács [7] developed the Theory of Queues book. The aim of this book is to give an introduction to the probabilistic treatment of mass servicing. It is dealt with different models which can be applied successfully to the theory of telephone traffic, airplane traffic, road traffic, storage, operation of dams, serving of customer, and others. His interest is chiefly in the time-dependent or transient behavior of these processes. The purpose of the Queues and Inventory text by Prabhu [8] is to give the similarities between the mathematical formalisms of queuing and inventory models which had been observed at a fairly early stage of their development. He has made no attempt, however, to establish a unified theory. He has treated each topic separately by unified methods, using modified notation, and the similarities between the results have been pointed out wherever they exist. Frederick [10, 11] wrote two articles. The purpose of the first is to discuss and illustrate how to use queuing theory to analyze a wide range of industrial problems, especially those involving discussions regarding the amount of capacity to provide. Cost models, an example, and a discussion of how to determine the relevant costs are presented to give detailed guidance on how to conduct such an analysis. Special consideration is given to the case where priorities are used in selecting members of the queue for service, and a survey of the available results for queuing models of this type is briefly presented. The second article gives the practical application of queuing theory to actual industrial problems. The many examples presented illustrate the wide usage of problems that can be formulated as queuing models. A broad conceptual picture of the general approach to most of these problems is given, with emphasis on the underlying cost consideration. Arjan and Anand [12] described the basic structure of queuing problem, the authors developed the philosophy of queuing theory in terms of the components of the queuing system and its characteristics drawing upon examples from everyday life. A general discussion of approaches to the solution of queuing problems is also presented. In 1961, Saaty [4] developed a textbook in queuing theory. The principal purpose of this book is to produce a general text and a summary of scattered papers and monographs on the subject of queues. Secondarily, it includes a wide Bibliography and indicates some unsolved problems. It also includes a descriptive introductory chapter most of which is aimed at the layman. Many ideas are illustrated with examples, and a number of exercises are intended to fill in some of the omitted detail and develop results along indicated lines. In Operations Economy by Fabrycky and Torgersen [2], deterministic and probabilistic models for waiting line situations are presented as a means for achieving economic operation of queuing systems.

From the above, it can be easily seen that there has not been such a trial made in sensitivity analysis of queuing theory. None of the references previously cited consider the sensitivity analysis of queuing models. It is this author's opinion that this is perhaps one of the most significant aspects for decision models. Arne Mjosund [16] mentioned in his article of Operations Economy book review the following:

> The first two introductory chapters expose the reader to scientific approach in problem solving in industry and to the construction and manipulation of models. The ideas are well formulated and defined. One important part of model testing, sensitivity analysis, is, however, lacking.

In this treatise, sensitivity analysis of queuing models has been considered.

CHAPTER II

GENERAL CLASSIFICATION AND DESCRIPTION OF QUEUING THEORY

Basic Structure of Queuing Problems

A common phenomenon occurring in everyday life is that of "queuing" or waiting in a line. Queues (waiting lines) are formed at bus stops, supermarket counters, and ticket booths. Queues are also found in industry, in shops where the machines wait to be repaired, at a tool crib where mechanics wait to receive tools, in a warehouse where the parts wait to be used, and in a sales department where the incoming customer orders wait to be processed.

In general, a queue is formed when either units requiring services - commonly referred to as customers - wait for service, or the service facilities stand idle and wait for customers. Some customers wait when the total number of customers requiring service exceeds the number of service facilities; some service facilities stand idle when the total number of service facilities exceeds the number of customers requiring service.

Queuing theory can be applied to a wide variety of operational situations where this imperfect matching between the customers and service facilities is caused by ones inability to predict accurately the arrival and service times of customers. In particular, it can be used for determining suitable number and type of service facilities.

Queuing Process

Basically, a queuing process is centered around a service system which has one or more service facilities. Customers requiring service are granted at different times by an input source, commonly known as population. These customers arrive at the service system and may or may not enter the system depending upon the queue conditions. Any customer entering the service system joins a queue for service (a queue may be of zero length). The service facilities select customers for service by some rule, commonly known as service discipline. After the service is completed, the customer leaves the service system. The queuing process is illustrated in Figure 1.

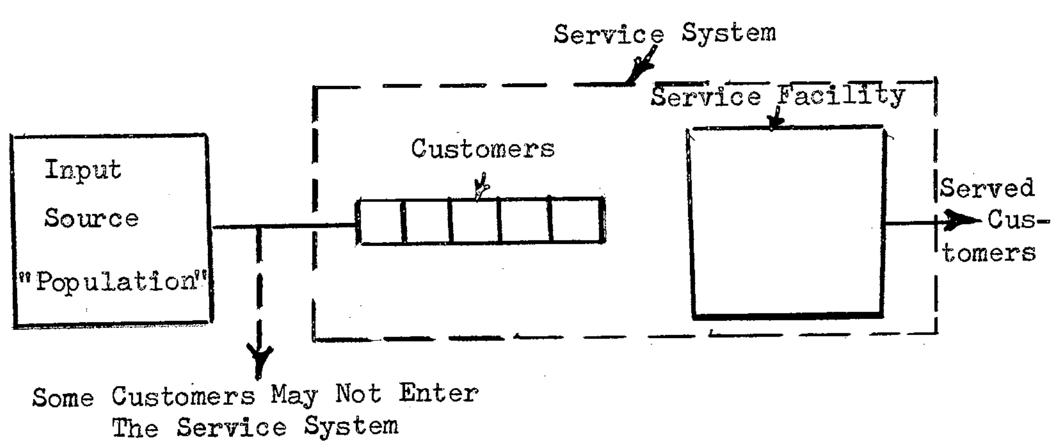


Figure 1. The Queuing Process

Consider the example of mechanics arriving at a tool crib to obtain tools. All the mechanics eligible for service, but excluding those at the crib, comprise the input source. The input source generates an input (a customer requiring service) in the form of a mechanic needing a tool. The mechanic then leaves the input source and arrives at the service system. In the example, he always enters the system, that is, he joins the queue irrespective of facilities. If a clerk is free, that is, if he is not serving another mechanic, the incoming mechanic receives service immediately. If all the clerks are busy, the incoming mechanic waits at the end of the queue (assuming that the clerk serves mechanics on a first-come, first-served basis). Any clerk becoming free serves the mechanic at the front of the queue. Thus, the mechanic joining the queue has to wait in the queue until all the mechanics in front of him are served. After a mechanic has received the required tools, he leaves the service system and joins the input source, thus becoming once more a potential customer.

Many practical situations can be put in the queuing framework. The elements of a queuing process, namely, the input source, queue, service facilities, and the service discipline are discussed in more detail.

Input Source (Population)

An input source is characterized by:

1. Its size
2. The arrival time distribution of the customers
3. The attitude of the customers.

<u>Size</u>

An input source (population) is infinite or finite. It is considered infinite if the rate at which the source generates the customers is not appreciably affected by the number of customers in the service system (customers in queue and those in service). Alternately, the source is considered finite if the rate is affected by the number of customers in the system. In practice, the population is considered infinite if the number of customers in the system is not likely to be appreciable fraction of the size of the population of potential customers.

Consider the example of a motel on a national highway. The total number of guests in the motel at any time is a very small fraction of the total population of potential customers (motorists driving on the highway). It may be said, therefore, that the customers arriving at the motel come from an infinite population. Now, consider the example of maintenance of machines by a repair crew. Here it is possible that an appreciable fraction of the machines will be out of order at any one time. Hence, the customers (machines) in this example are regarded as coming from a finite population.

Arrival Time Distribution

The periods between the arrival of individual customers may be constant or scattered in some fashion. In a clinic, patients may be given appointments in such a manner that they arrive at the clinic at specified equal intervals of time. On the other hand, the arrival times of customers in a restaurant are distributed more or less randomly and cannot be predicted. The arrival times can nevertheless be described. For example, customers arrive at a restaurant for service at intervals described in Table I.

TABLE I

ARRIVAL TIME DISTRIBUTION

Time Between Customer Arrivals	Percentage of Arrivals	Cumulative Percentage of Arrivals
0-4.99 minutes	42	42
5-9.99 minutes	23	65
10-14.99 minutes	18	83
15-19.99 minutes	11	94
20-24.99 minutes	4	98
25 minutes and over	2	100

This table represents what is called the arrival-time distribution. It shows, for instance, that a time period of 15 minutes or less precedes the arrival of 83 out of every 100 customers.

One could also obtain the average time between the customer arrivals and the average arrival rate (number of customers arriving per unit time) from the arrival-time distribution. Many of these distributions found in practice can be approximated by one of the well-known mathematical distributions, such as:

1. Constant time
2. Exponential
3. Erlang
4. Hyperexponential.

Exponential time distribution is a special case of Erlang as well as hyperexponential distributions; whereas, the constant time distribution is a special case of the Erlang distribution.

Arrival-time distributions concerning many practical situations such as failures of machines and arrivals of customers in restaurants are found to be exponential. However, there are some operational situations which have an arrival-time distribution appreciably different from exponential. Many of these nonexponential distributions can be approximated by Erlang and hyperexponential distributions.

It can be shown that the exponential arrival-time distribution gives rise to "Poisson" arrivals (and vice versa). In general, the arrivals will follow the Poisson distribution whenever the following assumption is satisfied.

The total number of arrivals during any given time interval is independent of the number of arrivals that have already occurred prior to the beginning of the time interval.

Attitude of Customers

If a customer, on joining the service system, does not get immediate service, the customer may:

1. Stay in the system until served.
2. Wait for certain time and leave the system if service is not commenced by that time.
3. Estimate the waiting time and then decide whether to leave.

The first type of customers who stay in the system (either voluntarily or involuntarily) until served, no matter how long they have to wait, are called "patient" customers. Machines moved to an internal maintenance shop in a plant for repairs are patient customers; the machines usually must be repaired whatever the waiting time.

The last two types of customers are classified as "impatient" customers, and an example is that of a customer arriving in a clothing store. If the salesmen are busy with other customers, a customer may wait for some time but then suddenly leave when his patience is exhausted, or he may estimate that his waiting time will be excessive so that he leaves immediately.

Queue

A queue refers to the customers waiting for service. This does not include the customers being served. Some operational situations allow a queue for any size to form; in others, the queue is characterized by its maximum permissible size, which may be infinite or finite.

In a sales department where the customer orders are received, there is no restriction on the number of orders that can come in so that a queue of any size can form. When there is no limit on its size, the permissible queue is said to be infinite.

In a gasoline station, the space for waiting of cars is usually limited. If a motorist arrives when all the space is occupied, he goes elsewhere for service. Thus, the maximum size of the queue is limited by the space for waiting. In many other situations, an incoming customer may not enter the service system if a certain number of customers are already waiting even though additional waiting space is available. Queue size in this case is controlled by the attitude of the customers. For instance, if a motorist needing gasoline finds all the spaces on both sides of the gas pumps occupied, he does not (in most cases, at least) stop at this gas station but goes to another. Because of this attitude of the the customers, the maximum queue size will equal the number of gas pumps.

When there is a limit on its size, the permissible queue is said to be finite.

In some finite queue systems, the maximum permissible queue is of zero length; that is, no queue is allowed to form. An example of such a situation is a parking lot. When all the service facilities (parking spaces) are busy, the incoming customers (motorists) do not wait but go elsewhere.

An interesting feature of the situations with finite queues is that if any customers arrive at times when the queue length is full; that is, the maximum permissible value, they do not enter the service system, and are, therefore, lost. The interest in queuing theory for such situations centers on the number of customers lost.

In some cases, there are more than one queue for one service station. An example of this, the case where an automatic machine has two input queues from previous different operations.

So whatever the queue, finite or infinite, it can be classified as single queue or multiple queue.

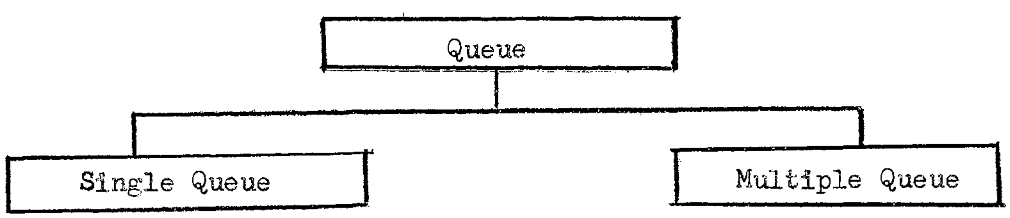


Figure 2. Classification of Queue

Service Facilities

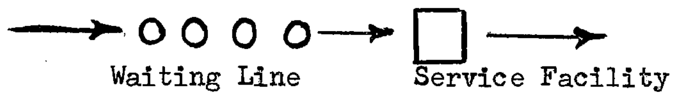
There are four basic structures of waiting line situations which describe the general conditions at the servicing facility. The simplest situation is where arriving units form a single line to be served by single processing facility, for example, a one-man barber shop. This is called the single-channel, single phase case. If the number of processing stations is increased (two or more barbers), but still draws on the one waiting line, it is called a multiple-channel, single phase case, since a customer can be served by any one of the barbers. A simple assembly line has a number of service facilities in series or tandom and is the single-channel, multiple phase case. The last one is the multiple-channel, multiple-phase case which might be illustrated by two or more parallel production lines. Figure 3 shows the four cases diagrammed and labeled.

Therefore, the service facilities are characterized by their arrangement, phases and channels. Figure 4 shows the schematic classification of the four cases.

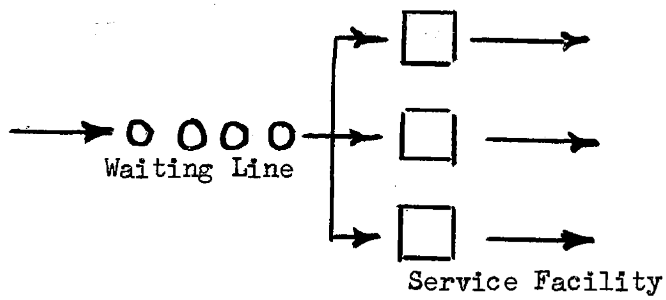
The service facilities have specific characteristics with respect to service-time distributions as it is explained below.

Service-Time Distributions

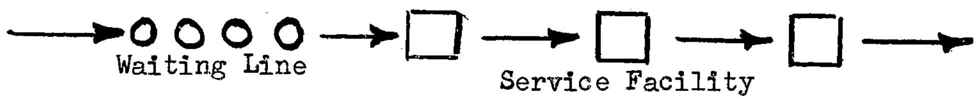
The time interval from the commencement of service to the completion of service for a customer is known as the



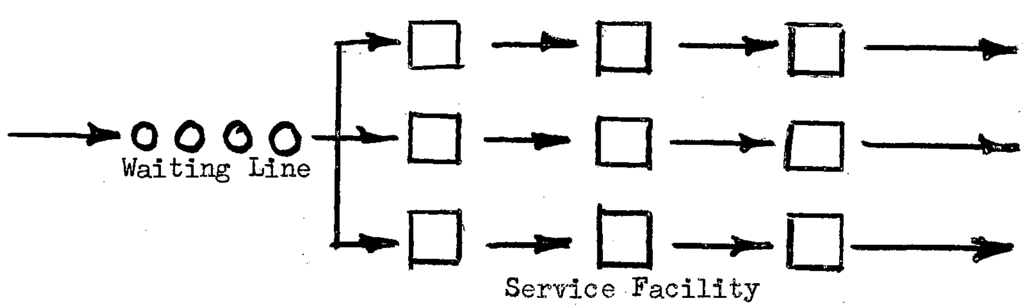
(a) Single Channel, Single Phase Case



(b) Multiple Channel, Single Phase Case



(c) Single Channel, Multiple Phase Case



(d) Multiple Channel, Multiple Phase Case

Figure 3. Four Basic Structures of Service Facility

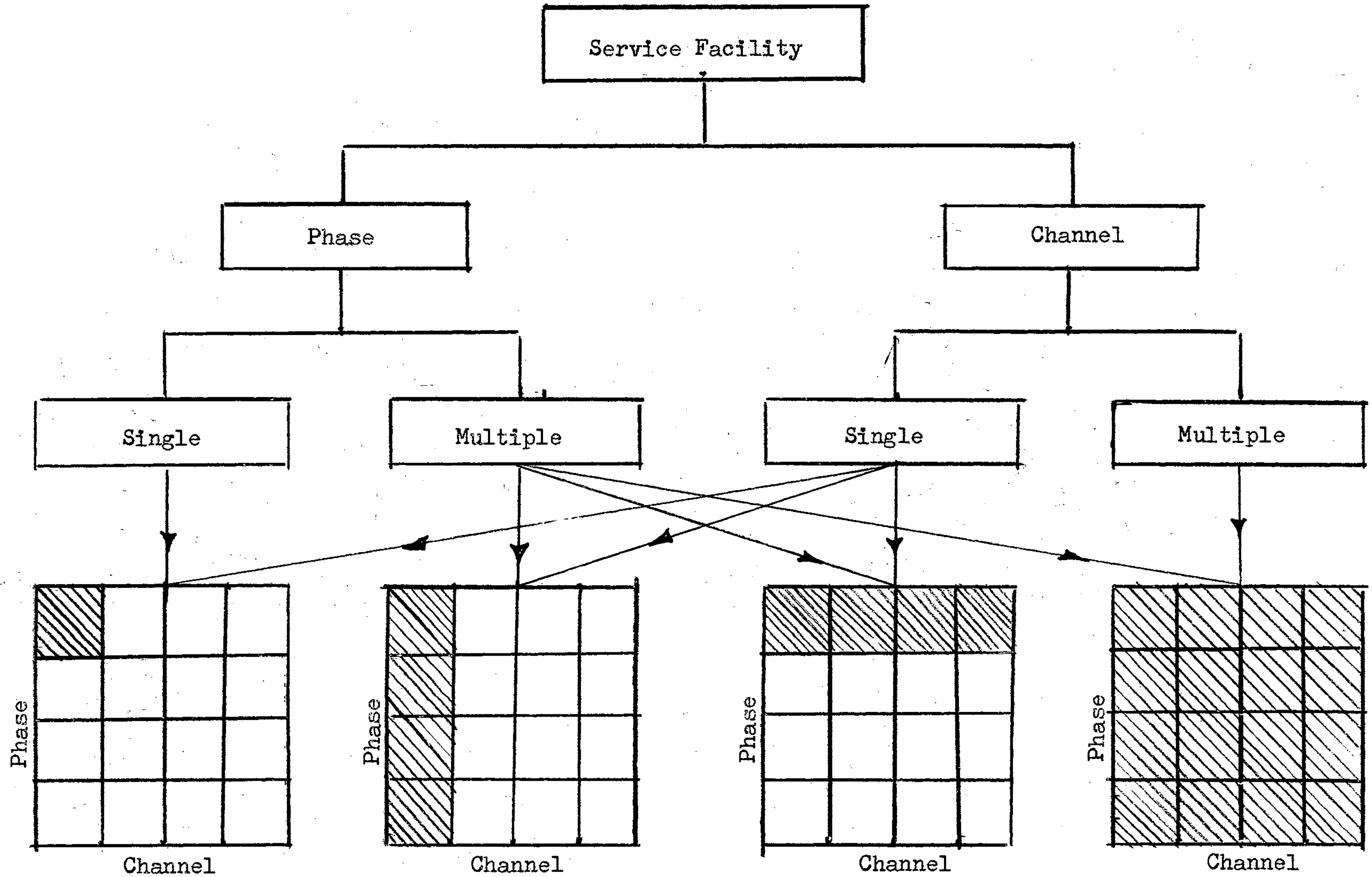


Figure 4. Classification of Service Facility

service time for the customer. The service times may be either constant or scattered in some fashion for different customers. The service-time distributions can be described in terms quite analogous to arrival time distributions. Thus, the service time distribution may be:

1. Constant time
2. Exponential
3. Erlang
4. Hyperexponential.

From the service-time distributions, one can obtain the average service time and the average service rate (services completed per unit time).

An example of constant service time is a nonstop subway train service between two stations in a city. Here, the arrival of customers (passengers) may be Poisson, but their service time (time spent in the trains while traveling between the two stations) is constant. Scattered service times are found in supermarkets where the girls at the cash register serve the customers, at the tool cribs where the clerks serve the mechanics, and in many other operational situations. Most of these have been found to follow the exponential service time distributions. However, there are some operational situations in which the service time distribution is appreciably different from exponential. Many of these nonexponential distributions can be approximated by Erlang and hyperexponential distributions.

In general, service times follow the exponential distribution whenever the chance of prolongation of service (for a customer) is independent of how long ago the service started.

Service Discipline

If any of the service facilities are free, the incoming customer is taken into service immediately. Should all the services be busy, the customers in the queue may be handled in a number of ways when a service facility becomes free; some of these being:

1. The customers are taken into service in order of their arrival. This is known as "first-come, first-served" service discipline. This may, for instance, be found at airports, where taxicabs queue while waiting for passengers. The taxicabs in this case are served (allotted passengers) on a "first-come, first-served" basis.
2. The customers are selected for service at random. This is known as the "random" service discipline. This is found in many operational situations where the customers do not wait in a well organized line.
3. The customers are assigned priorities. The service facility becoming free commences service on the customer with the highest

> priority. If there is more than one customer of the same priority in the queue, the service facility may select a customer from among these, either on the "first-come, first-served" or "random" basis. The processing of jobs on a computer in some industries is done. For example, on a priority basis with the debugging jobs having a higher priority than the regular checked out jobs.

A variation of the service discipline involving priorities is found in operational situations where the service for a customer with a lower priority is interrupted and the service for the incoming unit with higher priority commenced. This is known as the "preemitive priority" service discipline.

Classification of Queuing Systems

In summary, a queuing theory can be classified by source input, queue, and service facility. Figure 5 gives a clear classification for the whole possible cases in queuing theory. Queuing theory is classified with respect to the source input as finite or infinite. That is, the population which the customers call for service is either finite or infinite in number. It is classified with respect to queue as either the service system has a single queue or a multiple queue. The service facility is classified with respect to the channels and phases, either

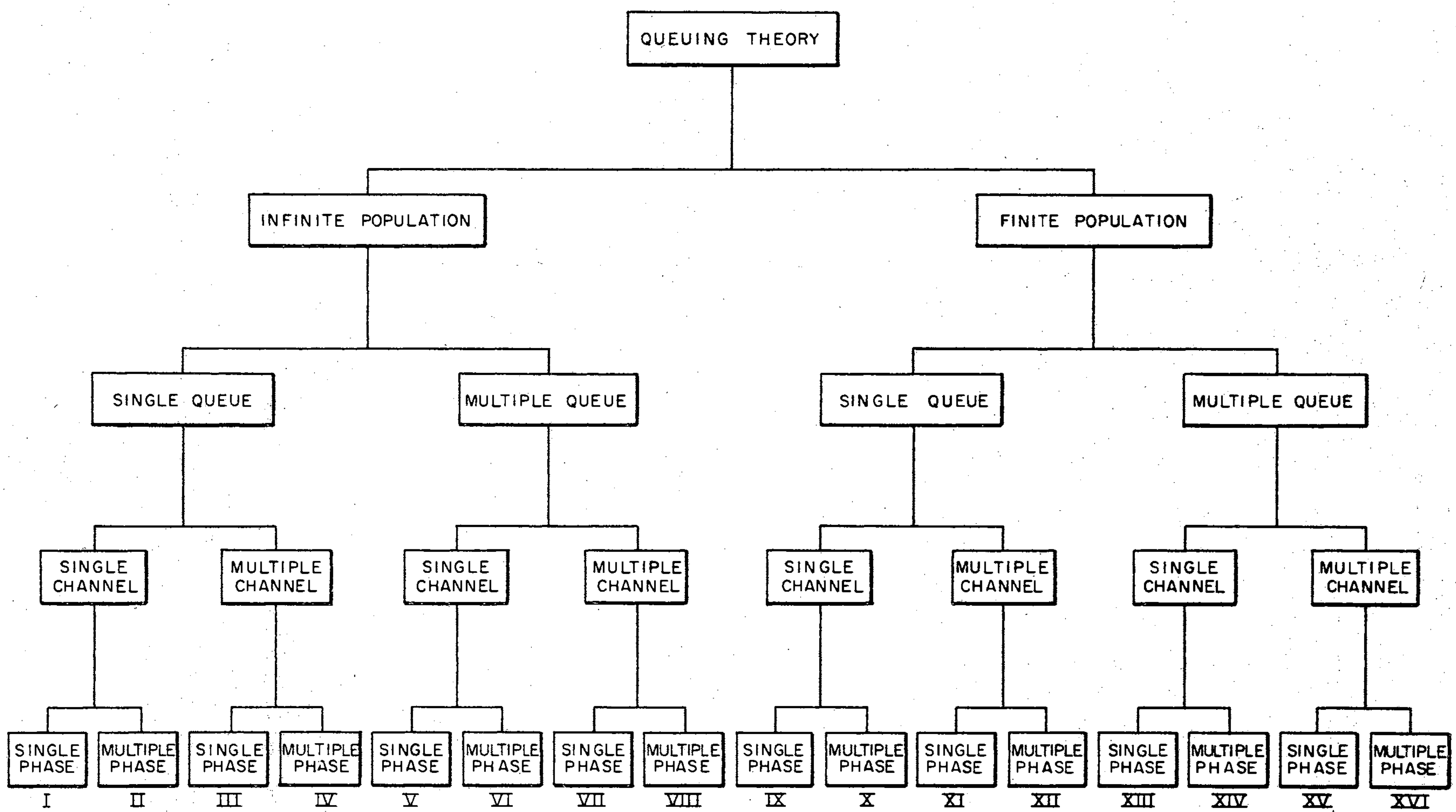


Figure 5. General Classification of Queuing Theory

single or multiple as shown in Figure 5. For example, case IV as shown in Figure 5 is classified as follows:

Source input: infinite

Queue : single

Service facility: Multiple channel, multiple phase.

The previous classification is called the general classification. Should assumptions be made to solve the situation in a particular way it would be called a specific classification. For example, the attitude of the customer is patient, the arrival time distribution is exponential, the service time distribution is constant, etc. Table II shows the general and specific classifications of the system.

TABLE II

GENERAL AND SPECIFIC CLASSIFICATION OF THE SYSTEM

Source	General Classification	Specific Classification
Input Source	i. Finite ii. Infinite	i. Arrival time distribution ii. Attitude of customers
Queue	i. Single ii. Multiple	i. Finite ii. Infinite
Service Facility	i. Channel a. Single b. Multiple ii. Phase a. Single b. Multiple	i. Service time distribution ii. Service discipline: a. First come, first served b. Random c. Priority ......, etc.

CHAPTER III

DETERMINISTIC WAITING LINE MODELS

Solution of Queuing Problems

In the previous chapter, the description and the classification of the waiting line system was discussed. Here, the solution of queuing problem is taken into consideration. For an operational situation that has been formulated as a queuing problem, the input source, the service system, and the service discipline are identified. Two methods are available for solving the problem, the mathematical approach and the simulation approach.

In the mathematical approach, the actual arrival and service-time distributions are approximated by one of the well-known mathematical distributions to arrive at relationships that describe the queuing process. From those relationships, one can determine the various operational characteristics. Further, if the relevant costs are known, conditions can be determined under which the operational situation gives optimal results; that is, minimizes the total cost or maximizes the profit.

In the simulation approach, statistics concerning arrivals and service times are duplicated mechanically, either from historical or assumed data. This is known as the Monte Carlo method and is particularly valuable when a computer is available. By duplicating a large number of arrivals with the operational situation on paper, one can determine the operational characteristics and the costs or profit (if the relevant costs due to waiting, idle time, etc., can be determined) resulting from changes in conditions such as the number of facilities, service rates, service discipline, etc.

The simulation approach is used in preference to the mathematical approach when the queuing system cannot be easily analyzed by mathematical means. Most of the analytical work for the solution of queuing problems has been concerned with operational situations in which the arrivals are Poisson and the service times follow the constant time or exponential distributions.

The objective of this chapter is to determine first the capacity of the service facility in the light of the relevant costs and the characteristics of the arrival pattern so that the sum of all costs associated with the waiting line system will be minimized. Second to draw the insignificant limits for the optimized situation. The mathematical approach is used for developing the models. All models developed here are based on the assumption that the arrival and service mechanisms are deterministic. That is, the future demand for service and the service duration considered are known with certainty. The model under consideration with these arrival and service mechanisms in this restricted sense is only an approximation of reality.

The Decision Model

The primary objective of waiting line system is to meet the demand for service at minimum cost. This requires the establishment of an appropriate level of service capacity by constructing and manipulating a mathematical model in the form

$$E = f(x_i , y_j)$$

where:

E = the measure of effectiveness sought (minimize total system cost)

x_i = the policy variable concerning the level of service capacity to provide

y_j = the environmental parameters of the arrival pattern, the waiting cost, and the service facility cost.

The following sections are devoted to developing deterministic decision models with the complete sensitivity analysis. The following symbolism will be used:

TC = total system cost per period

A = number of periods between arrivals

S = number of periods to complete one service

C_W = cost ot total waiting in the system per unit per period

C_f = cost of providing service facility of unit rate capacity.

Additional notation will be adopted and defined as required for deriving specific decision models.

Models for No Initial Waiting Line

In this section, assume that a queuing situation begins with no units in the system and that arrivals occur at regular intervals of length A period. The first arrival occurs at the beginning of the process. Service time is constant and equal to S periods. Since each unit serviced will require S periods, it is essential that S be less than or equal to A periods if a single phase-single channel is employed. If single phase-multiple channels "M" are to be used, it is required that S be less than or equal to MA periods. If these restrictions are violated, a waiting line will form which will grow beyond bound.

A Single Channel-Single Phase Model

This case is presented as single phase, single channel, single queue, infinite population in the general classification in Chapter II. This is shown in Figure 5 as Case I. Due to the general and the specific classifications combined, it is as follows.

TABLE III

MODEL CLASSIFICATIONS

Source	General Classification	Specific Classification
Input Source	Infinite	i. Constant arrival time distribution ii. Patient customers
Queue	Single	Infinite
Service Facility	i. Single channel ii. Single phase	i. Constant service time distribution ii. First come, first served

The system may be represented schematically as shown in Figure 6. The heavy dot represents an arrival every five periods. The slanting path represents a service operation requiring three periods. Since S is less than A, no waiting line will ever form.

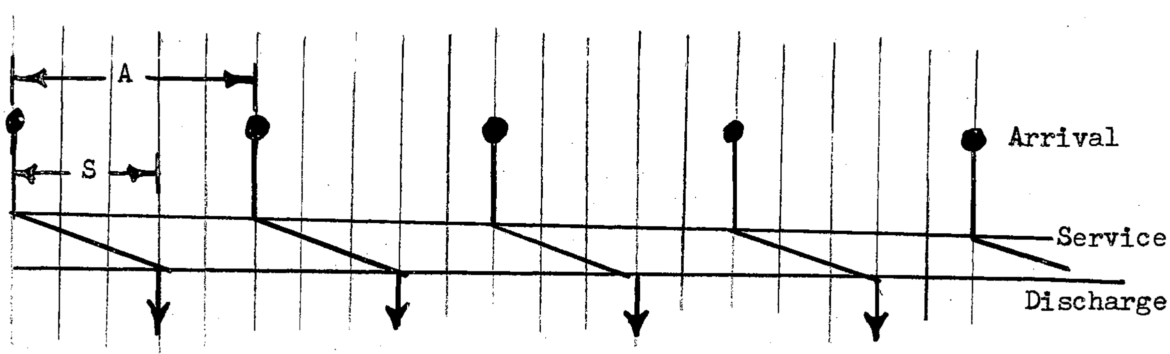


Figure 6. Single Channel System

The total system cost per period will be the sum of the waiting cost for the period and the service facility cost for the period; that is,

$$TC = WC + FC.$$

The waiting cost per period will be the product of the cost of total waiting in the system per unit per period and the number of units waiting each period, or

$$WC = C_W(\frac{S}{A}).$$

The service facility cost for the period will be the product of the number of units served during the period and the cost of serving one unit, or

$$FC = C_f(\frac{1}{S}).$$

The total system cost per period will be the sum of the waiting cost per period and the facility cost per period, or

$$TC = C_W(\frac{S}{A}) + C_f(\frac{1}{S}). \qquad (3\text{-}1)$$

A minimum cost service interval may be found by differentiating with respect to S, setting the result equal to zero, and solving for S as follows:

$$\frac{dTC}{ds} = \frac{C_W}{A} - \frac{C_f}{S^2} = 0$$

$$S^2 = \frac{C_f A}{C_W}$$

$$S_o = \sqrt{\frac{C_f A}{W}} \quad , \quad S \leq A. \tag{3-2}$$

The total system cost per period at S_o, (TC_o), can be calculated as follows:

$$TC_o = \frac{C_W}{A}\left(\sqrt{\frac{C_f A}{C_W}}\right) + C_f\sqrt{\frac{C_W}{C_f A}}$$

$$= \sqrt{\frac{C_W C_f}{A}} + \sqrt{\frac{C_f C_W}{A}}$$

$$TC_o = 2\sqrt{\frac{C_W C_f}{A}}. \tag{3-3}$$

As an application of the foregoing model, consider the following example. A unit will arrive every five periods. The cost of waiting is $5 per unit per period. One unit can be served at a cost of $9. Waiting cost per period, facility cost per period, and total cost per period may be tabulated as a function of S to illustrate the nature of the cost components. The results shown in Table III were developed from Equation (3-1). Inspection of the tabulated values indicates that waiting cost per period is directly proportional and that facility cost per period is inversely proportional to S. The minimum cost service interval is three periods and may be provided at a facility cost of $3 per period.

The minimum cost service interval may be found directly by substituting into Equation (3-3) as follows:

$$S_o = \sqrt{\frac{9(5)}{5}} = 3 \text{ periods.}$$

TABLE IV

COST COMPONENTS FOR A SINGLE CHANNEL MODEL

S	WC	FC	TC
0	$0.00	$∞	$∞
1	1.00	9.00	10.00
2	2.00	4.50	6.50
3	3.00	3.00	6.00
4	4.00	2.25	6.25
5	5.00	1.80	6.80

Under the conditions assumed, the decision maker would provide a single service channel capable of serving one unit every three periods as shown in Figure 6. The total system cost, TC_o, at S_o would be

$$TC_o = 2\sqrt{\frac{9(5)}{5}} = \underline{\underline{\$6}} \text{ per period.}$$

The graphical solution of the case is shown in Figure 7.

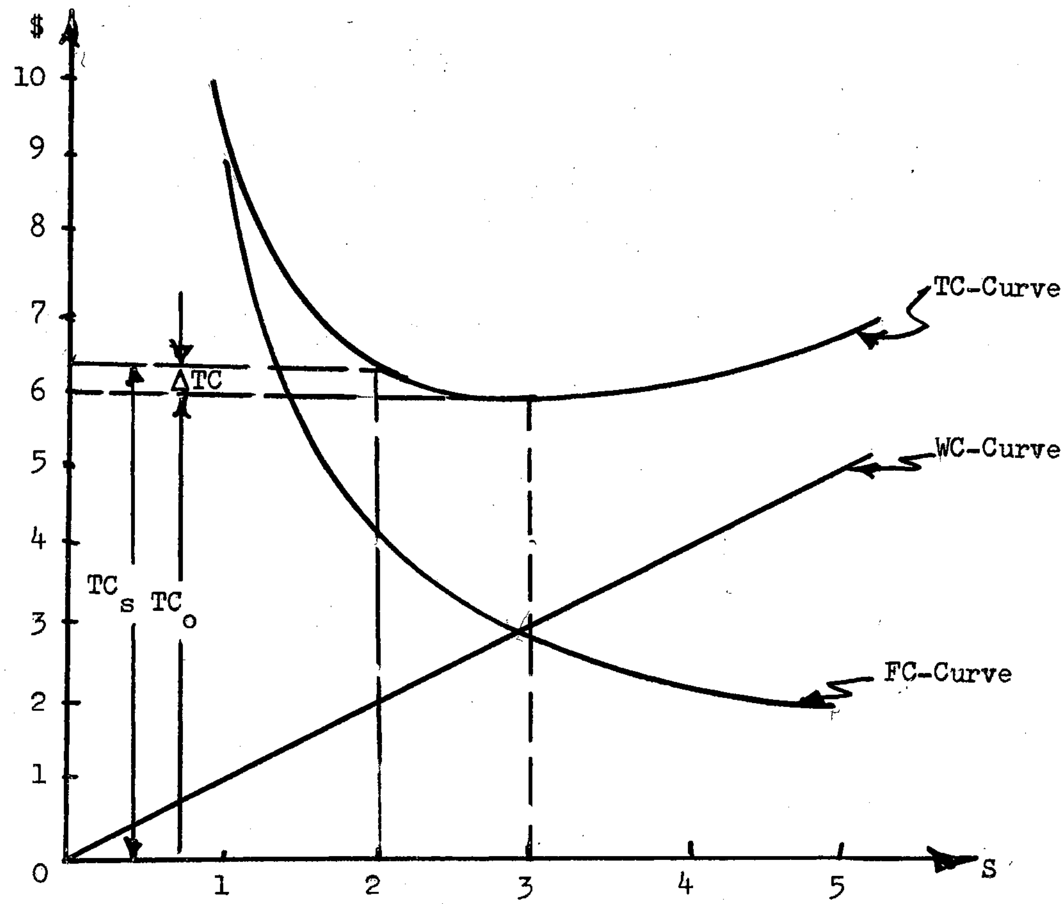


Figure 7. Cost Components for a Single Channel Model

TC = WC+FC

Sensitivity Analysis

The minimum cost service interval and the minimum total cost have been found in the last section. Should the decision maker change the real situation under consideration to the minimum total cost situation calculated? To answer this question, it is necessary to know about the sensitivity of the system as the optimum policy for minimizing the total system cost is known. Two kinds of sensitivity are considered in this chapter:

1. Sensitivity in terms of the behavior of the total system cost if the service interval deviates from the optimum policy.
2. Sensitivity of the optimum policy to changes in various system parameters, C_f and C_W.

Consider the first type of sensitivity. Let

TC_S = Total system cost at S periods of service

TC_o = Minimum total system cost at S_o periods of service

$$\Delta TC = TC_S - TC_o$$

and
$$K = \frac{\Delta TC}{TC_o} = \frac{TC_S - TC_o}{TC_o}.$$

By substituting the values for TC_S and TC_o, the last equation would be:

$$K = \frac{\left[C_W(\frac{S}{A}) + C_f(\frac{1}{S})\right] - 2\sqrt{\frac{C_f C_W}{A}}}{2\sqrt{\frac{C_f C_W}{A}}}$$

$$K = \frac{1}{2\sqrt{\frac{C_f C_W}{A}}}\left[C_W(\frac{S}{A}) + C_f(\frac{1}{S})\right] - 1$$

$$\therefore \quad 2(K+1) = \sqrt{\frac{C_W}{C_f A}}\, S + \sqrt{\frac{C_f A}{C_W}}\,(\frac{1}{S}).$$

Let $\sqrt{\frac{C_W}{C_f A}} = Z.$

Substituting for Z in the last equation, it would be

$$2(K + 1) = ZS + \frac{1}{ZS}.$$

Multiply both sides by S,

$$2(K + 1)S = ZS^2 + \frac{1}{Z}.$$

$$\therefore ZS^2 - (2K + 2)S + \frac{1}{Z} = 0.$$

The previous equation is quadratic, solve for S, or

$$S = \frac{(2K + 2) \pm \sqrt{4(K + 1)^2 - 4}}{2Z}$$

$$S = \frac{2[(K + 1) \pm \sqrt{(K + 1)^2 - 1}]}{2Z}$$

$$S = \frac{(K + 1) \pm \sqrt{(K + 1)^2 - 1}}{Z}$$

$$S = \frac{1}{Z}[(K + 1) \pm \sqrt{(K + 1 - 1)(K + 1 + 1)}]$$

$$S = \frac{1}{Z}[(K + 1) \pm \sqrt{K(K + 2)}].$$

As shown from the previous equation, there are two values of S for each value of K. Let the positive sign indicate the higher value of S, S_h, and the negative sign indicates the lower value of S, S_L. That is,

$$S = \frac{1}{Z}[(K+1) \pm \sqrt{K(K+2)}]$$

$$S_L = \frac{1}{Z}[(K+1) - \sqrt{K(K+2)}] \tag{3-4}$$

$$S_h = \frac{1}{Z}[(K+1) + \sqrt{K(K+2)}]$$

But as previously indicated, the value of Z is

$$Z = \sqrt{\frac{C_W}{C_f A}} \tag{3-5}$$

$$Z = \frac{1}{\sqrt{\frac{C_f A}{C_W}}}$$

$$\text{Since } S_o = \sqrt{\frac{C_f A}{C_W}}$$

$$\therefore Z = \frac{1}{S_o}. \tag{3-6}$$

Equation (3-6) shows the relationship between Z and S_o. Substituting from (3-5) in the set of equations (3-4), it would be,

$$S = \sqrt{\frac{C_f A}{C_W}}\,[(K+1) \pm \sqrt{K(K+2)}]$$

$$S_L = \sqrt{\frac{C_f A}{C_W}}\,[(K+1) - \sqrt{K(K+2)}] \tag{3-7}$$

$$S_h = \sqrt{\frac{C_f A}{C_W}}\,[(K+1) + \sqrt{K(K+2)}]$$

In the set of Equations (3-7), the service time per unit, S, is a function of K, the ratio of the incremental cost, ΔT, and minimum total system cost at certain values of the parameters. So S_L and S_h indicate the insignificant higher and lower limits of the service time per unit at a certain per cent of incremental cost to minimum system total cost. That is, if the decision maker sets a certain value of K, which gives no significance in cost difference, he can get the insignificance limits of S. Beyond these limits he will approve the change of the real situation to the optimum case. Within these limits he will keep the situation as it is since the benefit he will get is less than outlay for the change. In other words, the benefit he will get by the change to the optimum case has no value to him with respect to the money he will spend on the change.

As an application of the foregoing, consider the previous example which has C_W = 5 \$/unit/period, C_f = 9 \$/unit, and A = 5 periods. By substituting for the parameters' values in Equations (3-7), the upper and lower insignificant limits as a function of K would be

$$S = \sqrt{\frac{9(5)}{5}}\,[(K + 1) \pm \sqrt{K(K + 2)}].$$

$$\therefore S_u = 3[(K + 1) + \sqrt{K(K + 2)}]$$

$$S_L = 3[(K + 1) - \sqrt{K(K + 2)}].$$

Table V shows the upper and lower insignificant limit values, S_u and S_L, at different values of K. The graphical solution is given in Figure 8.

It is easily seen from the graph that the ratio of change of S_h is greater than the rate of change of S_L as the values of K increases. This was the sensitivity analysis in terms of the behavior of the total system cost if the service interval deviates from the optimum policy.

The second type of sensitivity, the sensitivity of the optimum policy to changes in system parameter, is illustrated as follows:

Define the parameter R as the ratio $^{C_W}/_{C_f}$. Table VI and Figure 9 give three curves of Equation (3-7) at different values of the parameter R, R = $^5/_9$, R = 1, and R = $^9/_5$. It is easily to see that as the value of the parameter R increases the optimum interval S_o increases and the range of the insignificant limits, $S_h - S_L$, increases. This is also shown in Figures 10 and 11 which their values are calculated at different values of A.

Figure 12 shows the locus of S_o, and the surface of insignificant limits at constant value of the parameter R, due to the change of A value in 3-dimensional analysis. Therefore, as A increases the value of S_o increases at certain parameter, R = 1. Also, if other surfaces are drawn at different values of R, R = $^9/_5$ and R = $^5/_9$, it can be seen that the locus of S_o shifts to the right, increases in value, as the parameter R decreases. This

TABLE V

TABULATION OF INSIGNIFICANT LIMITS WITH RESPECT TO K

K	$S_u = S_o[(K+1)+\sqrt{K(K+2)}]$	$S_L = S_o[(K+1)-\sqrt{K(K+2)}]$
0	3	3
0.1	4.674	1.926
0.2	5.589	1.611
0.3	6.393	1.407
0.4	7.14	1.260
0.5	7.954	1.146
0.6	8.547	1.053
0.7	9.225	0.975
0.8	9.837	0.909
0.9	10.548	0.852
1.0	11.196	0.804

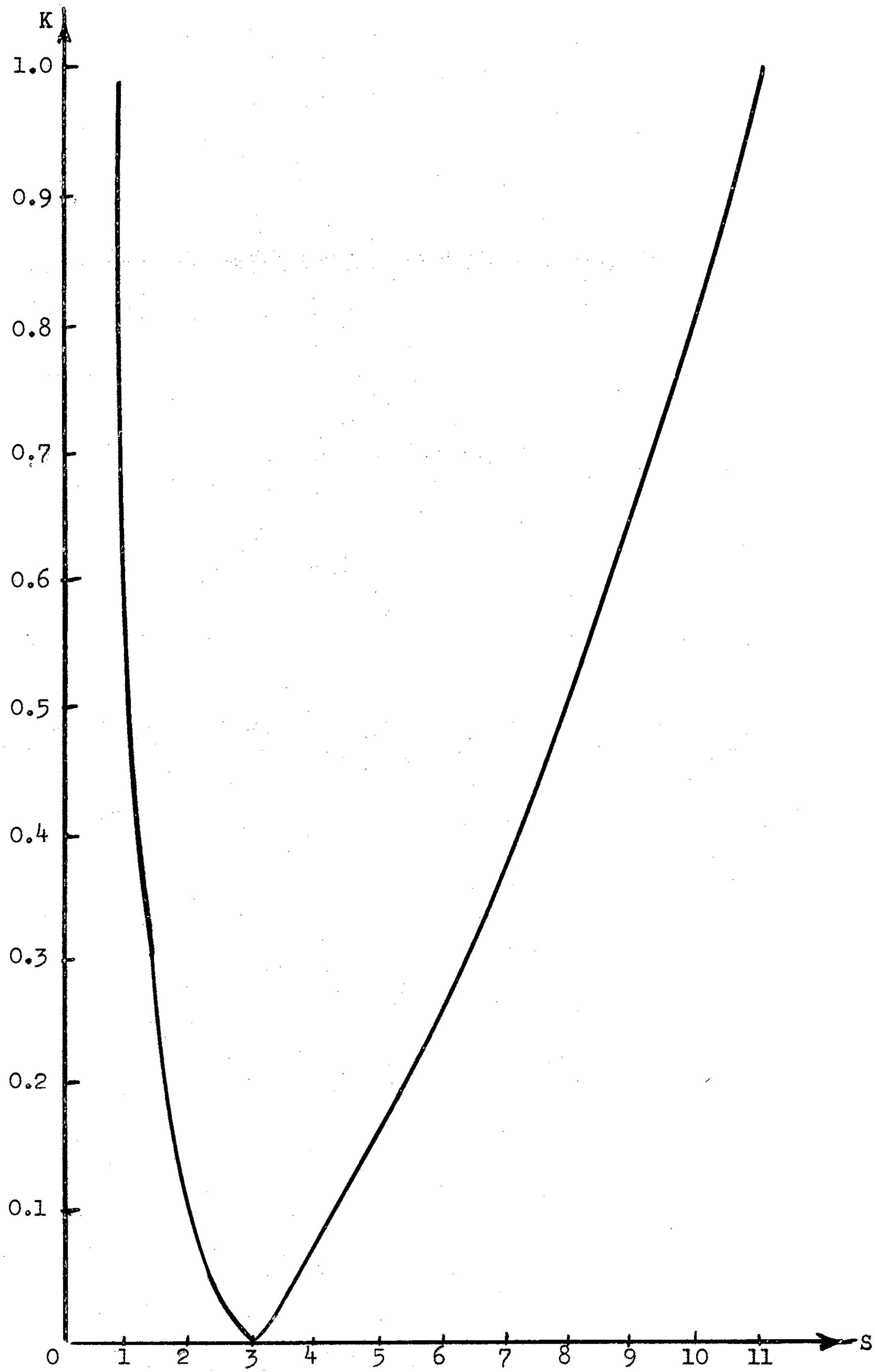


Figure 8. K-S Curve

TABLE VI

INSIGNIFICANT LIMITS AT A = 5 PERIODS AND DIFFERENT VALUES OF THE PARAMETER R

K	R = 5/9		R = 1		R = 9/5	
	S_L	S_u	S_L	S_u	S_L	S_u
0	3	3	2.236	2.236	1.667	1.667
0.1	1.926	4.674	1.436	3.484	1.070	2.600
0.2	1.611	5.589	1.201	4.166	0.895	3.106
0.3	1.407	6.393	1.049	4.765	0.782	3.552
0.4	1.260	7.14	0.939	5.322	0.700	3.967
0.5	1.146	7.954	0.854	5.854	0.637	4.364
0.6	1.053	8.547	0.783	6.373	0.583	4.751
0.7	0.975	9.225	0.727	6.878	0.542	5.126
0.8	0.909	9.837	0.678	7.372	0.505	5.496
0.9	0.852	10.548	0.635	7.862	0.473	5.861
1.0	0.804	11.196	0.600	8.345	0.447	6.221

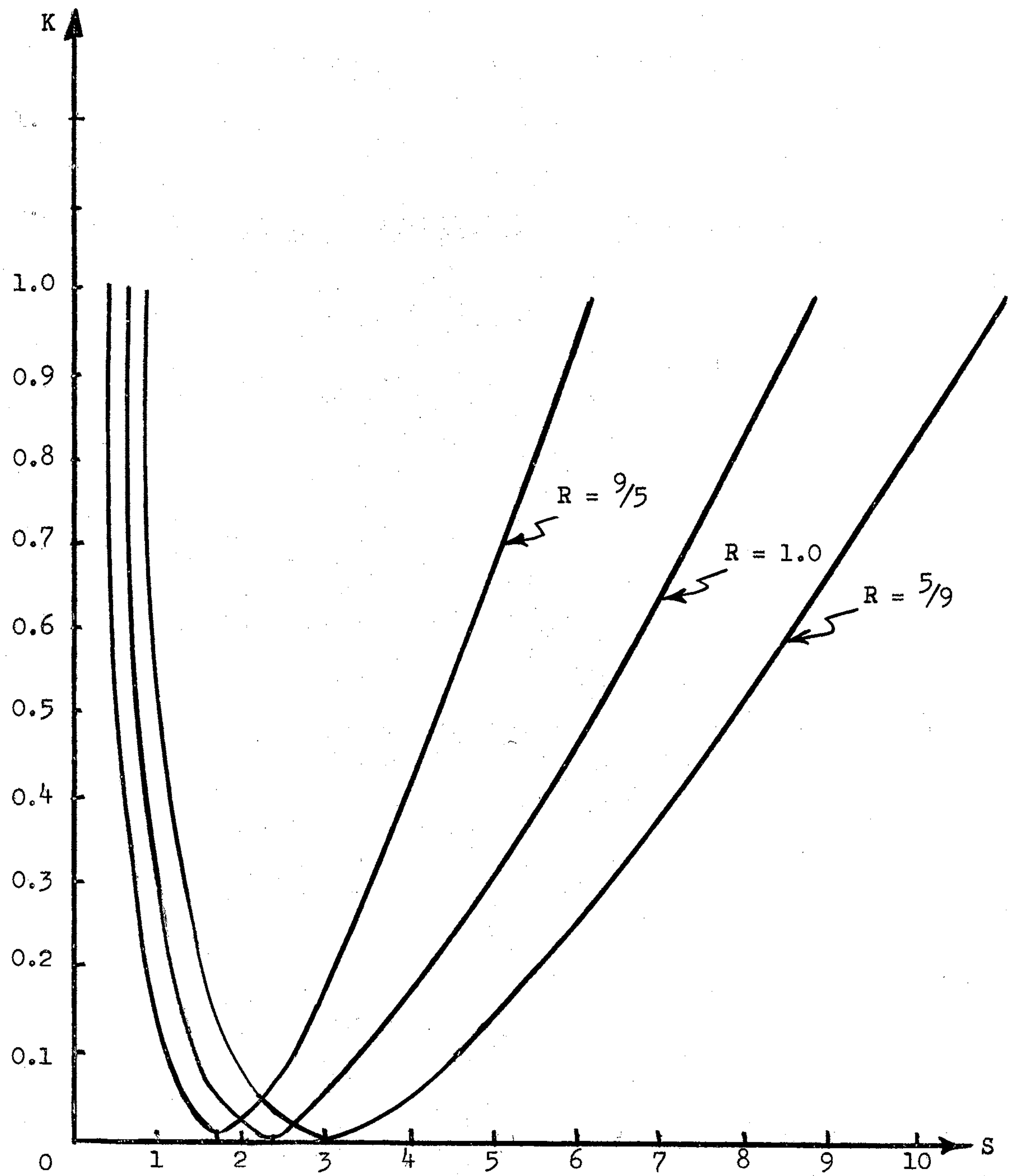


Figure 9. Insignificant Limits at A = 5 and Different Values of R

TABLE VII

INSIGNIFICANT LIMITS AT A = 4 AND DIFFERENT VALUES OF PARAMETER R

	R = 5/9		R = 1		R = 9/5	
K	S_L	S_u	S_L	S_u	S_L	S_u
0	2.683	2.683	2.00	2.00	1.491	1.491
0.1	1.722	4.180	1.284	3.116	0.957	2.323
0.2	1.441	4.998	1.074	3.726	0.801	2.778
0.3	1.258	5.717	0.938	4.262	0.699	3.177
0.4	1.127	6.386	0.840	4.76	0.626	3.548
0.5	1.024	7.024	0.746	5.236	0.569	3.903
0.6	0.939	7.646	0.70	5.70	0.522	4.249
0.7	0.872	8.250	0.650	6.190	0.485	4.589
0.8	0.813	8.846	0.606	6.594	0.452	4.916
0.9	0.762	9.433	0.568	7.032	0.423	5.242
1.0	0.719	10.013	0.536	7.464	0.399	5.564

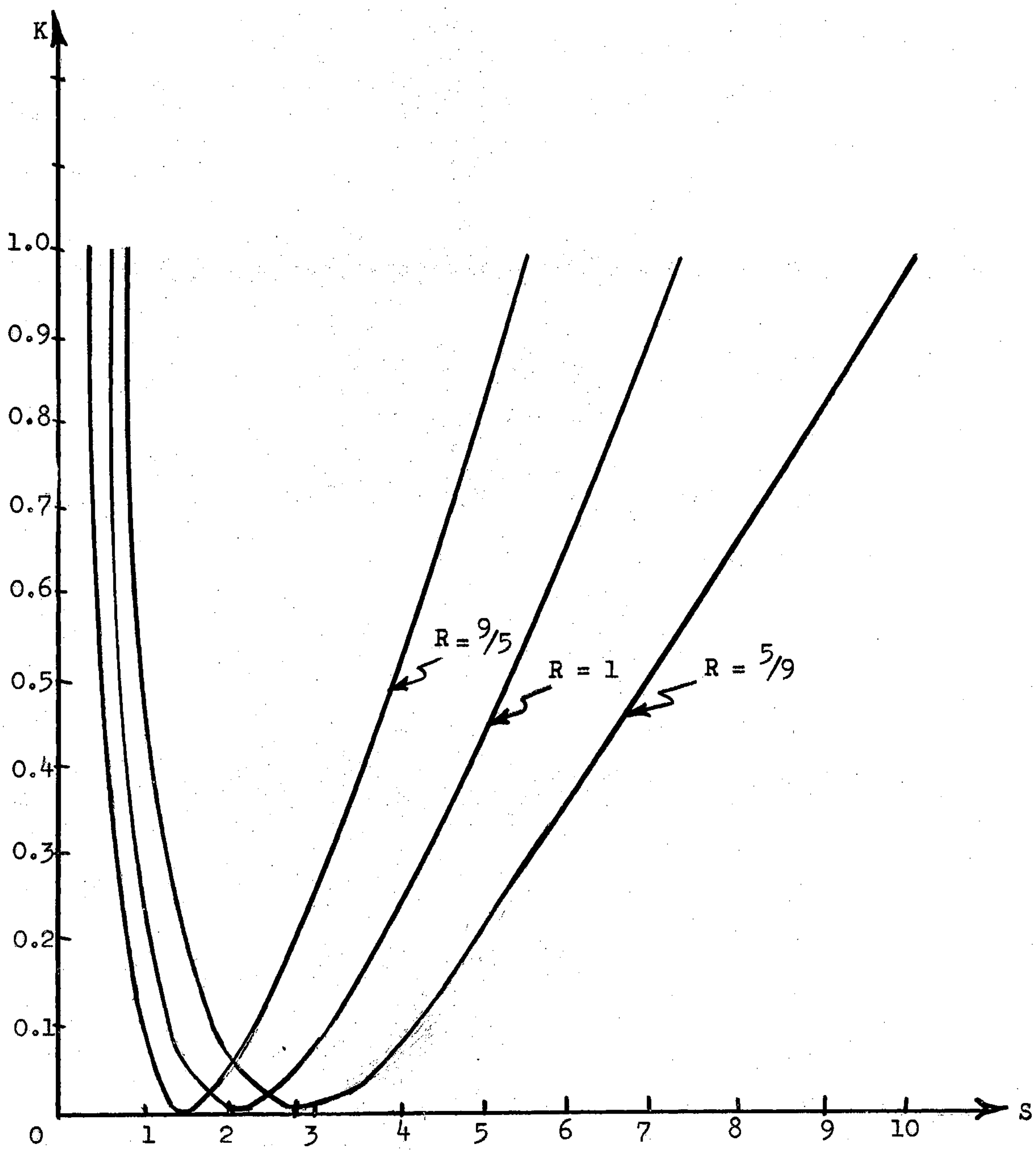


Figure 10. Insignificant Limits at A = 4 and Different Values of R

TABLE VIII

INSIGNIFICANT LIMITS AT A = 3 PERIODS AND DIFFERENT VALUES OF THE PARAMETER R

K	R = 5/9		R = 1		R = 9/5	
	S_L	S_u	S_L	S_u	S_L	S_u
0	2.323	2.323	1.732	1.732	1.288	1.288
0.1	1.491	3.619	1.112	2.698	0.827	2.006
0.2	1.247	4.328	0.93	3.227	0.692	2.400
0.3	1.089	4.950	0.812	3.691	0.604	2.745
0.4	0.976	5.529	0.727	4.122	0.541	3.065
0.5	0.887	6.082	0.662	4.534	0.492	3.372
0.6	0.813	6.621	0.606	4.936	0.451	3.671
0.7	0.755	7.143	0.563	5.326	0.419	3.961
0.8	0.704	7.659	0.525	5.710	0.390	4.246
0.9	0.660	8.168	0.492	6.090	0.366	4.529
1.0	0.623	8.669	0.464	6.464	0.345	4.807

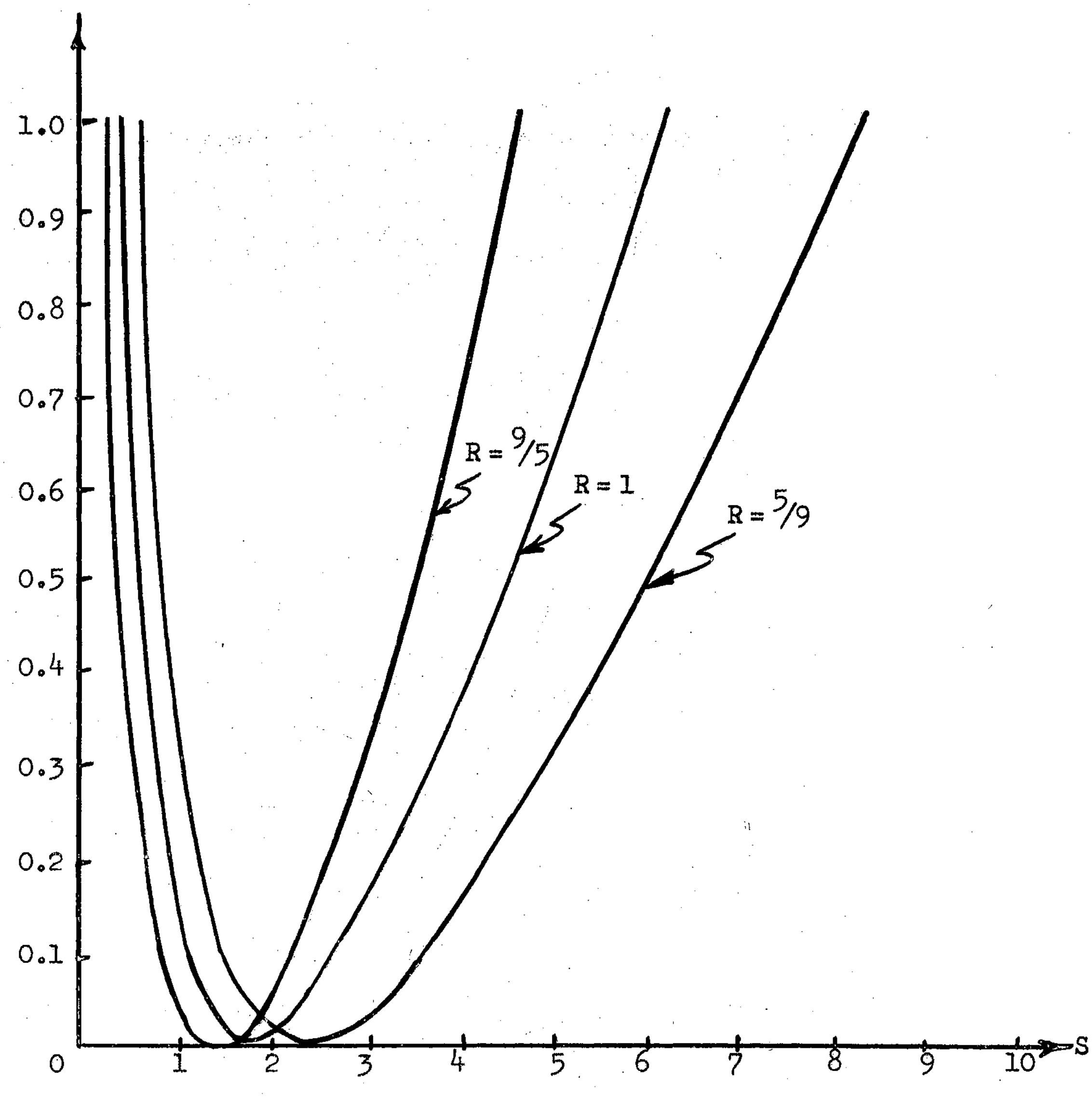


Figure 11. Insignificant Limits at A = 3 Periods and Different Values of R

can be shown from the values calculated in the previous tables.

Indifferent Range

The indifferent range, r, is calculated as the difference between the higher and the lower insignificant limits S_h and S_L, respectively; i.e.,

$$r = S_h - S_L.$$

Substitute in the previous equation for S_h and S_L from Equation (3-7) and the result would be

$$r = \sqrt{\frac{C_f A}{C_W}}[(K+1) + \sqrt{K(K+2)}] - \sqrt{\frac{C_f A}{C_W}}[(K+1) - \sqrt{K(K+2)}]$$

$$r = \sqrt{\frac{C_f A}{C_W}}[2\sqrt{K(K+2)}]$$

$$r = 2\sqrt{\frac{C_f A K(K+2)}{C_W}}.$$

By putting $R = \frac{C_W}{C_f}$

$$r = 2\sqrt{\frac{AK}{R}(K+2)}. \qquad (3\text{-}8)$$

The indifferent range, r, is a function of K, A, and R. Table X and Figure 13 give the interpretation of the previous equation by the aid of numerical values of K, A, and R. As the value of R increases at constant K, the

TABLE IX

INSIGNIFICANT LIMITS OF R = 1 AND DIFFERENT VALUES OF A

K	A = 3		A = 4		A = 5	
	S_L	S_h	S_L	S_h	S_L	S_h
0	1.732	1.732	2.00	2.00	2.236	2.236
0.1	1.112	2.698	1.284	3.116	1.436	3.484
0.2	0.930	3.227	1.074	3.726	1.201	4.166
0.3	0.812	3.691	0.938	4.262	1.049	4.765
0.4	0.727	4.122	0.840	4.76	0.939	5.322
0.5	0.662	4.534	0.764	5.236	0.854	5.854
0.6	0.606	4.936	0.700	5.70	0.783	6.373
0.7	0.563	5.326	0.650	6.190	0.727	6.878
0.8	0.525	5.710	0.606	6.594	0.678	7.372
0.9	0.492	6.090	0.568	7.032	0.635	7.862
1.0	0.464	6.464	0.536	7.464	0.600	8.345

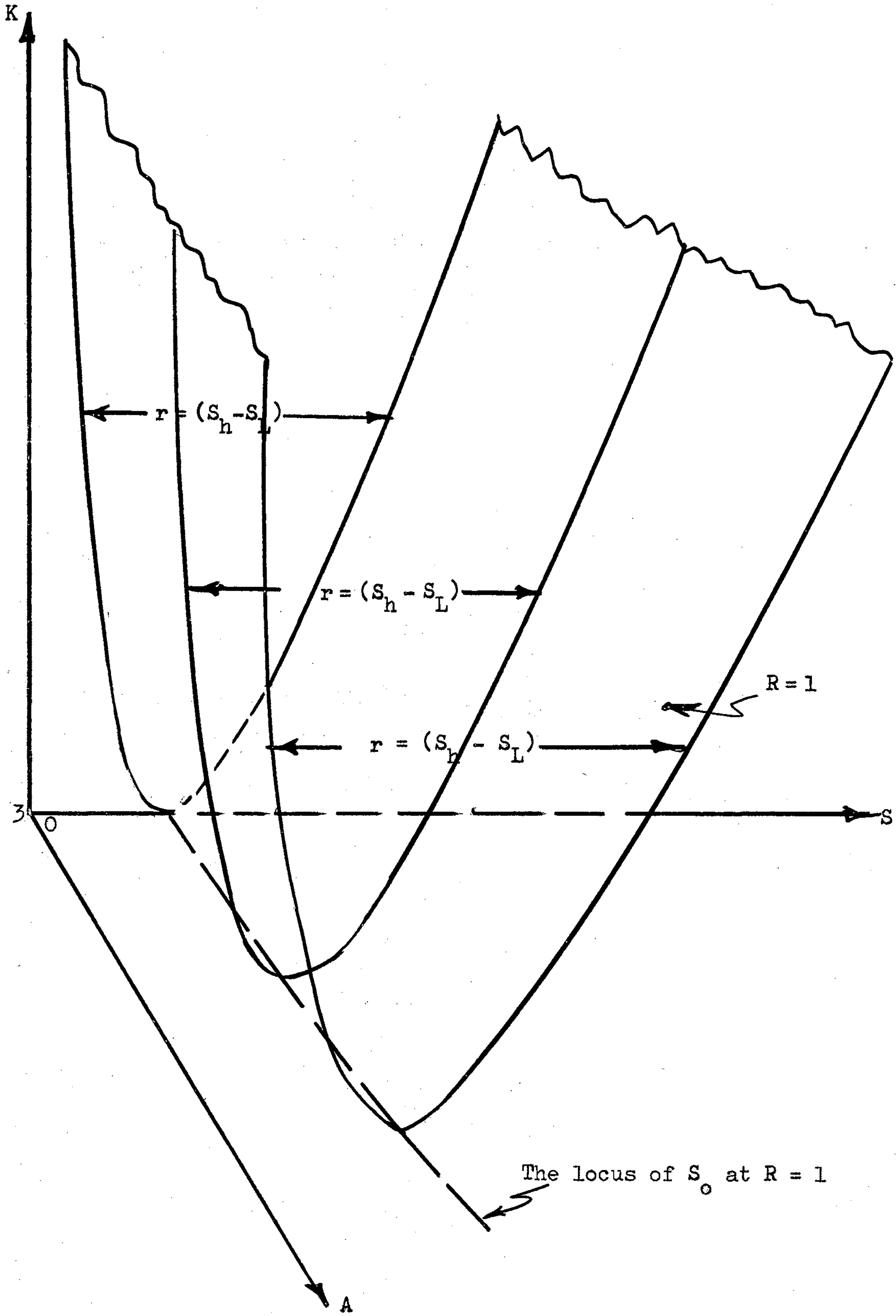


Figure 12. The Locus of S_o at R = 1

TABLE X

THE RANGE OF INSIGNIFICANT LIMITS TABLE

K	Range "r"								
	A = 5			A = 4			A = 3		
	R=5/9	R=1	R=9/5	R=5/9	R=1	R=9/5	R=5/9	R=1	R=9/5
0	0	0	0	0	0	0	0	0	0
0.1	2.748	2.048	1.530	2.458	1.832	1.366	2.128	1.586	1.179
0.2	3.978	2.965	2.211	3.557	2.652	1.977	3.081	2.297	1.708
0.3	4.986	3.671	2.77	4.459	3.324	2.418	3.861	2.879	2.141
0.4	5.88	4.383	3.267	5.259	3.920	2.922	4.553	3.395	2.524
0.5	6.808	5.000	3.727	6.00	4.472	3.334	5.195	3.872	2.880
0.6	7.494	5.590	4.168	6.707	5.00	3.727	5.808	4.33	3.22
0.7	8.250	6.191	4.584	7.370	5.50	4.100	6.388	4.763	3.542
0.8	8.92	6.694	4.991	8.033	5.988	4.464	6.955	5.185	3.856
0.9	9.694	7.227	5.388	8.671	6.464	4.819	7.508	5.598	4.163
1.0	10.392	7.745	5.774	9.294	6.928	5.165	8.046	6.00	4.462

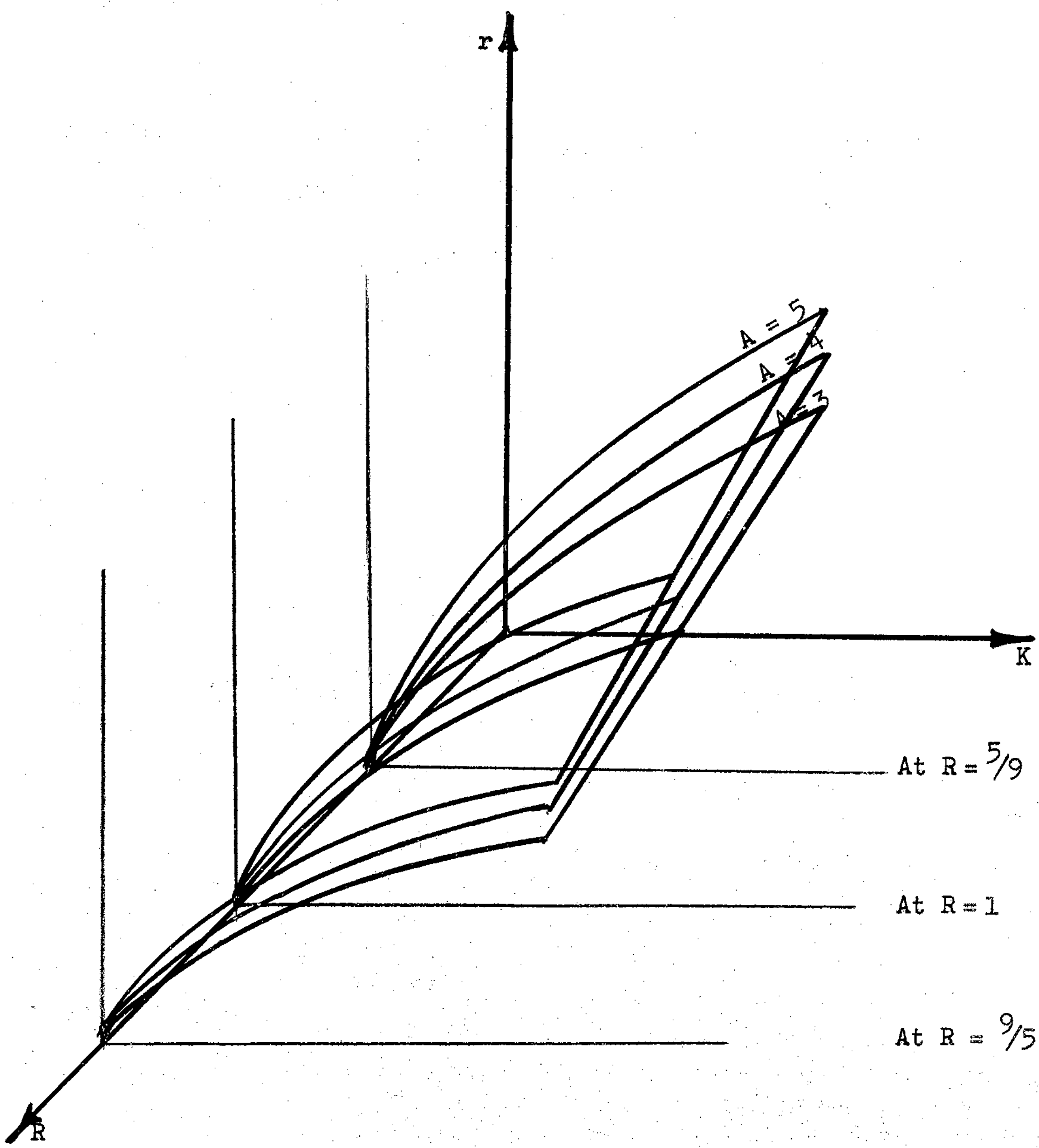


Figure 13. Indifferent Range Surface

indifferent range decreases. Also as K or A increases, the range increases when R is equal to constant value.

A Multiple Channel, Single Phase Model

This is the Case III shown in Figure 5. Its general classification is as follows: Infinite population, single queue, multiple channel, and single phase case.

The general and the specific classification combined is shown in the following table.

TABLE XI

GENERAL AND SPECIFIC CLASSIFICATION OF THE CASE

Source	General Classification	Specific Classification
Input Source	Infinite	1. Constant arrival time distribution 2. Patient customers
Queue	Single	Infinite
Service Facility	1. Single Phase 2. Multiple Channel	1. Constant service time distribution 2. First in, first served

A two channel waiting line system may be represented schematically as shown in Figure 14. The heavy dot represents an arrival and the slanting path represents a service operation. The second arrival finds the first service channel busy and goes immediately into the second. The third arrival finds the second channel busy and goes immediately into the first. Since S is less than MA, no waiting line will ever form.

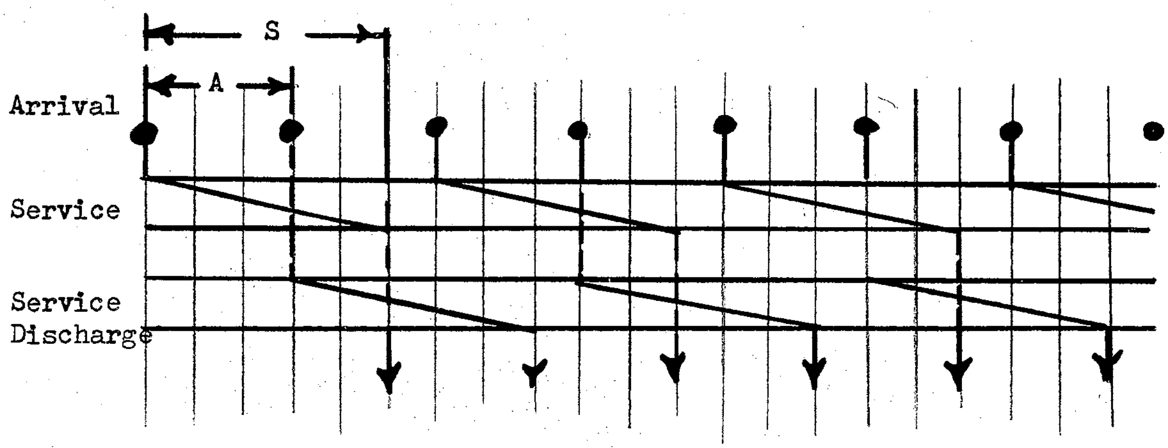


Figure 14. A Multiple Channel, Single Phase System

The total system cost for the period will be the sum of the waiting cost for the period and the facility cost for the period; that is,

$$TC = WC + FC.$$

The waiting cost per period will be the product of the cost of total waiting in the system per unit per period and the number of units waiting per period, or

$$WC = C_W(\frac{S}{A}).$$

The facility cost for the period will be the product of the number of units served during the period, the cost of serving one unit, and the number of channels in operation, or

$$FC = C_f(\frac{1}{S})M.$$

The total system cost per period will be the sum of the waiting cost per period and the facility cost per period, or

$$TC = C_W(\frac{S}{A}) + C_f(\frac{1}{S})M.$$

A minimum cost service interval may be found by differentiating with respect to S, setting the result equal to zero, and solving for S as follows:

$$\frac{dTC}{ds} = \frac{C_W}{A} - \frac{C_f M}{S^2} = 0$$

$$S_o^2 = \frac{C_f MA}{C_W}$$

$$\therefore S_o = \sqrt{\frac{C_f MA}{C_W}}, \quad S \leq MA. \qquad (3\text{-}10)$$

Equation (3-10) reduces to Equation (3-3) for the single channel-single phase case; M = 1.

The minimum total system cost can be obtained by substituting for S in Equation (3-9) from Equation (3-10), or

$$TC_o = \frac{C_W}{A}\sqrt{\frac{C_f MA}{C_W}} + C_f M\sqrt{\frac{C_W}{C_f MA}}$$

$$TC_o = \sqrt{\frac{C_f MC_W}{A}} + \sqrt{\frac{C_f MC_W}{A}}$$

$$TC_o = 2\sqrt{\frac{C_f MC_W}{A}}. \qquad (3\text{-}11)$$

To illustrate the application of the foregoing model, consider the following example. A unit will arrive every four periods. The cost of waiting is $2 per unit per period. A unit can be served for a cost of $9. Two channels are to be used. The minimum cost service interval may be found by substituting into Equation (3-10) as follows:

$$S_o = \sqrt{\frac{9(2)(4)}{2}} = 6 \text{ periods}$$

and the minimum total system cost may be found by substituting in Equation (3-11), or

$$TC_o = 2\sqrt{\frac{9(2)(2)}{4}} = 6 \text{ \$/period.}$$

The minimum total system cost can be written as a function of S_o by substituting from Equation (3-10) into Equation (3-11) as follows:

$$TC_o = 2\sqrt{\frac{C_f MC_W}{A} \cdot \frac{AC_W}{AC_W}}$$

$$= 2\sqrt{\frac{C_f MA}{C_W} \cdot \frac{C_W^2}{A^2}}$$

$$TC_o = \frac{2\ S_o\ C_W}{A}. \qquad (3\text{-}12)$$

From the previous example, it is known that

S_o = 6 periods

C_W = 2 \$/period/unit

A = 4 periods

$\therefore\ TC_o = \frac{2(6)2}{4} = 6$ \$/period

which is the same result obtained from Equation (3-11). Once S_o is known, use Equation (3-12) to facilitate the calculation.

Under the conditions assumed, the decision maker would provide two service channels, each with the capacity of serving one unit every six periods. Figure 15 shows the TC surface in xyz-space. It also shows the locus of S_o at different values of A on TC-surface and locus of S_o on S-A plane.

Sensitivity Analysis

Now to define K as before

$$K = \frac{TC_s - TC_o}{TC_o}.$$

Substitute for TC_s and TC_o in the previous equation, it would be

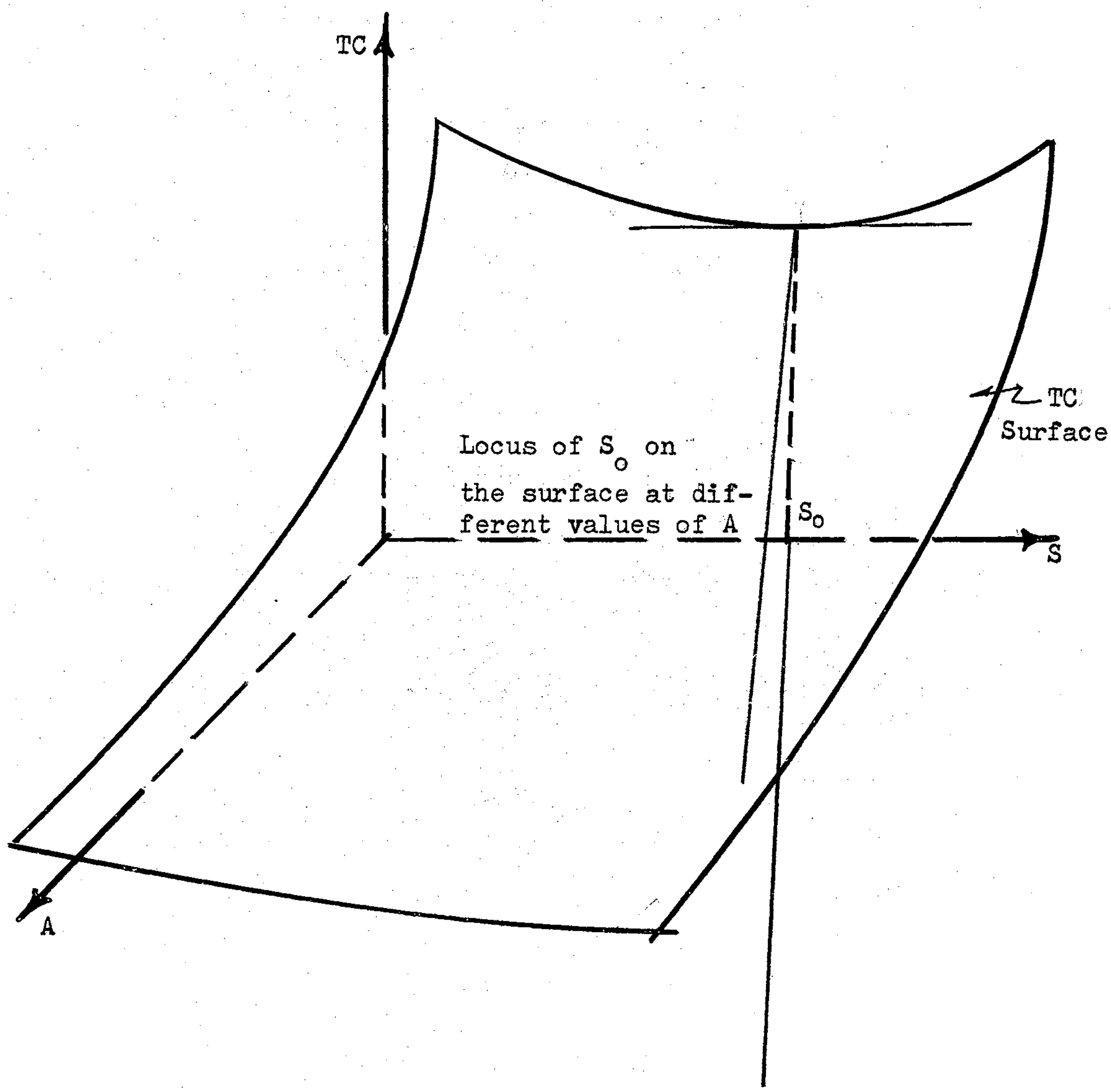


Figure 15. The Locus of S_o

$$K = \frac{\left[C_W(\frac{S}{A}) + C_f M(\frac{1}{S})\right] - 2\sqrt{\frac{C_f M C_W}{A}}}{2\sqrt{\frac{C_f M C_W}{A}}}$$

$$K = \frac{1}{2}\sqrt{\frac{A}{C_f MA}}\left[C_W(\frac{S}{A}) + C_f M(\frac{1}{S})\right] - 1$$

$$(2K + 2) = \sqrt{\frac{C_W}{CfMA}} \cdot S + \sqrt{\frac{AMC_f}{C_W}} \cdot \frac{1}{S}.$$

But $S_o = \sqrt{\frac{C_f MA}{C_W}}$

$$\therefore (2K + 2) = \frac{1}{S_o} \cdot S + S_o \cdot \frac{1}{S}.$$

Multiply both sides by S, or

$$(2K + 2)S = \frac{1}{S_o} S^2 + S_o$$

$$\therefore \frac{1}{S_o} S^2 - (2K + 2)S + S_o = 0. \qquad (3\text{-}13)$$

Equation (3-13) is a second degree in S. Solving for S, S would be

$$S = \frac{(2K + 2) \pm \sqrt{(2K + 2)^2 - 4}}{2(\frac{1}{S_o})}$$

$$S = \frac{2[(K + 1) \pm \sqrt{(K + 1)^2 - 1}]}{2(\frac{1}{S_o})}$$

$$S = S_o[(K + 1) \pm \sqrt{(K + 1 - 1)(K + 1 + 1)}]$$

$$S = S_o[(K + 1) \pm \sqrt{K(K + 2)}].$$

Substituting for S_o would be:

$$S = \sqrt{\frac{C_f MA}{C_W}[(K+1) \pm \sqrt{K(K+2)}]}$$

$$S_L = \sqrt{\frac{C_f MA}{C_W}[(K+1) - \sqrt{K(K+2)}]} \qquad (3\text{-}14)$$

$$S_h = \sqrt{\frac{C_f MA}{C_W}[(K+1) + \sqrt{K(K+2)}]}$$

Figure 16 shows the shape of S_h-curve and S_L-curve. They have exactly the same shape as in single channel-single phase curves of Equation (3-7). Equation (3-14) can be reduced to Equation (3-7) by putting $M = 1$. By comparison of the Equations (3-7) and (3-14), it can be said that by multiplying the insignificant limits in Equation (3-7) by $\sqrt{M}$ would be the insignificant limits for multiple channel-single phase case. So the multiple channel-single phase has the same analysis as the single channel-single phase case with scale $1 = \sqrt{M}$.

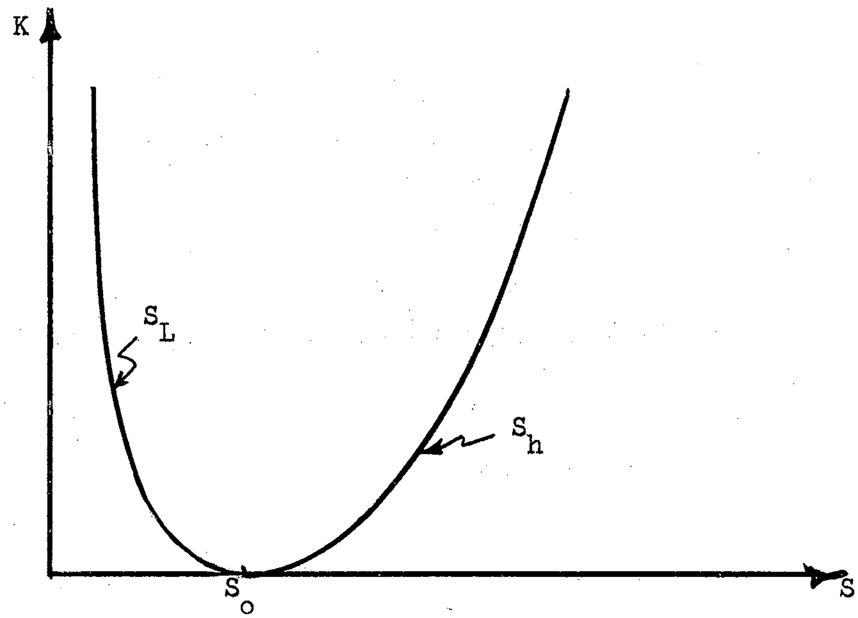


Figure 16. S_L and S_h-curves for Multiple Channel-Single Phase Case

Indifferent Range

The indifferent range can be calculated as follows:

$$r = S_h - S_L.$$

By substitution for S_h and S_L in the previous equation from Equation (3-14), it would be

$$r = \sqrt{\frac{C_f MA}{C_W}}\left[\{(K+1) + \sqrt{K(K+2)}\} - \{(K+1) - \sqrt{K(K+2)}\}\right]$$

$$r = \sqrt{\frac{C_f MA}{C_W}} \cdot 2\sqrt{K(K+2)}$$

$$r = 2\sqrt{\frac{C_f M\ A\ K}{C_W}(K+2)}. \qquad (3\text{-}15)$$

Put $R = \frac{C_W}{C_f}$

$$\therefore r = 2\sqrt{\frac{MAK}{R}(K+2)}. \tag{3-16}$$

Substituting for M = 1 in the last equation, it would be

$$r = 2\sqrt{\frac{AK}{R}(K+2)}$$

which is the single channel case as shown in Equation (3-8). Therefore, the multiple channel indifferent range is the single channel indifferent range multiplied by $\sqrt{M}$.

CHAPTER IV

PROBABILISTIC WAITING LINE MODELS

In this chapter, the deterministic restriction on arrival time and service time does not apply. Ordinarily, both the arrival rate and the service rate are expected values from specified probability distributions. A probabilistic waiting line system will result; however, if either the arrival and/or the service time is a random variable. The arrival time and the service time can follow any distribution, empirical or analytical. It depends on the case under consideration for study. The models and their sensitivity analysis considered here have the characteristic that the arrival rate is assumed to be an expected value from a Poisson distribution. This assumption is mathematically convenient and has a sound practical basis in many situations. Also, the service rate is an expected value from Poisson distribution. The models are based on the assumption of an infinite population, in that the size of the population is large relative to the arrival rate.

In this case, individuals leaving the population do not significantly affect the arrival potential of the remaining units.

Poisson Arrivals With Exponential Service

Assume that both the arrival rate and the service rate are expected values from independent Poisson distributions. This assumption holds when the rates are independent of time, queue length, or any other property of waiting line system. The expected number of arrivals per period may be expressed as $1/A_m$ or λ. The expected number of service completions per period may be expressed as $1/S_m$ or μ. Where A_m and S_m are the mean time between arrivals and the mean service time in periods, for the assumed distributions, respectively. If the number of arrivals per period or the number of services per period have a Poisson distribution, then the time between arrivals A_x, or the service duration, S_x, will have an exponential distribution.[1] (It is assumed that μ is greater than λ, and that the arrival population is infinite.)

The Probability of n Units in the System

Under the foregoing assumptions, the probability that an arrival occurs between time t and time $t + \Delta t$ is $\lambda \Delta t$.[2]

---

[1]For a mathematical proof, see C. W. Churchman, R. L. Ackoff, and E. L. Arnoff, Introduction to Operations Research (New York: John Wiley and Sons, Inc.), 1957, pp. 398-400.

[2]See W. Feller, An Introduction to Probability Theory and Its Applications, 2nd ed. (New York: John Wiley and Sons, Inc.), 1957, p. 400.

Likewise, the probability that a service is completed in the time interval from t to $t + \Delta t$, given that a unit is being served at time t (conditional probability) is $\mu \Delta t$. Let

n = number of units in the system at time t

$P_n(t)$ = probability of n units in the system at time t.

If it is assumed that the probabilities of more than one unit arriving or being served during the small time interval Δt are negligible, and if $n \geq 1$, the probability that there will be n units in the system at time $t + \Delta t$ may be expressed as the sum of three independent compound probabilities as follows:

(1) The product of the probabilities that there are n units in the system at time t, no arrivals occur during time Δt, and no services are completed during time Δt, which is $[P_n(t)][1 - \lambda(\Delta t)][1 - \mu(\Delta t)]$.

(2) The product of the probabilities that there are $(n + 1)$ units in the system at time t, there is one unit serviced during time Δt, and there are no arrivals during time Δt, which is $[P_{n+1}(t)][\mu(\Delta t)][1 - \lambda(\Delta t)]$.

(3) The product of the probabilities that there are $n - 1$ units in the system at time t, there is one arrival during time Δt, and

there are no units serviced during Δt,

which is:

$[P_{n-1}(t)][\lambda(\Delta t)][1 - \mu(\Delta t)]$.

All other possibilities that might be enumerated will yield terms in Δt of higher order. These are assumed to be negligible.

The probability of n units in the system, for $n \geq 1$, at time $(t + \Delta t)$ is obtained by adding the preceding probabilities.

$$P_n(t+\Delta t) = [P_n(t)][1-\lambda(\Delta t)][1-\mu(\Delta t)] + [P_{n+1}(t)][\mu(\Delta t)][1-\lambda(\Delta t)] + [P_{n-1}(t)][\lambda(\Delta t)][1-\mu(\Delta t)].$$

Since the time interval Δt is small, the probabilities at time $t+\Delta t$ are equivalent to those at time t. By substituting $P_n(t)$ for $P_n(t+\Delta t)$, expanding, and dropping terms in Δt of higher order, the foregoing expression becomes

$$P_n(t) = P_n(t)[1-\lambda(\Delta t)-\mu(\Delta t)] + P_{n+1}(t)[\mu(\Delta t)] + P_{n-1}(t)[\lambda(\Delta t)]$$

$$P_{n+1}(t)[\mu(\Delta t)] = P_n(t) - P_n(t)[1-\lambda(\Delta t)-\mu(\Delta t)] - P_{n-1}(t)[\lambda(\Delta t)]$$

$$P_{n+1}(t) = P_n(t)\,\frac{\lambda+\mu}{\mu} - P_{n-1}(t)\,\frac{\lambda}{\mu}. \qquad (4\text{-}1)$$

The probability of no units in the system, $n=0$, at time $t + \Delta t$, is the sum of two independent probabilities as follows:

(1) The product of the probability that there are no units in the system at time t, and the probability that there are no arrivals during time Δt, which is $P_0(t)[1 - \lambda(\Delta t)]$.

(2) The product of the probabilities that there is one unit in the line at time t, then one unit is served during time Δt, and there are no arrivals during Δt, which is $P_1(t)[\mu(\Delta t)][1 - \lambda(\Delta t)]$.

All other possibilities that might be enumerated will yield terms in Δt of higher order. As before, these are assumed to be negligible.

The probability of no units in the system at time $t + \Delta t$ is obtained by adding the foregoing probabilities.

$$P_0(t + \Delta t) = P_0(t)[1 - \lambda(\Delta t)] + P_1(t)[\mu(\Delta t)][1 - \lambda(\Delta t)].$$

Since the time interval Δt is small, the probabilities at time $t + \Delta t$ are equivalent to those of time t. By substituting $P_n(t)$ for $P_n(t + \Delta t)$, expanding and dropping terms of Δt in higher order, the foregoing expression becomes

$$P_0(t) = P_0(t) - P_0(t)[\lambda(\Delta t)] + P_1(t)[\mu(\Delta t)]$$

$$P_1(t) = P_0(t)\,\frac{\lambda}{\mu}. \qquad (4\text{-}2)$$

Equations (4-1) and (4-2) may be solved by successive substitution for P_0 in terms of $P_0, P_1, P_2, \ldots, P_n$. Assuming that $P_n(t)$ is independent of t, and equal to P_n, results in:

$P_0 = P_0$

$P_1 = P_0(\frac{\lambda}{\mu})$ from Equation (4-2)

$P_2 = P_0(\frac{\lambda}{\mu})^2$ letting n = 1 in Equation (4-1) and substituting for P_1

$P_3 = P_0(\frac{\lambda}{\mu})^3$ letting n = 2 in Equation (4-1) and substituting for P_1

.
.
.
.
.

$P_n = P_0(\frac{\lambda}{\mu})^n$ letting n = n - 1 in Equation (4-1) and substituting for P_{n-1}.

Summing the left and the right sides of the preceding series results in the equality

$$\sum_{n=0}^{\infty} P_n = P_0 \sum_{n=0}^{\infty} (\frac{\lambda}{\mu})^n.$$

But it is obvious that

$$\sum_{n=0}^{\infty} P_n = 1.$$

And from the sum of an infinite geometric series

$$\sum_{n=0}^{\infty} (\frac{\lambda}{\mu})^n = \frac{1}{1 - (\lambda/\mu)}.$$

Therefore,

$$P_o\left[\frac{1}{1 - (\lambda/\mu)}\right] = 1$$

$$P_o = 1 - (\lambda/\mu).$$

Substituting this expression for P_o into the previous relationship for P_n gives

$$P_n = (1 - \frac{\lambda}{\mu})(\frac{\lambda}{\mu})^n \tag{4-3}$$

The Mean Number of Units in the System

The mean number of units in the system, n_m, may be expressed as

$$n_m = \sum_{n=0}^{\infty} n\ P_n$$

$$= \sum_{n=0}^{\infty} n(1 - \frac{\lambda}{\mu})(\frac{\lambda}{\mu})^n$$

$$= (1 - \frac{\lambda}{\mu}) \sum_{n=0}^{\infty} n(\frac{\lambda}{\mu})^n$$

$$= (1 - \frac{\lambda}{\mu})\left[\frac{\lambda}{\mu} + 2(\frac{\lambda}{\mu})^2 + 3(\frac{\lambda}{\mu})^3 + \ldots\right]. \tag{4-4}$$

Let:

$$g = \frac{\lambda}{\mu} + 2(\frac{\lambda}{\mu})^2 + 3(\frac{\lambda}{\mu})^3 + \ldots .$$

And let

$$g(\frac{\lambda}{\mu}) = (\frac{\lambda}{\mu})^2 + 2(\frac{\lambda}{\mu})^3 + \ldots .$$

Subtracting the second series from the first gives

$$g(1 - \frac{\lambda}{\mu}) + 1 = 1 + \frac{\lambda}{\mu} + (\frac{\lambda}{\mu})^2 + (\frac{\lambda}{\mu})^3 + \ldots .$$

The right hand side is now an infinite geometric series; therefore,

$$g(1 - \frac{\lambda}{\mu}) + 1 = \frac{1}{(1 - \lambda/\mu)}$$

$$g = \frac{\lambda/\mu}{[1 - (\lambda/\mu)]^2}.$$

Substituting for g in Equation (4-4) gives

$$n_m = \frac{\lambda}{\mu}\left[\frac{1}{1 - (\lambda/\mu)}\right]$$

$$n_m = \frac{\lambda}{\mu - \lambda} \qquad (4\text{-}5)$$

The Mean Waiting Time

The expected time an arrival spends in the system, W_m, can be shown to be:

$$W_m = \frac{n_m}{\lambda}.$$

Substituting Equation (4-5) for n_m gives

$$W_m = \frac{1}{\mu - \lambda}. \qquad (4\text{-}6)$$

The Expected Total System Cost

The expected total system cost per period is the sum of the expected waiting cost per period and the expected facility cost per period; that is,

$$TC_m = WC_m + FC_m.$$

The expected waiting cost per period is the product of the cost of waiting per unit per period and the mean number of units in the system during the period, or

$$\begin{aligned} WC_m &= C_W(n_m) \\ &= \frac{C_W \lambda}{\mu - \lambda}. \end{aligned}$$

The expected service cost per period is the product of the cost of providing service facility of unit rate capacity and the service rate in units per period, or

$$FC_m = C_f \, \mu .$$

The expected total cost per period is the sum of these components and may be expressed as

$$TC_m = \frac{C_W \lambda}{(\mu - \lambda)} + C_f \, \mu. \qquad (4\text{-}7)$$

Therefore, the expected total cost is a function of the two variables, service rate, μ, and arrival rate, λ. The graphical representation for TC_m is shown in Figure 17.

At the plane $\lambda = \lambda_1$, which is parallel to $TC_m - \mu$ plane, the minimum expected total cost, TC_{m_1}, is occurred at $\mu = \mu_1$. At the plane $\lambda = \lambda_2$, the minimum expected

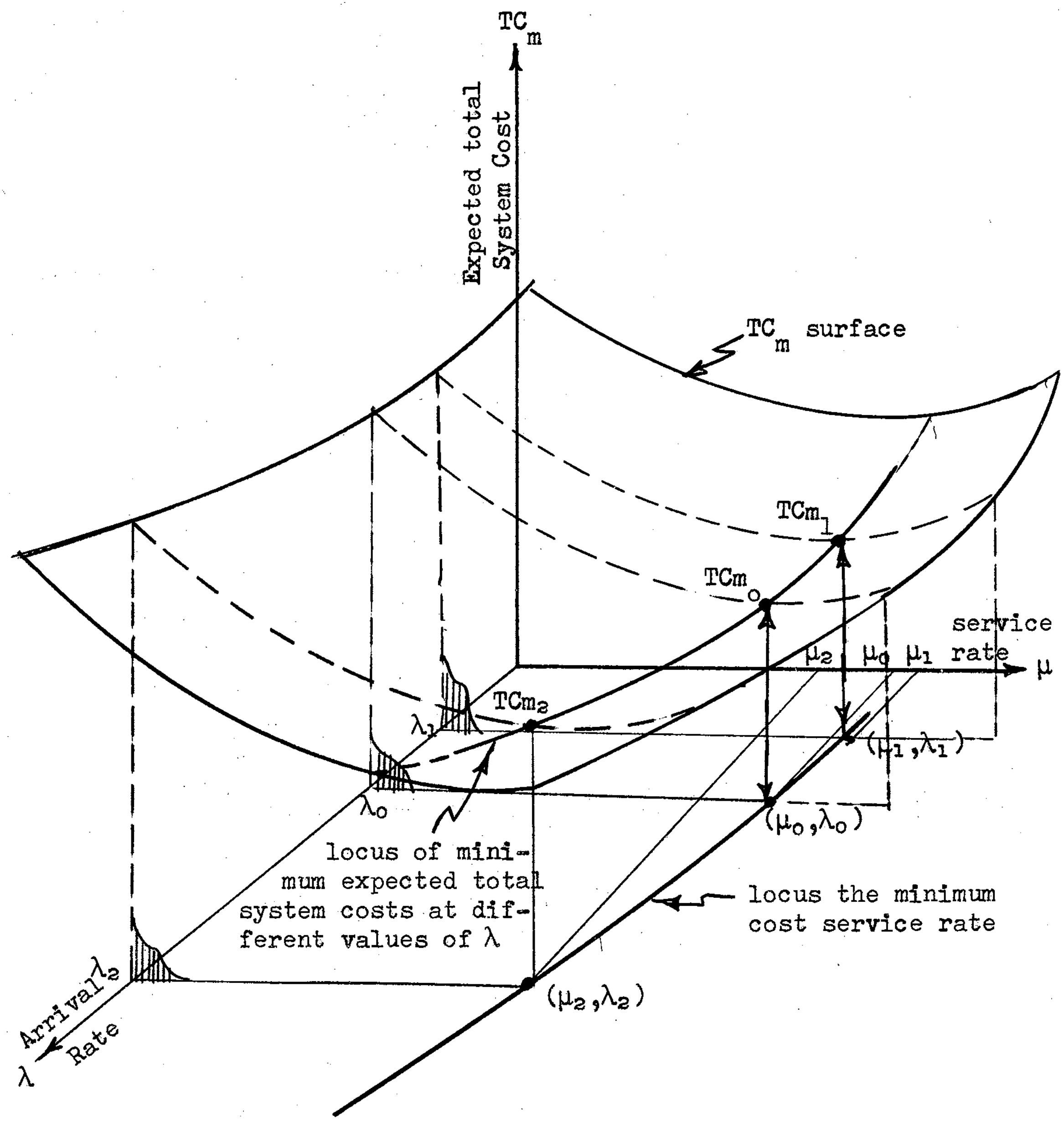


Figure 17. Expected Total System Cost Surface

total cost, TC_{m_2}, is occurred at $\mu = \mu_2$. The minimum cost service rate, μ, at certain arrival rate can be calculated by differentiating partially the expected total system cost equation, TC_m, with respect to μ and equating the result to zero, or

$$\frac{\partial TC_m}{\partial \mu} = -C_W \lambda(\mu - \lambda)^{-2} + C_f = 0$$

$$\therefore C_f = C_W \lambda(\mu - \lambda)^{-2}$$

$$C_f(\mu - \lambda)^2 = C_W \lambda$$

$$(\mu - \lambda)^2 = \frac{C_W \lambda}{C_f}$$

$$\mu - \lambda = \sqrt{\frac{C_W \lambda}{C_f}}$$

$$\mu = \lambda + \sqrt{\frac{C_W \lambda}{C_f}}. \qquad (4\text{-}8)$$

Equation (4-8) is the locus of the minimum cost service rate for different values of λ, C_W, and C_f. The locus of the minimum expected total system cost at different values of λ can be found by substituting from Equation (4-8) in Equation (4-7). This is shown in Equation (4-9).

Let: TC_{m_i} = the minimum expected total cost at the i^{th} arrival rate, λ_i, and at the i^{th} service rate, μ_i.

$$TC_{m_i} = \frac{C_W \lambda_i}{\left[\lambda_i + \sqrt{\frac{C_W}{C_f} \lambda_i} - \lambda_i\right]} + C_f\left[\lambda_i + \sqrt{\frac{C_W \lambda_i}{C_f}}\right]$$

$$= \frac{C_W \lambda_1}{\sqrt{\frac{C_W \lambda_1}{C_f}}} + C_f \lambda_1 + C_f \sqrt{\frac{C_W \lambda_1}{C_f}}$$

$$TC_{m_1} = C_W \lambda_1 \sqrt{\frac{C_f}{C_W \lambda_1}} + C_f \lambda_1 + \sqrt{C_W C_f \lambda_1}$$

$$= \sqrt{C_W C_f \lambda_1} + C_f \lambda_1 + \sqrt{C_W C_f \lambda_1}$$

$$TC_{m_1} = C_f \lambda_1 + 2\sqrt{C_W C_f \lambda_1} \qquad (4\text{-}9)$$

Equation (4-9) is shown graphically in Figure 17 as the curve drawn on the TC_m-surface. The allocation of the minimum of the minimum expected total system costs, TC_{m_o}, can be found by differentiating partially Equation (4-9) with respect to λ_1 and equating to zero or

$$\frac{\partial TC_{m_1}}{\partial \lambda_1} = C_f + 2\sqrt{C_W C_f} \frac{1}{2\sqrt{\lambda_1}} = 0$$

$$\sqrt{\frac{C_f C_W}{\lambda_1}} = -C_f .$$

Square both sides it would be

$$\frac{C_f C_W}{\lambda_1} = C_f^2$$

$$\lambda_1 = \frac{C_W}{C_f} .$$

Let TC_{m_o} occur at $\lambda_1 = \lambda_o$

$$\therefore \lambda_o = \frac{C_W}{C_f} \qquad (4\text{-}10)$$

The minimum of the minimum expected total costs can be calculated bu substituting from (4-10) in Equation (4-9).

$$TC_{m_o} = C_f \frac{C_W}{C_f} + 2\sqrt{C_W Cf \frac{C_W}{C_f}}$$

$$= C_W + 2\ C_W$$

$$TC_{m_o} = 3C_W. \qquad (4\text{-}11)$$

$\therefore$ TC_{m_o} is a function of C_W, the waiting cost/unit/ period. It is equal to the trible value of C_W and it occurs at arrival rate $\lambda_o = \frac{C_W}{C_f}$, and service rate μ_o where μ_o can be calculated from Equation (4-8) as follows:

$$\mu_o = \lambda_o + \sqrt{\frac{\lambda_o\ C_W}{C_f}}$$

$$\mu_o = \frac{C_W}{C_f} + \sqrt{\left(\frac{C_W}{C_f}\right)^2}$$

$$\mu_o = 2\left(\frac{C_W}{C_f}\right). \qquad (4\text{-}12)$$

To sum up,

$$\left.\begin{aligned} TC_{m_o} &= 3C_W \\ \text{occurs at} \quad \lambda_o &= \frac{C_W}{C_f} \\ \mu_o &= 2\left(\frac{C_W}{C_f}\right). \end{aligned}\right\} \qquad (4\text{-}13)$$

This is shown in Figure 18.

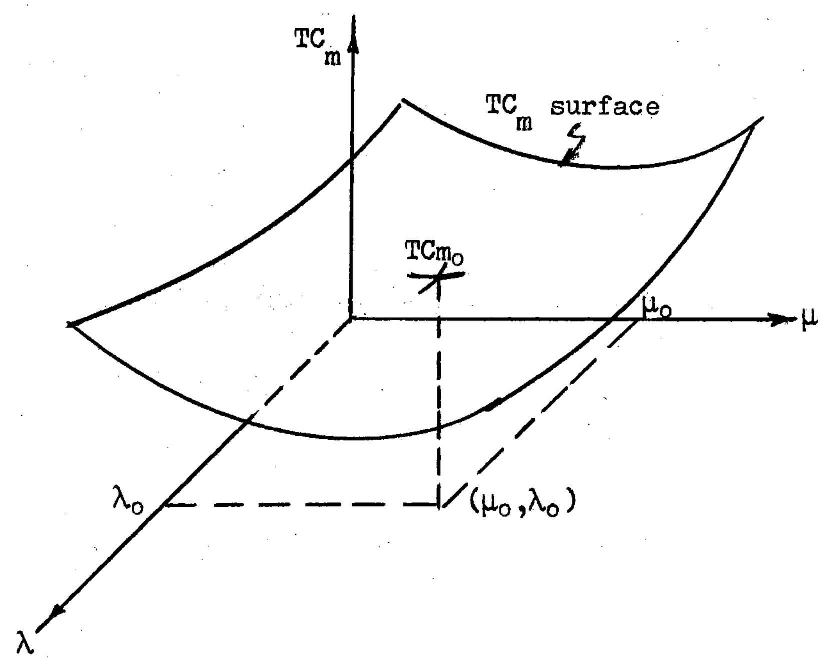


Figure 18. Description of Optimal Case

<u>Sensitivity Analysis</u>

For the calculation of the insignificant limits, consider the same definition of K as previously stated where

$$K = \frac{TC_m - TC_{m_o}}{TC_{m_o}} = \frac{\Delta TC_m}{TC_{m_o}}$$

$$\therefore (TC_{m_o})K = \Delta TC_m$$

$$(3C_W)K = \left[\frac{C_W \lambda}{(\mu - \lambda)} + C_f \mu\right] - 3C_W$$

$$3(K + 1)C_W = \frac{C_W\lambda}{(\mu - \lambda)} + C_f\mu.$$

Multiplying both sides by $(\mu - \lambda)$, it would be

$$3(K + 1)(\mu - \lambda)C_W = C_W\lambda + C_f(\mu - \lambda)\mu$$

$$3(K + 1)C_W\mu - 3C_W(K + 1)\lambda = C_W + C_f\mu^2 - C_f\lambda\mu$$

$$C_f\mu^2 - [3C_W(1 + K) + \lambda C_f]\mu + [3C_W(1 + K) + C_W]\lambda = 0 \tag{4-14}$$

The previous equation is quadratic; its solution would be:

$$\mu = \frac{[3C_W(1 + K) + \lambda C_f] \pm \sqrt{[3C_W(1 + K) + \lambda C_f]^2 - 4C_f\lambda[3C_W(1 + K) + C_W]}}{2C_f}$$

$$\mu = \frac{[3C_W(1 + K) + \lambda C_f] \pm \sqrt{9C_W^2(1 + K)^2 + \lambda^2 C_f^2 + 6C_W C_f\lambda(1 + K) - 12C_f\lambda C_W(1 + K) - 4C_f C_W\lambda}}{2C_f}$$

$$\mu = \frac{[3C_W(1 + K) + \lambda C_f] \pm \sqrt{9C_W^2(1 + K)^2 + \lambda^2 C_f^2 - 6C_f\lambda C_W(1 + K) - 4C_f C_W\lambda}}{2C_f}$$

$$\mu = \frac{1}{2C_f}[3C_W(1 + K) + \lambda C_f] \pm \frac{1}{2C_f}\sqrt{9C_W^2(1 + K)^2 + \lambda^2 C_f^2 - (6 + 6K + 4)C_f\lambda C_W}$$

$$\mu = \frac{1}{2C_f}[3C_W(1 + K) + \lambda C_f] \pm \frac{1}{2C_f}\sqrt{9C_W^2(1 + K)^2 + \lambda^2 C_f^2 - (10 + 6K)C_f\lambda C_W}$$

$$\mu = \left[\frac{3}{2}\frac{C_W}{C_f}(1 + K) + \frac{1}{2}\lambda\right] \pm \sqrt{\frac{9}{4}\left(\frac{C_W}{C_f}\right)^2(1 + K)^2 + \frac{\lambda^2}{4} - \frac{1}{4}(10 + 6K)\left(\frac{C_W}{C_f}\right)\lambda}$$

Since the previous equation is to calculate the insignificant limits of the service rate at the plane $\lambda = \lambda_0$; therefore, the value of λ can be substituted in the last equation by the value of λ_0, which is equal to $\frac{C_W}{C_f}$. By substituting for λ it gives,

$$\mu = \left[\frac{3}{2}\frac{C_W}{C_f}(1+K) + \frac{1}{2}\frac{C_W}{C_f}\right] \pm \sqrt{\frac{9}{4}\left(\frac{C_W}{C_f}\right)^2(1+K)^2 + \frac{1}{4}\left(\frac{C_W}{C_f}\right)^2 - \frac{1}{4}(10+6K)\left(\frac{C_W}{C_f}\right)^2}$$

$$\mu = \frac{C_W}{C_f}\left[\frac{3}{2} + \frac{3}{2}K + \frac{1}{2}\right] \pm \sqrt{\left(\frac{C_W}{C_f}\right)^2\left[\frac{9}{4}(1+K)^2 + \frac{1}{4} - \frac{1}{4}(10+6K)\right]}$$

$$\mu = \frac{C_W}{C_f}\left[(2 + \frac{3}{2}K) \pm \sqrt{\frac{9}{4} + \frac{9}{2}K + \frac{9}{4}K^2 + \frac{1}{4} - \frac{10}{4} - \frac{6}{4}K}\right]$$

$$\mu = \frac{C_W}{C_f}\left[(2 + \frac{3}{2}K) \pm \sqrt{(\frac{3}{2}K)^2 + \frac{12}{4}K}\right]$$

$$\left.\begin{aligned} \mu &= \frac{1}{2}\frac{C_W}{C_f}\left[(4+3K) \pm \sqrt{K(9K+12)}\right] \\ \mu_h &= \frac{1}{2}\frac{C_W}{C_f}\left[(4+3K) + \sqrt{K(9K+12)}\right] \\ \mu_L &= \frac{1}{2}\frac{C_W}{C_f}\left[(4+3K) - \sqrt{K(9K+12)}\right] \end{aligned}\right\} \qquad (4\text{-}15)$$

Figure 19 shows the upper and lower limits of the service rate μ. They are calculated at $\lambda = \lambda_o$ plane. That is the minimum expected total system cost at that plane is the lowest cost on the surface. The insignificant limits for this case would be for the minimum service rate μ_o. These limits shown in the figure are the graphical presentation for Equation (4-15).

Indifferent Range

The indifferent range can be calculated for the previous case by subtracting the higher limit from the lower

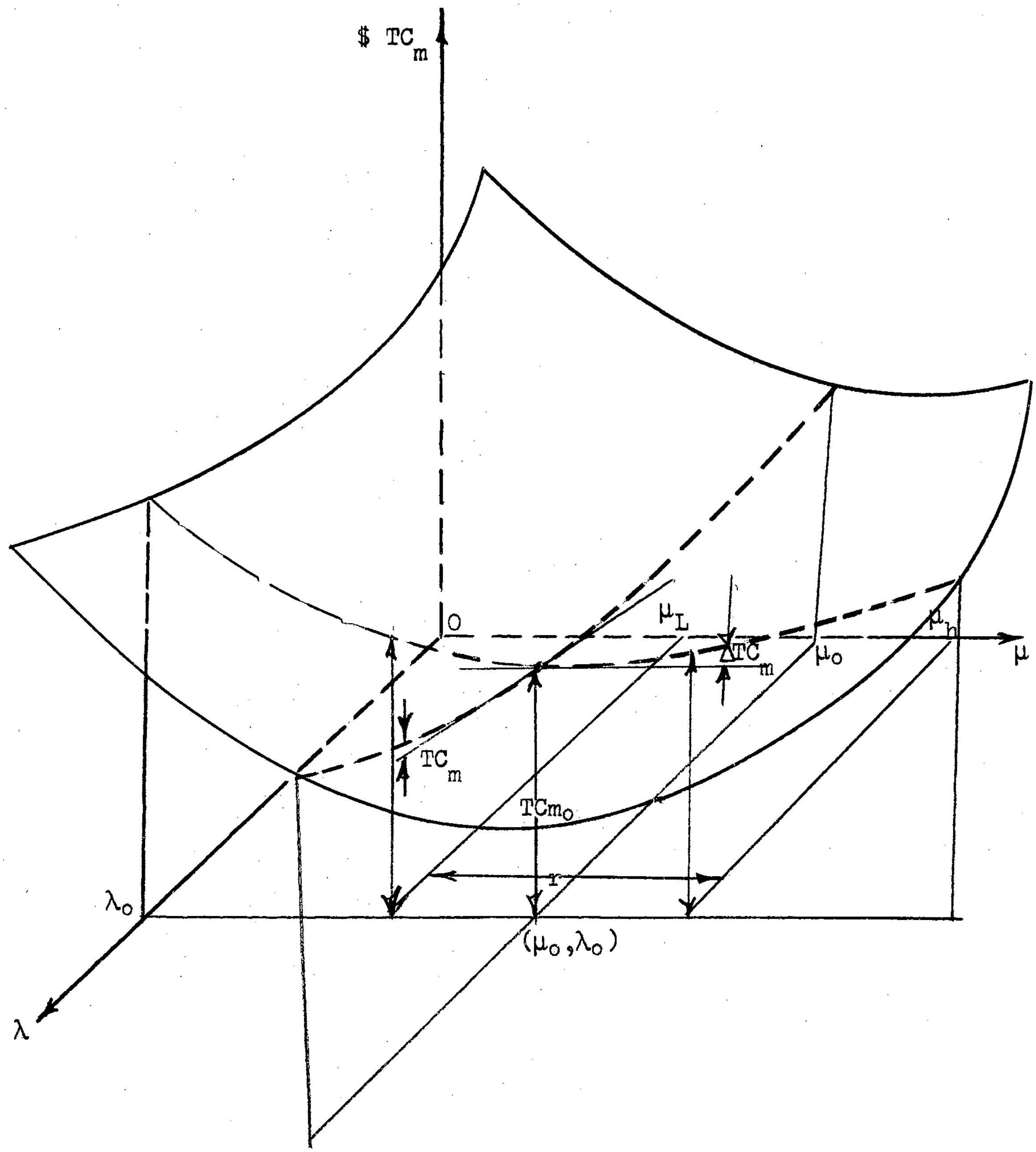


Figure 19. The Insignificant Limits From the Lowest Expected Total System Cost on the Surface at λ_o-plane

limit, or

$$r = \mu_h - \mu_L$$

$$= \frac{1}{2}\frac{C_W}{C_f}\left[(4 + 3K) + \sqrt{K(9K + 12)} - (4 + 3K) + \sqrt{K(9K + 12)}\right]$$

$$= \frac{1}{2}\frac{C_W}{C_f}\left[2\sqrt{K(9K + 12)}\right]$$

$$r = \frac{C_W}{C_f}\sqrt{K(9K + 12)}. \qquad (4\text{-}16)$$

As an illustration, consider the case where the cost of waiting per unit period, C_W, is \$4.00 and the service facility cost for serving one unit, C_f, is \$2.00. For this situation, the best arrival rate and service rate to meet the lowest expected total system cost applying Equation (4-13) is as follows:

$$\lambda_o = \frac{C_W}{C_f} = \frac{4}{2} = 2 \text{ units/period}$$

$$\mu_o = 2\left(\frac{C_W}{C_f}\right) = 2\left(\frac{4}{2}\right) = 4 \text{ units/period.}$$

By this policy the lowest expected total cost per period is:

$$TC_{m_o} = 3C_W = 3(4) = \$12.$$

The insignificant limits for the service rate at $K = 0.04$ is:

$$\mu = \frac{1}{2}\left(\frac{C_W}{C_f}\right)\left[(4 + 3K) \pm \sqrt{K(9K + 12)}\right]$$

$$\mu_h = \frac{1}{2}\left(\frac{4}{2}\right)\left[(4 + 0.12) + \sqrt{0.04(.36 + 12)}\right] \approx 5$$

$$\mu_L = \frac{1}{2}\left(\frac{4}{2}\right)\left[(4.12) - \sqrt{0.04(12.36)}\right] \approx 3$$

and the indifferent range, r, is

$$r = \mu_h - \mu_L = 2$$

or this result can be obtained by following Equation (4-16).

The insignificant limits and indifferent range, shown in Equation (4-15) and (4-16), respectively, are calculated at the plane $\lambda = \lambda_o$ passing through the lowest cost TC_{m_o}.

In the following section, consideration is given for the case where the plane λ is not passing through the lowest cost TC_{m_o}. This is the general case. The calculation for the insignificant limits and indifferent range will be related also to the lowest cost, TC_{m_o}, which is easy to calculate.

General Case

The general case is shown in Figure 20. This is the case where it is required to calculate the insignificant limits and/or the indifferent range for any arrival rate λ. In other words, the insignificant limits and

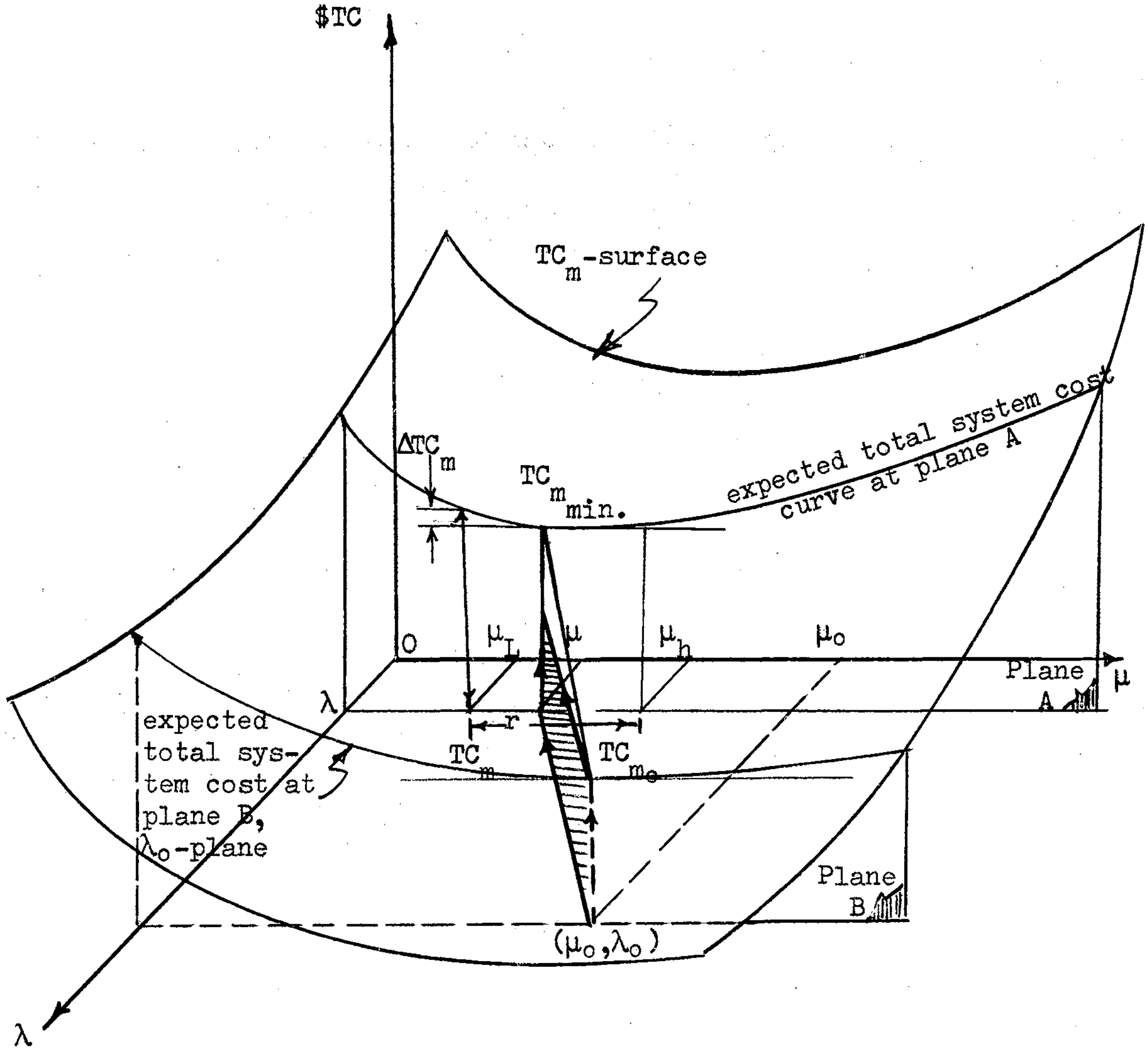


Figure 20. The Insignificant Limits From the Minimum Expected Total System Cost at λ-plane

range are to be calculated at the plane A. Let plane A be at a certain arrival rate λ, or

$$\lambda = K_2 \lambda_o$$

where λ_o = the arrival rate at which the lowest expected total cost, TC_{m_o}, occurs

K_2 = constant value

since $\lambda_o = \frac{C_W}{C_f}$

$$\therefore \lambda = K_2 \frac{C_W}{C_f}$$

$$K_2 = \frac{C_f}{C_W} \lambda \qquad (4\text{-}17)$$

The numerical value of K_2 may be less or greater than one and it is always positive or zero. It depends on the values of C_f, C_W, and λ.

The minimum expected total system cost value at the plane A, $TC_{m_{min.}}$, can be calculated by following Equation (4-9), or

$$TC_{min.} = 2\sqrt{C_W C_f \lambda} + C_f \lambda .$$

Let the lowest expected total system cost, TC_{m_o}, occur on the curve of intersection between plane B and TC_m surface. Since TC_{m_o} is the lowest point on the surface, therefore $TC_{m_{min.}}$ is always greater than TC_{m_o}. Let

K_1 = the ratio of $TC_{m_{min.}}$ and TC_{m_o}

i.e. $$K_1 = \frac{TC_{m_{min.}}}{TC_{m_o}}.$$

Substitute for $TC_{m_{min.}}$ and TC_{m_o} values from Equations (4-9) and (4-11) in the previous equation, it would be

$$K_1 = \frac{1}{3C_W}\left[2\sqrt{C_W C_f \lambda} + C_f\lambda\right] \qquad (4\text{-}18)$$

where $K_1 \geq 1$.

In the case where $K = 1$ and $K_2 = 1$, plane A coincides on plane B and $TC_{m_{min.}}$ is equal to and coincides on TC_{m_o}.

Following the same procedure as in the last section,

$$\therefore K = \frac{TC_m - TC_{m_{min.}}}{TC_{m_{min.}}}.$$

Since $$TC_{m_{min.}} = K_1 TC_{m_o}$$

$$\therefore K = \frac{TC_m - K_1 TC_{m_o}}{K_1 TC_{m_o}}$$

$$KK_1 TC_{m_o} = TC_m - K_1 TC_{m_o}.$$

But $$TC_{m_o} = 3C_W$$

and $$TC_m = \frac{C_W\lambda}{\mu - \lambda} + C_f\mu$$

$$\therefore\ 3KK_1 C_W = \frac{C_W \lambda}{\mu - \lambda} + C_f \mu - 3K_1 C_W$$

$$3K_1 C_W (K+1) = \frac{C_W \lambda}{\mu - \lambda} + C_f\ \mu .$$

Multiply both sides by $(\mu - \lambda)$

$$3K_1 C_W (K+1)(\mu - \lambda) = C_W \lambda + C_f \mu (\mu - \lambda)$$

$$3K_1 C_W (K+1)\mu - 3K_1 C_W (K+1)\lambda = C_W \lambda + C_f \mu^2 - C_f \mu \lambda .$$

This equation leads to the following:

$$C_f \mu^2 - [3K_1 C_W (K+1) + C_f \lambda]\mu + [3K_1 C_W (K+1) + C_W]\lambda = 0.$$

The preceding equation is second degree equation in μ. By solving the quadratic equation, μ would be

$$\mu = \frac{\left[3K_1 C_W (K+1) + C_f \lambda\right] \pm \sqrt{\left[3K_1 C_W (K+1) + C_f \lambda\right]^2 - 4\lambda C_f \left[3K_1 C_W (K+1) + C_W\right]}}{2C_f}$$

$$\mu = \frac{1}{2C_f}\left[3K_1 C_W (K+1) + C_f \lambda\right] \pm \frac{1}{2C_f}\sqrt{9K^2 C_W^2 (K+1)^2 + C_f^2 \lambda^2 + 6K_1 C_W (K+1) C_f \lambda - 12\lambda C_f K_1 (K+1) C_W - 4\lambda C_W C_f}$$

$$= \frac{1}{2C_f}\left[3K_1 C_W (K+1) + C_f \lambda\right] \pm \frac{1}{2C_f}\sqrt{9K_1^2 C_W^2 (K+1)^2 + C_f^2 \lambda^2 - 6\lambda C_f K_1 C_W (K+1) - 4\lambda C_f C_W}$$

$$= \frac{1}{2C_f}\left[3K_1 C_W (K+1) + C_f \lambda\right] \pm \sqrt{\frac{9}{4} K_1^2 \left(\frac{C_W}{C_f}\right)(K+1)^2 + \frac{\lambda^2}{4} - \frac{6}{4}\lambda\ K_1 \left(\frac{C_W}{C_f}\right)(K+1) - \lambda\left(\frac{C_W}{C_f}\right)}$$

Substitute for λ from Equation (4-17) in the previous equation, or

$$\mu = \frac{C_W}{2C_f}\left[3K_1 (K+1) + K_2 \left(\frac{C_f}{C_W}\right)\left(\frac{C_W}{C_f}\right)\right] \pm \sqrt{\frac{9}{4}K_1 \left(\frac{C_W}{C_f}\right)^2 (K+1)^2 + \frac{K_2^2}{4}\left(\frac{C_W}{C_f}\right)^2 - \frac{6}{4}K_1 K_2 \left(\frac{C_W}{C_f}\right)^2 (K+1) - K_2\left(\frac{C_W}{C_f}\right)^2}$$

$$\mu = \frac{1}{2}\left(\frac{C_W}{C_f}\right)\left[3K_1 (K+1) + K_2 \pm \sqrt{9K_1^2 (K+1)^2 + K_2^2 - 6K_1 K_2 (K+1) - 4K_2}\right]$$

$$= \frac{1}{2}\left(\frac{C_W}{C_f}\right)\left[3K_1(K+1) + K_2 \pm \sqrt{9K_1^2(K+1)^2 - 6K_1K_2(K+1) + K_2(K_2-4)}\right]$$

from which

$$\mu_h = \frac{1}{2}\left(\frac{C_W}{C_f}\right)\left[3K_1(K+1) + K_2 + \sqrt{9K_1^2(K+1)^2 - 6K_1K_2(K+1) + K_2(K_2-4)}\right]$$

$$\mu_L = \frac{1}{2}\left(\frac{C_W}{C_F}\right)\left[3K_1(K+1) + K_2 - \sqrt{9K_1^2(K+1)^2 - 6K_1K_2(K+1) + K_2(K_2-4)}\right] \qquad (4\text{-}19)$$

where:

$$K = \frac{\Delta TC_m}{TC_{m_{min.}}}$$

$$K_1 = \frac{1}{3C_W}[2\sqrt{C_W C_f \lambda} + C_f\lambda]$$

$$K_2 = \frac{C_f}{C_W}\lambda .$$

Indifferent Range

Since $r = \mu_h - \mu_L$.

Substitute in the preceding equation from (4-19), r would be

$$r = \frac{C_W}{C_f}\sqrt{9K_1^2(K+1)^2 - 6K_1K_2(K+1) + K_2(K_2-4)}. \qquad (4\text{-}20)$$

By putting $K_1 = 1$, and $K_2 = 1$, Equation (4-19) reduces to Equation (4-15) and Equation (4-20) reduces to Equation (4-16), which is the special case for the calculation of insignificant limits and range at the plane $\lambda = \lambda_o$.

In the last illustration, the insignificant limits and the indifferent range are calculated when the situation is working at arrival rate λ_o. Consider the same situation but working at different arrival rate, $\lambda = 8$. Therefore, the optimum service rate here, at which the minimum expected total system cost occurs, can be calculated by following Equation (4-8), or

$$\mu = \lambda + \sqrt{\frac{C_W \lambda}{C_f}}$$

$$= 8 + \sqrt{\frac{4 \times 8}{2}} = 8 + 4 = 12.$$

At this service rate, the minimum expected total cost can be calculated by following Equation (4-9), or

$$TC_{m_{min.}} = C_f \lambda + 2\sqrt{C_W C_f \lambda}$$

$$= 2(8) + 2\sqrt{4(2)8}$$

$$= 16 + 2(8) = \$32 \text{ per period.}$$

Now, to calculate the insignificant limits and the indifferent range, K_1 and K_2 should be calculated as follows:

$$K_1 = \frac{32}{12} = 2.67$$

$$K_2 = \frac{C_f}{C_W} \lambda = \frac{2}{4}(8) = 4.$$

Apply Equation (4-19), μ_h and μ_L can be calculated by assuming $K = 0.04$ as follows:

$$\mu_h = \frac{1}{2}\left(\frac{4}{2}\right)\left[3\left(\frac{32}{12}\right)(1.04) + 4 + \right.$$

$$\left.\sqrt{9\left(\frac{32}{12}\right)^2(1.04)^2 - 6\left(\frac{32}{12}\right)(4)(1.04) + 0}\right]$$

$$= 8.32 + 4 + \sqrt{69.248 - 66.560}$$

$$= 12.32 + \sqrt{2.688} = 12.32 + 1.68 \approx 14$$

$$\mu_L = 12.32 - 1.68 \approx 11.$$

And the indifferent range = 14 - 11 = 3.

In other words, if the service rate for the considered case other than the optimum one, 12 units/period, and the decision maker does not care about 4% of the optimum total cost, as difference in costs between the working situation and the optimum one, he should decide to change to the 12 units as service rate if μ is beyond the insignificant limits 11 → 14 and leave the situation as it is if μ within the limits.

CHAPTER V

POISSON ARRIVALS WITH CONSTANT SERVICE TIME

When service is provided automatically by mechanical means, or when the service operation is mechanically paced, the service duration might be a constant. Under these conditions, the service time distribution has a variance of zero. The mean number of units in the system is given by

$$n_m = \frac{(\lambda/\mu)^2}{2[1 - (\lambda/\mu)]} + \left(\frac{\lambda}{\mu}\right). \qquad (5\text{-}1)$$

And the mean waiting time is

$$w_m = \frac{\lambda/\mu}{2\mu[1 - (\lambda/\mu)]} + \frac{1}{\mu}. \qquad (5\text{-}2)$$

The Expected Total System Cost

The expected total system cost per period is the sum of the expected waiting cost per period and the expected facility cost per period; that is,

$$TC_m = WC_m + FC_m.$$

The expected waiting cost per period is the product of the cost of waiting per unit per period and the mean number of units in the system during the period, or

$$WC_m = C_W n_m$$

$$= C_W\left\{\frac{(\lambda/\mu)^2}{2[1 - (\lambda/\mu)]} + \frac{\lambda}{\mu}\right\}. \qquad (5\text{-}3)$$

The expected facility cost per period is the product of the cost of providing service facility of unit capacity and the service rate in units per period.

$$FC_m = C_f(\mu). \qquad (5\text{-}4)$$

The expected total system cost per period is the sum of these cost components and may be expressed as

$$TC_m = C_W\left\{\frac{(\lambda/\mu)^2}{2[1 - (\lambda/\mu)]} + \frac{\lambda}{\mu}\right\} + C_f\mu. \qquad (5\text{-}5)$$

Define x as the load factor, which is the ratio of the arrival rate, λ, to the potential service rate, μ. Substitute for λ/μ in Equation (5-5), it would be

$$TC_m = C_W\left\{\frac{x^2}{2(1 - x)} + x\right\} + \frac{C_f\lambda}{x} \qquad (5\text{-}6)$$

The Minimum Expected Total System Cost

To find out the minimum expected total system cost load factor, differentiate the expected total system cost Equation (5-6) with respect to x and equate to zero, Equation (5-6) would be:

$$\frac{dTC_m}{dx} = C_W\left[\frac{2(1-x) - 2x + 2x^2}{4(1-x)^2} + 1\right] - \frac{C_f\lambda}{x^2} = 0$$

$$C_W\left[4x - 4x^2 + 2x^2 + 4(1-x)^2\right]x^2 - 4C_f\lambda(1-x)^2 = 0$$

$$C_W\left[4x - 4x^2 + 2x^2 + 4(1 - 2x + x^2)\right]x^2 - 4C_f\lambda(1-x)^2 = 0$$

$$C_W\left[4x - 4x^2 + 2x^2 - 8x + 4x^2 + 4\right]x^2 - 4C_f\lambda(1 + x^2 - 2x) = 0$$

$$C_W x^2(2x^2 - 4x + 4) - 4C_f\lambda(1 + x^2 - 2x) = 0.$$

Divide by C_f

$$2\frac{C_W}{C_f}x^4 - 4\frac{C_W}{C_f}x^3 + 4\frac{C_W}{C_f}x^2 - 4\lambda - 4\lambda x^2 + 8\lambda x = 0$$

let $R = \frac{C_W}{C_f}$

$$2Rx^4 - 4Rx^3 + 4Rx^2 - 4\lambda - 4\lambda x^2 + 8\lambda x = 0$$

divide by 2R

$$x^4 - 2x^3 + 2\left(1 - \frac{\lambda}{R}\right)x^2 + \frac{4\lambda}{R}x - \frac{2\lambda}{R} = 0 \qquad (5\text{-}7)$$

let:

$$p = -2$$

$$q = 2\left(1 - \frac{\lambda}{R}\right)$$

$$V = \frac{4\lambda}{R}$$

$$W = -\frac{2\lambda}{R}.$$

Substitute these values in Equation (5-7), it would be

$$x^4 + px^3 + qx^2 + Vx + W = 0. \qquad (5\text{-}8)$$

By solving Equation (5-8), the $X_{min.}$ can be found. To solve for X, first determine a, b, and c such that

$$x^4 + px^3 + qx^2 + VX + W + (ax + b)^2 = (x^2 + {}^p/_2\, x + c)^2 . \qquad (5\text{-}9)$$

The determination of a, b, and c is accomplished by equating the coefficient of like powers of x in the first and second numbers of Equation (5-9)

$$a^2 + q = 2C + {}^{p^2}/4 \qquad (5\text{-}10)$$

$$2ab + V = Cp \qquad (5\text{-}11)$$

$$b^2 + W = C^2 . \qquad (5\text{-}12)$$

Hence from the Equations (5-10), (5-11), and (5-12) it would be

$$(Cp - v)^2 = 4\ a^2 b^2 = 4(2C + {}^{p^2}/4 - q)(C^2 - W)$$

or

$$C^3 - {}^q/_2\, C^2 + \frac{1}{4}(pV - 4W)C + \frac{1}{8}(4qW - p^2 W - V^2) = 0.$$

(5-13)

Let:

$$F = -\frac{q}{2} = \frac{\lambda}{R} - 1$$

$$Y = \frac{1}{4}(pV - 4W)$$

$$= \frac{1}{4}\left[(-2)\left(\frac{4\lambda}{R}\right) - 4\left(\frac{-2\lambda}{R}\right)\right] = 0$$

$$E = \frac{1}{8}(4qW - p^2 W - V^2)$$

$$= \frac{1}{8}\left[8\left(1 - \frac{\lambda}{R}\right)\left(\frac{-2\lambda}{R}\right) - 4\left(\frac{-2\lambda}{R}\right) - \frac{16\lambda^2}{R^2}\right]$$

$$= \frac{1}{8}\left[-\frac{16\lambda}{R} + \frac{16\lambda^2}{R^2} + \frac{8\lambda}{R} - 16\frac{\lambda^2}{R^2}\right] = -\frac{\lambda}{R}.$$

Substitute in Equation (5-13), it would be

$$C^3 + FC^2 + YC + E = 0.$$

Since $Y = 0$, the last equation is

$$C^3 + FC^2 + E = 0. \qquad (5\text{-}14)$$

C can be found by solving Equation (5-14) and then obtain a and b by substitution in Equations (5-10) and (5-11).

A method of solving Equation (5-14) will now be explained. Equation (5-14) can first be transformed so as to remove the second degree term. Let

$$C = Z - \frac{F}{3}. \qquad (5\text{-}15)$$

Substitute in Equation (5-14) from (5-15) it would be

$$\left(Z - \frac{F}{3}\right)^3 + F\left(Z - \frac{F}{3}\right)^2 + E = 0$$

$$\left(Z - \frac{F}{3}\right)\left(Z^2 - \frac{2}{3}FZ + \frac{F^2}{9}\right) + F\left(Z^2 - \frac{2}{3}FZ + \frac{F^2}{9}\right) + E = 0$$

$$Z^3 - \frac{2}{3}FZ^2 + \frac{F^2 Z}{9} - \frac{F}{3}Z^2 + \frac{2}{9}F^2 Z - \frac{F^3}{27} + FZ^2 - \frac{2}{3}F^2 Z + \frac{F^3}{9} + E = 0$$

$$Z^3 - \frac{3F^2}{9}Z - \frac{F^3}{27} + \frac{F^3}{9} + E = 0$$

$$Z^3 - \frac{3}{9} F^2 Z + (\frac{2F^3}{27} + E) = 0$$

$$Z^3 + 3HZ + G = 0 \qquad (5\text{-}16)$$

where $$H = \frac{-F^2}{9} = -\frac{1}{9}(\frac{\lambda}{R} - 1)^2 \qquad (5\text{-}17)$$

and $$G = \frac{2F^3 + 27E}{27} = \frac{2}{27}\left(\frac{\lambda}{R} - 1\right)^3 - \frac{\lambda}{R}. \qquad (5\text{-}18)$$

It is customary to refer to (5-16) as the reduced cubic equation.

The roots of the given Equation (5-14) can be found from (5-15) when the roots of (5-16) are known.

Equation (5-16) can be solved for Z by Cardan's formulas as follows:

$$Z = \left[\frac{-G + \sqrt{G^2 + 4H^3}}{2}\right]^{1/3} + \left[\frac{-G - \sqrt{G^2 + 4H^3}}{2}\right]^{1/3}. \qquad (*)$$

To get the solution for c substitute Z in Equation (5-15), it would be

$$C = \left[\frac{-G + \sqrt{G^2 + 4H^3}}{2}\right]^{1/3} + \left[\frac{-G - \sqrt{G^2 + 4H^3}}{2}\right]^{1/3} - \frac{F}{3}. \qquad (5\text{-}19)$$

Substitute from Equations (5-17) and (5-18) in Equation (5-19), it would be

(*) Introduction to the Theory of Equations by Conkwright, Ginn and Company, 1941, pp. 70-71.

$$c = \left[-\frac{1}{2}\left\{\frac{2}{27}\left(\frac{\lambda}{R} - 1\right)^3 - \frac{\lambda}{R}\right\} + \frac{1}{2}\sqrt{\left\{\frac{2}{27}\left(\frac{\lambda}{R} - 1\right)^3 - \frac{\lambda}{R}\right\}^2 + 4\left\{-\frac{1}{9}\left(\frac{\lambda}{R} - 1\right)^2\right\}^3}\,\right]^{1/3} +$$

$$\left[-\frac{1}{2}\left\{\frac{2}{27}\left(\frac{\lambda}{R} - 1\right)^3 - \frac{\lambda}{R}\right\} - \frac{1}{2}\sqrt{\left\{\frac{2}{27}\left(\frac{\lambda}{R} - 1\right)^3 - \frac{\lambda}{R}\right\}^2 + 4\left\{-\frac{1}{9}\left(\frac{\lambda}{R} - 1\right)^2\right\}^3}\,\right]^{1/3} - \frac{1}{3}\left(\frac{\lambda}{R} - 1\right) \tag{5-20}$$

Consider the radical in (5-20) to reduce it

$$\sqrt{\left\{\frac{2}{27}\left(\frac{\lambda}{R} - 1\right)^3 - \frac{\lambda}{R}\right\}^2 + 4\left\{-\frac{1}{9}\left(\frac{\lambda}{R} - 1\right)^2\right\}^3} = \sqrt{\frac{4}{729}\left(\frac{\lambda}{R} - 1\right)^6 - \frac{4}{27}\left(\frac{\lambda}{R}\right)\left(\frac{\lambda}{R} - 1\right)^3 + \left(\frac{\lambda}{R}\right)^2 - \frac{4}{729}\left(\frac{\lambda}{R} - 1\right)^6}$$

$$= \sqrt{\left(\frac{\lambda}{R}\right)^2 - \frac{4}{27}\left(\frac{\lambda}{R}\right)\left(\frac{\lambda}{R} - 1\right)^3}$$

$$c = \left[\left\{-\frac{1}{27}\left(\frac{\lambda}{R} - 1\right)^3 + \frac{1}{2}\frac{\lambda}{R}\right\} + \frac{1}{2}\sqrt{\left(\frac{\lambda}{R}\right)^2 - \frac{4}{27}\left(\frac{\lambda}{R}\right)\left(\frac{\lambda}{R} - 1\right)^3}\,\right]^{1/3} + \left[\left\{-\frac{1}{27}\left(\frac{\lambda}{R} - 1\right)^3 + \frac{1}{2}\frac{\lambda}{R}\right\} - \frac{1}{2}\sqrt{\left(\frac{\lambda}{R}\right)^2 - \frac{4}{27}\left(\frac{\lambda}{R}\right)\left(\frac{\lambda}{R} - 1\right)^3}\,\right]^{1/3}$$

$$- \frac{1}{3}\left(\frac{\lambda}{R} - 1\right).$$

Let:

$$\alpha = -\frac{1}{27}\left(\frac{\lambda}{R} - 1\right)^3 + \frac{1}{2}\frac{\lambda}{R} \tag{5-21}$$

$$\beta = \frac{1}{2}\sqrt{\left(\frac{\lambda}{R}\right)^2 - \frac{4}{27}\left(\frac{\lambda}{R}\right)\left(\frac{\lambda}{R} - 1\right)^3} \tag{5-22}$$

$$\therefore c = (\alpha + \beta)^{1/3} + (\alpha - \beta)^{1/3} - \frac{1}{3}\left(\frac{\lambda}{R} - 1\right) \tag{5-23}$$

Having found c from Equation (5-23), then obtain a and b by substitution in (5-10) and (5-11). Note that it is not necessary to find all the roots of values of c, since any one will be suffice.

Now upon adding $(ax + b)^2$ to both members of Equation (5-8), an equation is obtained in which both members are perfect squares. It is, in fact,

$$(x^2 + \frac{p}{2}x + c)^2 = (ax + b)^2 .$$

Therefore, $$x^2 + \frac{p}{2} x + c = ax + b$$

or $$x^2 + \frac{p}{2} x + c = -ax - b.$$

Substitute for the value of p in the last two equations

$$x^2 - x + c = ax + b$$

or

$$x^2 - x + c = -ax - b$$

i.e.,

$$x^2 - (1 + a)x + (c - b) = 0 \qquad (5\text{-}24)$$

or

$$x^2 - (1 - a)x + (c + b) = 0. \qquad (5\text{-}25)$$

Therefore, the four roots of Equations (5-7) can be found by solving the quadratic Equations (5-24) and (5-25). Then, the roots are:

$$x_{1,2} = \frac{(1 + a) \pm \sqrt{(1 + a)^2 - 4(c - b)}}{2} \qquad (5\text{-}26)$$

$$x_{3,4} = \frac{(1 - a) \pm \sqrt{(1 - a)^2 - 4(c + b)}}{2}. \qquad (5\text{-}27)$$

To sum up, the preceding is the method of finding the minimum expected total system cost load factor, $x_{min.}$, also the summary of the procedure of finding out $x_{min.}$ is as follows.

Procedure

Suppose that it is given λ, C_W, and C_f and required the minimum expected total cost service rate, μ_o.

Step 1: Calculate

$$R = \frac{C_W}{C_f}$$

$$p = -2$$

$$q = 2(1-\frac{\lambda}{R})$$

$$V = \frac{4\lambda}{R}$$

$$\alpha = -\frac{1}{27}(\frac{\lambda}{R} - 1)^3 + \frac{1}{2}\frac{\lambda}{R}$$

$$\beta = \frac{1}{2}\sqrt{(\frac{\lambda}{R})^2 - \frac{4}{27}(\frac{\lambda}{R})(\frac{\lambda}{R} - 1)^3}.$$

Step 2: Calculate c from Equation (5-23) or

$$c = (\alpha + \beta)^{1/3} + (\alpha - \beta)^{1/3} + \frac{1}{6} q.$$

Step 3: Using Equations (5-10) and (5-11) calculate a and b, or

$$a = \sqrt{2c - q + 1}$$

$$b = \frac{-(2c + V)}{2a}$$

$\therefore$ two sets are considered in solution, either $(-a, + b)$ or $(a, - b)$, which both of them give the same solution.

Step 4: Calculate the minimum expected total cost load factor, which is the positive, less than one, value out of the four following values:

$$x_{1,2} = \frac{(1 + a) \pm \sqrt{(1 + a)^2 - 4(c - b)}}{2}$$

$$x_{3,4} = \frac{(1 - a) \pm \sqrt{(1 - a)^2 - 4(c + b)}}{2}.$$

Step 5: Find μ_o as

$$x_o = x_{min.} = \frac{\lambda}{\mu_o}$$

$$\therefore \mu_o = \frac{\lambda}{x_{min.}}. \qquad (5\text{-}28)$$

Step 6: Calculate TC_o, minimum expected total system cost by using Equation (5-6), and minimum load factor

$$TC_o = C_W\left\{\frac{x_o^2}{2(1 - x_o)} + x_o\right\} + \frac{C_f\lambda}{x_o}$$

or using Equation (5-5) and μ_o. Both of them lead to the same value of TC_o.

Sensitivity Analysis

Having followed the previous procedure to get $TC_{min.}$ at x_o, define K as previous chapters, such as:

$$K = \frac{TC - TC_{min.}}{TC_{min.}}$$

i.e.,

$$K = \frac{\left\{C_W\left[\frac{x^2}{2(1 - x)} + x\right] + \frac{\lambda C_f}{x}\right\} - TC_{min.}}{TC_{min.}}$$

$$K = \frac{C_W}{2TC_{min.}} \frac{x^2}{(1-x)} + \frac{C_W}{TC_{min.}} x + \frac{C_f \lambda}{TC_{min.}} \frac{1}{x} - 1. \tag{5-29}$$

As known from queuing theory, the load factor, x, does not exceed unity, otherwise the system will build to an infinite queue. That is, it will be explosive case. This is why on the values of x less than one are considered in Step 4 of the previous procedure of finding minimum expected total system cost.

By finding the set of values of K for the range of load factor (0, ..., 1), these values shown in Table XII. Figure 21 shows the shape of the K - x relation. After finding out the values of x_h and x_L from Table XII or Figure 21. μ_L and μ_h, the insignificant limits of expected total system cost can be found.

TABLE XII

TABULATION OF EQUATION (5-29) VALUES

x	$\frac{C_W}{2TC_{min.}} \frac{x^2}{(1-x)}$	$\frac{C_W}{TC_{min.}} x$	$\frac{C_f}{TC_{min.}} \frac{1}{x}$	-1	K
(1)	(2)	(3)	(4)	(5)	(6) = 2+3+4+5
0	----	----	∞	-1	∞
	.	.	.	.	.
	.	.	.	.	.
	.	.	.	.	.
1	∞	---	---	-1	∞

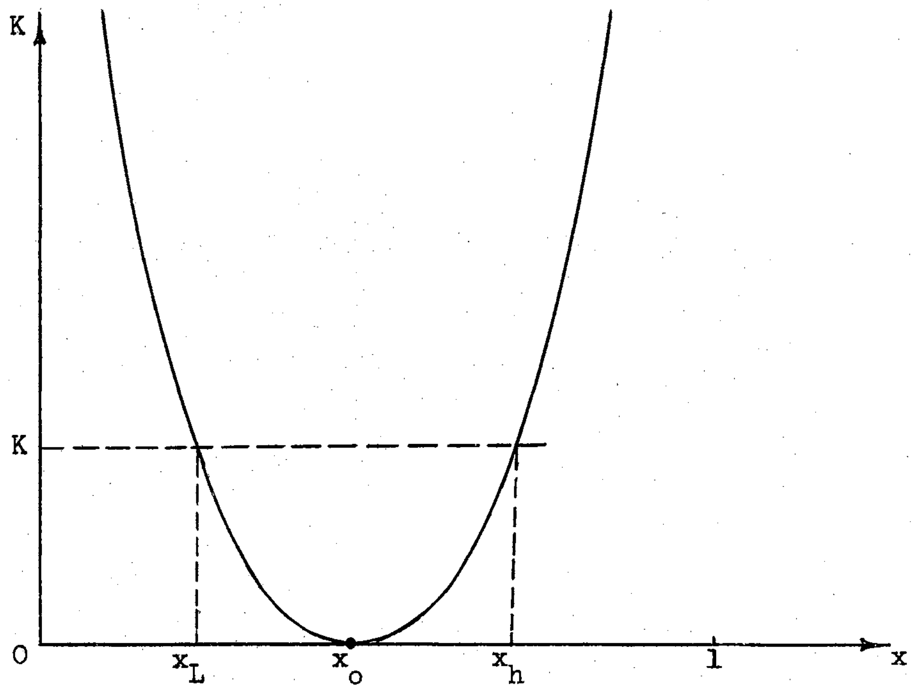


Figure 21. Graph of Equation (5-29)

$$\mu_L = \frac{\lambda}{x_h}$$

$$\mu_h = \frac{\lambda}{x_L} \qquad (5\text{-}30)$$

As an illustration, consider the case where

$C_W = 0.10$ $/unit/period

$C_f = 0.165$ $/unit/period

$\lambda = 0.125$ units/period.

Following the procedure given previously.

Step 1:

$$R = \frac{.100}{.165} = 0.606$$

$$p = -2$$

$$q = 2(1 - \frac{0.125}{0.606})$$

$$= 2(1 - 0.206)$$

$$= 2(0.794) = 1.588$$

$$v = 4(0.206) = 0.824$$

$$\alpha = -\frac{1}{27}(0.206 - 1)^3 + \frac{1}{2}(0.206)$$

$$= (0.2646)^3 + 0.103$$

$$= 0.0185 + 0.103 = 0.1215$$

$$\beta = \frac{1}{2}\sqrt{(0.206)^2 + 4(0.206)(0.0185)}$$

$$= 0.5\sqrt{0.0424 + (.824)(0.0185)}$$

$$= 0.5\sqrt{0.0424 + 0.01524}$$

$$= \sqrt{\frac{0.05764}{4}} = \sqrt{0.01441} = \underline{\underline{0.12}}.$$

Step 2:

$$c = (0.2415)^{1/3} + (0.0015)^{1/3} + \frac{1.588}{6}$$

$$= 0.623 + 0.106 + 0.265 = 1.094 \approx \underline{\underline{1.1}}$$

Step 3:

$$a = \sqrt{2.2 - 1.588 + 1}$$

$$= \sqrt{3.2 - 1.6} = \sqrt{1.6} = \underline{\underline{1.265}}$$

$$b = \frac{-(2.2 + 0.824)}{2.530} = -\frac{3.024}{2.530} = -\underline{\underline{1.195}}.$$

Step 4:

$$x_{1,2} = \frac{2.265 \pm \sqrt{(2.265)^2 - 4(2.295)}}{2}$$

$x_{1,2}$ has imaginary radical which violates the assumption

$$x_{3,4} = \frac{-.265 \pm \sqrt{(.265)^2 - 4(-0.095)}}{2}$$

$$= -0.1325 \pm 0.5\sqrt{0.0722 + 0.3800}$$

$$= -0.1325 \pm \sqrt{\frac{.4522}{4}}$$

$$= -0.1325 \pm \sqrt{0.11305}$$

$$= -0.1325 \pm 0.3605$$

$x_{min.} = 0.228$ i.e., $x_o = 0.23$

$$TC_{min.} = C_W\left\{\frac{x_o^2}{2(1 - x_o)} + x_o\right\} + C_f \frac{\lambda}{x_o}$$

$$= 0.10\left\{\frac{(0.23)^2}{2(0.77)} + 0.23\right\} + 0.165\left(\frac{0.125}{0.23}\right)$$

$$= 0.10\left(\frac{.0529}{1.54}\right) + 0.023 + 0.165(0.54)$$

$$= 0.0030 + 0.023 + 0.0800$$

$$= \underline{\underline{0.106}}$$

$$\therefore K = \frac{0.1}{2(.106)} \frac{x^2}{(1 - x)} + \frac{0.1}{0.106} x$$

$$+ \frac{0.165(.125)}{0.106} \frac{1}{x} - 1$$

$$K = \frac{0.1}{.212} \frac{x^2}{(1 - x)} + 0.9433 x + \frac{1.5566}{8} \frac{1}{x} - 1$$

$$K = 0.472 \frac{x^2}{(1 - x)} + 0.9433 x + 0.194 \frac{1}{x} - 1 \tag{5-31}$$

Table XIII shows values of K at different values of x within the range from zero to one. And Figure 22 shows the graphical presentation of Equation (5-31). Using this graph the lower limit of the load factor, x_L, and the higher limit of the load factor, x_h, can be obtained at certain value of K. For example, at K = 0.4, $x_L = 0.14$, and $x_h = 0.62$. From x_L and x_h, μ_L and μ_h can be calculated using Equation (5-30) or,

$$\mu_h = \frac{0.125}{0.14} = 0.89 \text{ and } \mu_L = \frac{0.125}{0.62} = 0.2.$$

<u>Indifferent Range</u>

The indifferent range is the higher limit minus the lower limit or,

$$r = \mu_h - \mu_L$$

$$= \frac{\lambda}{x_L} - \frac{\lambda}{x_h}$$

TABLE XIII

TABULATION OF EQUATION (5-31)

x	$0.472\frac{x^2}{(1-x)}$	0.943x	$0.194\frac{1}{x}$	Total	K
(1)	(2)	(3)	(4)	(5) = (2)+(3)+(4)	(6) = (5) - 1
0	0	0	∞	∞	∞
0.1	0.00524	0.0943	1.940	2.03924	1.03924
0.2	0.0236	0.1886	0.970	1.1822	0.1822
0.3	0.0472	0.2829	0.681	1.0111	0.011
0.4	0.1460	0.3772	0.485	1.0082	0.0082
0.5	0.236	0.4715	0.388	1.0955	0.0955
0.6	0.425	0.5658	0.323	1.3138	0.3138
0.7	0.7694	0.6601	0.277	1.7065	0.7065
0.8	1.5104	0.7544	0.243	2.5078	1.5078
0.9	3.8232	0.8487	0.215	4.8869	3.8869
1.0	∞	0.943	0.194	∞	∞

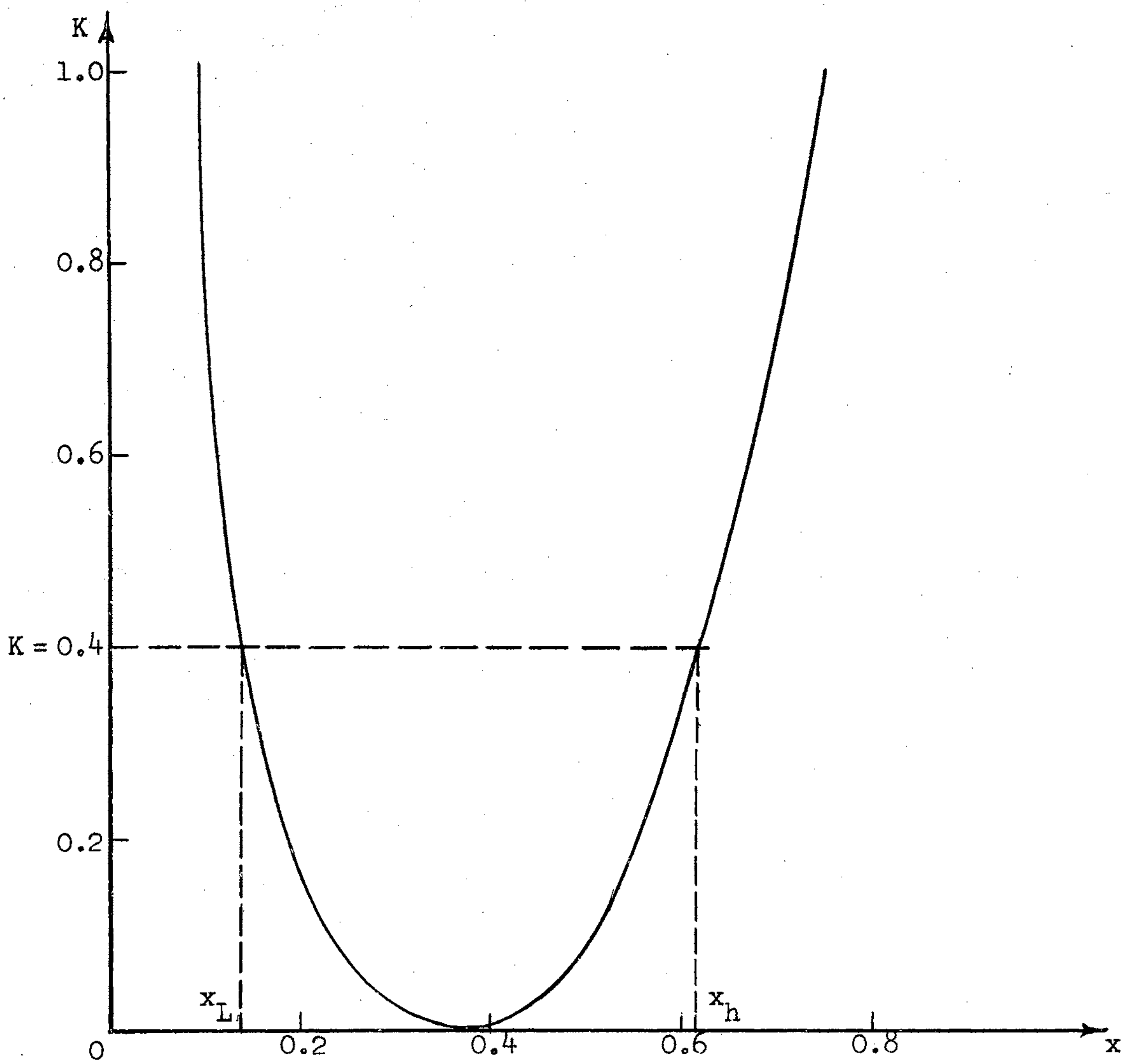


Figure 22. Graphical Presentation of Equation (5-31)

$$r = \frac{\lambda(x_h - x_L)}{x_L \, x_h}. \tag{5-32}$$

For the above example, the indifferent range is

$$r = 0.89 - 0.2 = \underline{\underline{0.79}}$$

CHAPTER VI

POISSON ARRIVALS WITH ANY SERVICE TIME DISTRIBUTION

For further generality, it is desirable to have expressions for pertinent system characteristics regardless of the form of the service time distribution. If σ^2 is the variance of the service time distribution, the mean number of units in the system is given by:

$$n_m = \frac{\left(\frac{\lambda}{\mu}\right)^2 + \lambda^2\sigma^2}{2\left[1 - \left(\frac{\lambda}{\mu}\right)\right]} + \left(\frac{\lambda}{\mu}\right).^{(*)} \qquad (6\text{-}1)$$

Consider, like the previous chapter, that $X = \frac{\lambda}{\mu}$, where X is called the load factor. The load factor is defined as the ratio of arrival rate to potential service rate. Taking this load factor into consideration Equation (6-1) would be:

$$n_m = \frac{X^2 + \lambda^2\sigma^2}{2[1 - X]} + X \qquad (6\text{-}2)$$

and the mean waiting time as is given

(*) For proof, see Appendix A.

$$W_m = \frac{(\frac{\lambda}{\mu^2}) + \lambda\sigma^2}{2[1 - (\frac{\lambda}{\mu})]} + \frac{1}{\mu} \quad (*)$$

$$= \frac{(\frac{\lambda}{\mu^2}) + \lambda\sigma^2}{2[1 - (\frac{\lambda}{\mu})]} + \frac{\lambda}{\lambda\mu}$$

$$= \frac{(\frac{\lambda}{\mu^2}) + \lambda\sigma^2}{2[1 - (\frac{\lambda}{\mu})]} + \frac{(\frac{\lambda}{\mu})}{\lambda}$$

$$W_m = \frac{\frac{X^2}{\lambda} + \lambda\sigma^2}{2[1 - X]} + \frac{X}{\lambda} \qquad (6\text{-}3)$$

The Expected Total System Cost

The expected total system cost per period is the sum of the expected waiting cost per period and the expected facility cost per period; that is,

$$TC_m = WC_m + FC_m$$

The expected waiting cost per period in the product of the cost of waiting per unit per period and the mean number of units in the system during the period. The expected facility cost per period may be taken as the product of the cost of providing service facility of unit rate capacity and the service rate in units per period. Therefore, the expected total system cost per period is

---

(*) For proof, see Appendix A.

$$TC_m = C_W \left[\frac{X^2 + \lambda\sigma^2}{2(1 - X)} + X\right] + C_f\mu$$

and since

$$X = \frac{\lambda}{\mu}$$

$$\therefore \mu = \frac{\lambda}{X}$$

$$\therefore TC_m = C_W \left[\frac{X^2 + \lambda\sigma^2}{2(1 - X)} + X\right] + \frac{C_f\lambda}{X} \qquad (6\text{-}4)$$

As shown in Equation (6-4) the expected total system cost is a function of load factor, expected arrival rate, and variance of the service distribution. To calculate minimum expected total system cost at constant λ, differentiate with respect to X and equate to zero.

$$\therefore \frac{dTC_m}{dX} = C_W\left[\frac{2(1 - X) \cdot 2X - (X^2 + \lambda\sigma^2)(-2)}{4(1 - X)^2} + 1\right] - \frac{C_f\lambda}{X^2} = 0$$

$$= C_W\left[\frac{2X(1 - X) + (X^2 + \lambda\sigma^2) + 2(1 - X)^2}{2(1 - X)^2}\right] - \frac{C_f\lambda}{X^2} = 0$$

$$= C_W X^2 [2X(1 - X) + (X^2 + \lambda\sigma^2) + 2(1 - X)^2]$$

$$- 2(1 - X)^2 C_f\lambda = 0$$

Let $R = \frac{C_W}{C_f}$, and divide the previous equation by C_f,

$$2RX^3 - 2RX^4 + RX^4 + R\lambda\sigma^2X^2 + 2RX^2 - 4RX^3 + 2RX^4$$

$$- 2\lambda + 4\lambda X - 2\lambda X^2 = 0$$

$$RX^4 - 2RX^3 + (R\lambda\sigma^2 - 2\lambda + 2R)X^2 + 4\lambda X - 2\lambda = 0$$

divide by R:

$$X^4 - 2X^3 + (\lambda\sigma^2 - \frac{2\lambda}{R} + 2)X^2 + 4\frac{\lambda}{R}X - \frac{2\lambda}{R} = 0 \qquad (6\text{-}5)$$

Equation (6-5) can be simplified by defining the following:

$$P = -2$$

$$q = [\lambda\sigma^2 + 2(1 - \frac{\lambda}{R})]$$

$$V = \frac{4\lambda}{R}$$

$$W = -\frac{2\lambda}{R}$$

Put P, q, V, and W in Equation (6-5),

$$X^4 + PX^3 + qX^2 + VX + W = 0 \qquad (6\text{-}6)$$

X_o the minimum expected total system cost load factor, can be calculated by solving Equation (6-6).

Solution of the Equation (6-6)

We first determine a, b, and C such that

$$X^4 + PX^3 + qX^2 + VX + W + (aX + b)^2 = (X^2 + \frac{P}{2}X + C)^2.$$

The determination of a, b, and C is accomplished by equating the coefficient of the powers of X in the first and second members of the last equation

$$a^2 + q = 2C + \frac{P^2}{4} \qquad (6\text{-}7)$$

$$2ab + V = CP \tag{6-8}$$

$$b^2 + W = C^2 \tag{6-9}$$

Hence from the Equations (6-7), (6-8), and (6-9), it would be

$$(CP - V)^2 = 4a^2 b^2 = 4(2C + \frac{P^2}{4} - q)(C^2 - W)$$

or

$$C^3 - \frac{q}{2} C^2 + \frac{1}{4}(PV - 4W)C + \frac{1}{8}(4qW - P^2W - V^2) = 0. \tag{6-10}$$

Let

$$F = -\frac{q}{2} = \frac{-\lambda\sigma^2}{2} + (\frac{\lambda}{R} - 1)$$

$$F = (\frac{\lambda}{R} - 1) - \frac{\lambda\sigma^2}{2}$$

$$Y = \frac{1}{4}(PV - 4W)$$

$$= \frac{1}{4}\left[-2\left(\frac{4\lambda}{R}\right) - 4\left(\frac{-2\lambda}{R}\right)\right] = 0$$

$$Y = 0$$

and

$$E = \frac{1}{8}(4qW - P^2W - V^2)$$

$$= \frac{1}{8}\left[4\left(\frac{-2\lambda}{R}\right)\left(\lambda\sigma^2 + 2 - \frac{2\lambda}{R}\right) - 4\left(\frac{-2\lambda}{R}\right) - \left(\frac{4\lambda}{R}\right)^2\right]$$

$$= \frac{1}{8}\left[\frac{-8\lambda^2\sigma^2}{R} - \frac{16\lambda}{R} + \frac{16\lambda^2}{R^2} + \frac{8\lambda}{R} - \frac{16\lambda^2}{R^2}\right]$$

$$= \frac{-\lambda}{R}(\lambda\sigma + 1)$$

$$E = -\frac{\lambda}{R}(\lambda\sigma^2 + 1).$$

Substitute in Equation (6-10)

$$C^3 + FC^2 + YC + E = 0 \tag{6-11}$$

C can be found by solving Equation (6-11) and then obtain a and b by substitution in Equation (6-7) and Equation (6-8).

A method of solving Equation (6-11) will now be explained. Equation (6-11) will be first transformed so as to remove the second degree term. Let

$$C = Z - \frac{F}{3} \tag{6-12}$$

Substitute from (6-12), Equation (6-11) would be

$$Z^3 + 3HZ + G = 0 \;^{(*)} \tag{6-13}$$

where:

$$H \overset{(*)}{=} \frac{3Y - F^2}{9} = -\frac{1}{9}\left[\left(\frac{\lambda}{R} - 1\right) - \frac{\lambda\sigma^2}{2}\right]^2 \tag{6-14}$$

and

$$G \overset{(*)}{=} \frac{2F^3 - 9FY + 27E}{27}$$

$$G = \frac{2}{27}\left[\left(\frac{\lambda}{R} - 1\right) - \frac{\lambda\sigma^2}{2}\right]^3 - \frac{\lambda}{R}\left(\lambda\sigma^2 + 1\right) \tag{6-15}$$

The roots of the given Equation (6-11) can be found from Equation (6-12) when the roots of Equation (6-13) are known. Equation (6-13) can be solved for Z by following Cardan's formulas as follows:

$$Z = \left[\frac{-G + \sqrt{G^2 + 4H^3}}{2}\right]^{1/3} + \left[\frac{-G - \sqrt{G^2 + 4H^3}}{2}\right]^{1/3}$$

(*) For proof see Introduction to the Theory of Equations, by Conkwright, Ginn and Company, 1941, pp. 76-77.

To get the solution for C substitute Z in Equation (6-12),

$$C = \left[\frac{-G + \sqrt{G^2 + 4H^3}}{2}\right]^{1/3} + \left[\frac{-G - \sqrt{G^2 + 4H^3}}{2}\right]^{1/3} - \frac{F}{3} \quad (6\text{-}16)$$

Substitute for the values G, H, and F and get value of C. Having found C from the preceding equation, then obtain a and b from Equations (6-7) and (6-8) by substitution for the value C. Note that it is not necessary to find all the roots of values of C, since any one will suffice.

Now upon adding $(aX + b)^2$ to both members of Equation (6-5), an equation is obtained in which both members are perfect squares. It is, in fact,

$$(X^2 + \frac{P}{2} X + C)^2 = (aX + b)^2 \ .$$

Therefore,

$$X^2 + \frac{P}{2} X + C = aX + b$$

or

$$X^2 + \frac{P}{2} X + C = -aX - b \qquad (6\text{-}17)$$

Substitute for P values in Equation (6-17)

$$X^2 - X + C = aX + b \qquad (6\text{-}18)$$

or

$$X^2 - X + C = -aX - b \qquad (6\text{-}19)$$

From Equation (6-18), the first two roots of Equation (6-5) can be obtained. They would be:

$$X^2 - X + C - aX - b = 0$$

$$X^2 - (a + 1)X + (C - b) = 0 \qquad (6\text{-}20)$$

$$X_{1,2} = \frac{(a+1) \pm \sqrt{(a+1)^2 - 4(C-b)}}{2} \qquad (6\text{-}21)$$

The other two roots of Equation (6-5) can be found from Equation (6-19)

$$X^2 - X + C + aX + b = 0$$

$$X^2 - (1 - a)X + (b + C) = 0 \qquad (6\text{-}22)$$

$$X_{3,4} = \frac{(1-a) \pm \sqrt{(1-a)^2 - 4(b+C)}}{2} \qquad (6\text{-}23)$$

Only positive values, equal or less than one, of these four roots are considered in the solution. The imaginary values and the negative ones do not make sense for the queuing situation. The positive values of these roots, which are greater than one, are also neglected because they are violating the assumptions. That is, a load factor, which is greater than unity, means the arrival rate is greater than the service rate. This is not the case considered here.

As a summary, the procedure to find the minimum expected total system cost, $TC_{min.}$, is as follows:

The given data for the problem C_W, C_f, λ, and variance of the service distribution, σ^2.

Procedure to Find the Minimum Expected Total System Cost

Step 1: Calculate the following:

$$P = -2$$

$$q = \left[\lambda\sigma^2 + 2\left(1 - \frac{\lambda}{R}\right)\right]$$

$$V = \frac{4\lambda}{R}$$

$$W = \frac{-2\lambda}{R}$$

$$F = -\frac{\lambda}{R}(\lambda\sigma^2 + 1)$$

$$H = -\frac{1}{9}\left[\left(\frac{\lambda}{R} - 1\right) - \frac{\lambda\sigma^2}{2}\right]^2$$

$$G = \frac{2}{27}\left[\left(\frac{\lambda}{R} - 1\right) - \frac{\lambda\sigma^2}{2}\right]^3 - \frac{\lambda}{R}(\lambda\sigma^2 + 1)$$

Step 2: Calculate C using the following equation

$$C = \left[\frac{-G + \sqrt{G^2 + 4H^3}}{2}\right]^{1/3} + \left[\frac{-G - \sqrt{G^2 + 4H^3}}{2}\right]^{1/3} - \frac{F}{3}$$

Step 3: Calculate a and b from Equations (6-7) and (6-8) or

$$a = \sqrt{2C + \frac{P^2}{4} - q}$$

$$b = \frac{CP - V}{2a}$$

Step 4: Calculate the load factor values using Equations (6-21) and (6-23)

$$X_{1,2} = \frac{(a + 1) \pm \sqrt{(a + 1)^2 - 4(C - b)}}{2}$$

$$X_{3,4} = \frac{(1 - a) \pm \sqrt{(1 - a)^2 - 4(C + b)}}{2}$$

Step 5: Between the values of X, in Step 4, the one which is less than or equal to unity will be the minimum total system cost load factor, $X_{min.}$. If there are more than one having positive values less or equal to unity, the one which gives the minimum value in Equation (6-4) should be chosen as $X_{min.}$.

Step 6: Calculate $TC_{min.}$ by substituting $X_{min.}$ in Equation (6-4).

If σ^2 is set to zero in the above steps of the procedure, the steps reduce to the steps of the procedure in Chapter V, constant service rate case.

Sensitivity Analysis

For the calculation of the insignificant limits, consider the same definition of K as previously stated where

$$K = \frac{TC - TC_{min}}{TC_{min}}$$

Substitute for TC from Equation (6-4), the above equation would be

$$K = \frac{\left\{C_W\left[\frac{X^2 + \lambda\sigma^2}{2(1 - X)} + X\right] + \frac{\lambda C_f}{X}\right\} - TC_{min}}{TC_{min}}$$

or

$$K = \frac{C_W}{2TC_{min}} \cdot \frac{X^2}{(1-X)} + \frac{C_W\lambda\sigma^2}{2TC_{min}} \cdot \frac{1}{(1-X)} + \frac{C_W X}{TC_{min}} + \frac{C_f\lambda}{TC_{min}} \cdot \frac{1}{X} - 1 \qquad (6\text{-}24)$$

Equation (6-24) shows that for every value of K there are two values of X, X_h and X_L. Once the decision maker sets a value for K, the higher and the lower values of the load factor, X, can be calculated from Equation (6-24). Then the higher and the lower limits of the service rate, μ_h and μ_L, can be calculated from the following relations,

$$\mu_L = \frac{\lambda}{X_h}$$

$$\mu_h = \frac{\lambda}{X_L} \tag{6-25}$$

As an illustration, consider the case where the number of arrivals per hour has a Poisson distribution with a mean of 0.2 units. The cost of waiting per unit per hour is $2.10 and the cost of serving one unit is $4.05. The purpose is to find the minimum expected total system cost service rate with a service time variance of 3(hours)^2. i.e., <u>given</u>:

C_W = $2.00 per unit per period

λ = 0.2 units per period

σ^2 = 3(hours)^2

C_f = $4.00 per unit per period

<u>Required</u>: The minimum total system cost, μ_o. By following the procedure given in this chapter, the calculation of μ_o would be,

<u>Step 1</u>: $P = -2, \quad R = \frac{2.0}{4.0} = 0.50$

$$q = [(0.2)(3) + 2(1 - 0.4)]$$
$$= 0.6 + 1.2 = 1.8$$

$$V = \frac{4(0.2)}{0.5} = 1.6$$

$$W = \frac{-2(0.2)}{0.5} = -0.8$$

$$F = -\frac{0.2}{0.5}[0.2(3) + 1]$$

$$= -0.24 - 0.4 = -0.64$$

$$H = -\frac{1}{9}\left[\left(\frac{0.2}{0.5} - 1\right) - \frac{0.2(3)}{2}\right]^2$$

$$= -\frac{1}{9}[-0.6 - 0.3]^2 = -\frac{1}{9} \text{ X } 0.81 = -0.09$$

$$G = \frac{2}{27}\left[\left(\frac{0.2}{0.5} - 1\right) - \frac{0.2(3)}{2}\right]^3 - \frac{0.2}{0.5}(0.2 \text{ X } 3 + 1)$$

$$= \frac{2}{27}(-0.9)^3 - 0.4(1.6)$$

$$= -0.054 - 0.64 = -0.694$$

Step 2:

$$C = \left[\frac{0.694 + \sqrt{(.694)^2 + 4(0.09)^3}}{2}\right]^{1/3}$$

$$+ \left[\frac{0.694 - \sqrt{(.694)^2 + 4(-0.09)^3}}{2}\right]^{1/3} + \frac{0.64}{3}$$

$$= \left[\frac{0.694 + \sqrt{0.482 - 0.003}}{2}\right]^{1/3}$$

$$+ \left[\frac{0.694 - \sqrt{0.482 - 0.003}}{2}\right]^{1/3} + 0.213$$

$$= \left[\frac{0.694 + \sqrt{0.479}}{2}\right]^{1/3} + \left[\frac{0.694 - \sqrt{0.479}}{2}\right]^{1/3} + 0.213$$

$$= \left[\frac{0.694 + 0.69}{2}\right]^{1/3} + \left[\frac{0.694 - 0.69}{2}\right]^{1/3} + 0.23$$

$$= (0.692)^{1/3} + (0.002)^{1/3} + 0.23$$

$$= 0.885 + 0.126 + 0.230$$

$$= 1.241$$

Step 3: $a = \sqrt{2(1.241) + 1 - 1.8}$

$$= \sqrt{3.482 - 1.800}$$

$$= \sqrt{1.682} = 1.30$$

$$b = \frac{1.241(-2) - 1.6}{2(1.30)} = \frac{-2.482 - 1.6}{2.6} = -\frac{4.082}{2.6}$$

$$= -1.570$$

Step 4: $X_{1,2} = \frac{2.30 \pm \sqrt{(2.30)^2 - 4(2.811)}}{2}$

The radical for $X_{1,2}$ is imaginary, then $X_{3,4}$ is considered,

$$X_{3,4} = \frac{-0.30 \pm \sqrt{(0.30)^2 + 4(0.33)}}{2} = \frac{-0.30 \pm \sqrt{1.410}}{2}$$

$$= \frac{-0.30 \pm 1.187}{2}$$

$$X_o = \frac{0.887}{2} = 0.443$$

$$\because \quad X_0 = 0.443 = \frac{\lambda}{\mu_o} = \frac{0.2}{\mu_o}$$

$$\therefore \quad \mu_o = \frac{0.2}{0.443} = 0.45 \text{ units/period.}$$

Step 6: $TC_{min} = C_W \left[\frac{X_o^2 + \sigma^2}{2(1 - X_o)} + X_o\right] + \frac{C_f \lambda}{X_o}$

$$TC_{min} = 2.0 \left[\frac{(0.443)^2 + 0.2(3)}{2(1-0.443)} + 0.443\right] + \frac{4(0.2)}{0.443}$$

$$= \frac{0.196 + 0.6}{0.557} + 0.886 + \frac{0.8}{0.443}$$

$$= 1.429 + 0.886 + 1.806$$

$$= \$4.121 \approx 4.12$$

Substitute in Equation (6-24)

$$K = \frac{2}{2(4.12)} \frac{X^2}{(1-X)} + \frac{2(0.2)3}{2(4.12)} \frac{1}{(1-X)} + \frac{2.0(X)}{4.12} + \frac{4.0(0.2)}{4.12} \frac{1}{X} - 1$$

$$K = 0.243 \frac{X^2}{(1-X)} + 0.15 \frac{1}{(1-X)} + 0.485\ X + 0.194 \frac{1}{X} - 1 \qquad (6\text{-}27)$$

Table XIV shows tabulation values for Equation (6-27). Figure 23 shows the graphical presentation of Equation (6-27). Both the table and the figure show as X increases from zero to one, K value decreases to zero and increases again towards its infinity at X = 1. K reaches its zero value at $X = X_{min.}$, the minimum expected total system cost load factor. This is shown in the above calculation at K = 0, $X_{min} = 0.443$. Moreover, consider the decision maker sets the value of K = 0.6 at which the difference in the expected total system costs to the minimum is insignificant. From Figure 23, at K = 0.6 the values of the load factor are $X_L = 0.150$ and $X_h = 0.725$. The insignificant service rate limits, μ_L and μ_h, can be calculated as follows:

TABLE XIV

NUMERICAL VALUES OF EQUATION (6-27)

X	$\frac{0.243\ X^2}{(1 - X)}$	$\frac{0.15}{(1 - X)}$	0.485 X	$\frac{0.194}{X}$	-1	K
0	0	0.1500	0.485	∞	-1	∞
0.1	0.0027	0.1667	0.0485	1.9400	-1	1.1579
0.2	0.0122	0.1875	0.0970	0.9700	-1	0.2667
0.3	0.0312	0.2143	0.1455	0.6467	-1	0.0377
0.4	0.0648	0.2500	0.1940	0.4850	-1	0.0062
0.5	0.1215	0.30000	0.2425	0.3880	-1	0.0528
0.6	0.2187	0.3750	0.2910	0.3233	-1	0.2080
0.7	0.3967	0.5000	0.3395	0.2771	-1	0.5133
0.8	0.7776	0.7500	0.3880	0.2425	-1	1.1581
0.9	1.9683	1.5000	0.4365	0.2156	-1	3.1204
1.00	∞	∞	0.485	0.194	-1	∞

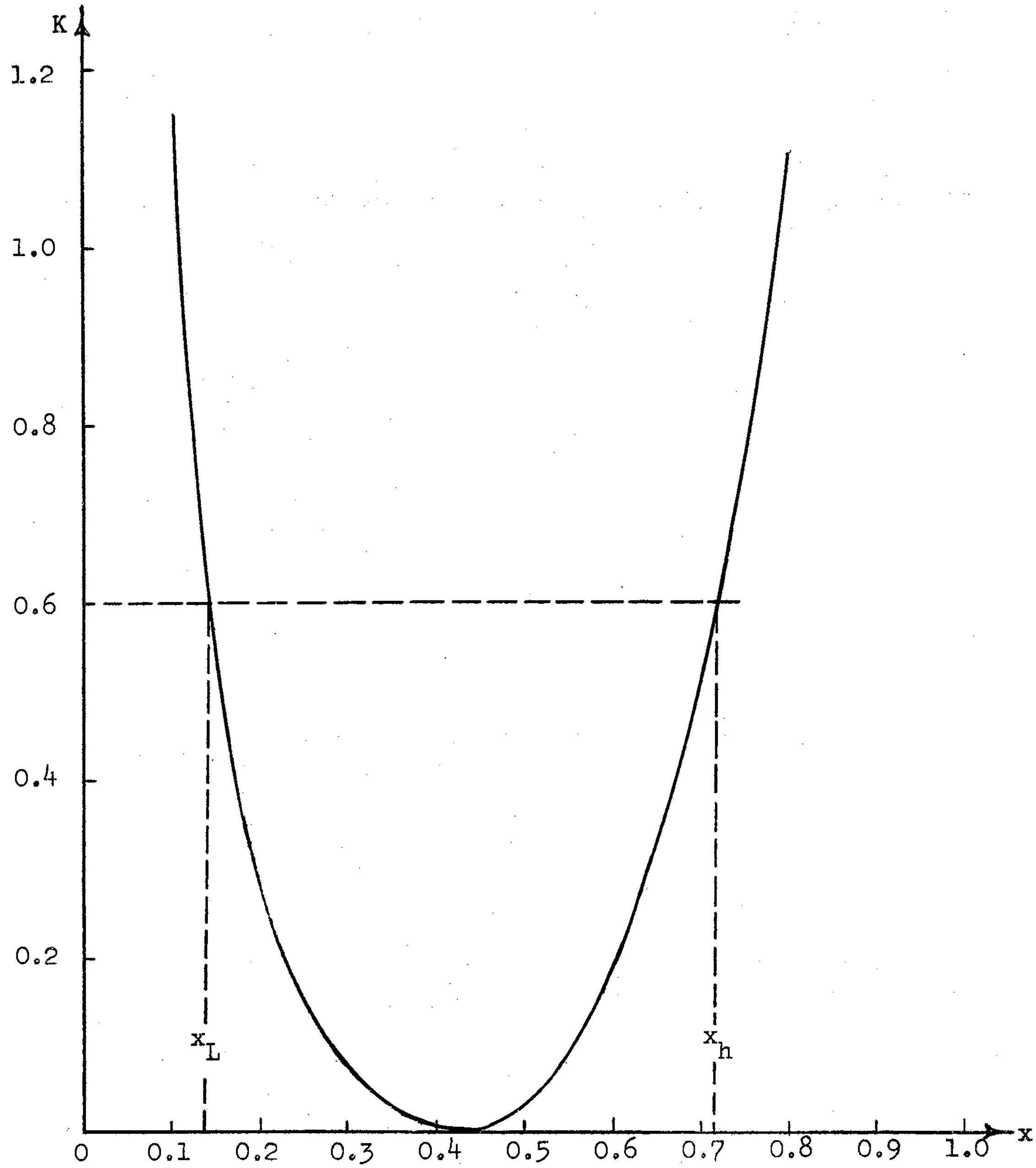


Figure 23. Graphical Presentation of Equation (6-27)

$$\mu_L = \frac{\lambda}{X_h} = \frac{0.2}{0.725} = 0.28 \quad \text{units/period}$$

$$\mu_h = \frac{\lambda}{X_L} = \frac{0.2}{0.150} = 1.33 \quad \text{units/period}$$

Indifferent Range

The indifferent range, r, is the difference between the higher limit service rate, μ_h, and the lower limit service rate, μ_L.

$$\therefore \quad r = \mu_h - \mu_L = 1.33 - 0.28 = 1.05$$

The Effect of Variance

Consider the above illustration again and calculate the minimum expected total system cost service rate, μ_0, the insignificant limits, μ_L and μ_h, and the indifferent range, r, at service time variance equal to one instead of three keeping the same values for the other variables. The calculations are as follows:

Step 1: $P = -2 \qquad R = \frac{2}{4} = \frac{1}{2}$

$$q = [0.2(1) + 2(1 - \frac{0.2}{0.5})] = 1.4$$

$$V = \frac{4(0.2)}{0.5} = 1.6$$

$$W = \frac{-2(0.2)}{0.5} = -0.8$$

$$F = -\frac{0.2}{0.5}[0.2(1) + 1] = -0.48$$

$$H = -\frac{1}{9}\left[\left(\frac{0.2}{0.5} - 1\right) - \frac{0.2(1)}{2}\right]^2 = -0.054$$

$$G = \frac{2}{27}[-0.6 - 0.1]^3 - \frac{0.2}{0.5}(1.2) = -0.6$$

Step 2:

$$C = \left[\frac{0.6 + \sqrt{0.36 - 0.0005}}{2}\right]^{1/3} + \left[\frac{0.6 - \sqrt{0.36 - 0.0005}}{2}\right]^{1/3} + 0.16$$

$$C = \sqrt[3]{0.6} + 0.16 = 1.00$$

Step 3: $a = \sqrt{2 + 1 - 1.4} = 1.265$

$$b = \frac{(1)(-2) - 1.6}{2(0.98)} = -1.84$$

Step 4:

$$X_{1,2} = \frac{(2.265) \pm \sqrt{5.13 - 4(2.84)}}{2}$$

The radical is imaginary. Consequently X_1 and X_2 are imaginary values

$$X_{3,4} = \frac{-0.265 \pm \sqrt{0.07 + 3.36}}{2}$$

from which

$$X_3 = \frac{-0.265 - 1.85}{2} = \text{negative value}$$

$$X_4 = \frac{-0.265 + 1.85}{2} = 0.292$$

$\therefore X_o = 0.29$

and

$$\mu_o = \frac{\lambda}{X_o} = \frac{0.2}{0.29} = 0.69$$

Step 6:

$$TC_{min} = 2\left[\frac{(0.29)^2 + 0.2(1)}{2(1 - 0.29)} + 0.29\right] + \frac{4(0.2)}{0.29} = \$3.73$$

By substituting in Equation (6-24) it would be

$$K = \frac{2}{2(3.73)} \frac{X^2}{(1-X)} + \frac{2(0.2)(1)}{2(3.73)} \frac{1}{(1-X)} + \frac{2X}{3.73} + \frac{4(.02)}{3.73} \frac{1}{X} - 1$$

$$K = \frac{X^2}{3.73(1-X)} + \frac{0.054}{(1-X)} + 0.536X + \frac{0.215}{X} - 1 \qquad (6\text{-}28)$$

The numerical and graphical presentation of Equation (6-28) are shown in Table XV and Figure 24, respectively. Consider the case where the decision maker sets up the value of K = 0.6 at which the difference in the expected total system costs to the minimum is insignificant. The insignificant service rate limits can be calculated using Figure 24. At K = 0.6, X_L and X_h are 0.225 and 0.299, respectively. μ_L and μ_h can be calculated from Equation (6-25)

$$\mu_L = \frac{0.2}{0.299} = 0.67$$

$$\mu_h = \frac{0.2}{0.225} = 0.89.$$

Indifferent range = $\mu_h - \mu_L$

$$r = 0.89 - 0.67 = 0.22$$

The effect of variance on the insignificant limits can be shown by comparing the results of the above illustration. By decreasing the variance from three to one, the minimum expected total system cost service rate, μ_o, decreases from 0.69 units/period to 0.29 units/period. This results from the decrease of the minimum load factor, X_o. The lower insignificant limit increases from 0.28 to 0.67 and the upper insignificant limit decreases from 1.33 to 0.89

TABLE XV

NUMERICAL VALUES OF EQUATION (6-28)

X	$\frac{X^2}{3.73(1-X)}$	$\frac{0.54}{(1-X)}$	0.536 X	$\frac{0.215}{X}$	-1	K
0	0	0.54	0	∞	-1	∞
0.1	0.00298	0.600	0.0536	2.15	-1	1.8036
0.2	0.0134	0.675	0.1072	1.075	-1	0.8706
0.3	0.0345	0.7714	0.1608	0.717	-1	0.6837
0.4	0.1074	0.900	0.2144	0.5375	-1	0.7593
0.5	0.1344	1.08	0.268	0.430	-1	0.9124
0.6	0.2416	1.35	0.3216	0.3583	-1	1.2715
0.7	0.4379	1.80	0.3752	0.3071	-1	1.9202
0.8	0.8579	2.70	0.4288	0.2688	-1	3.2555
0.9	2.1716	5.4	0.4824	0.2389	-1	7.2929
1.00	∞	∞	0.536	0.215	-1	∞

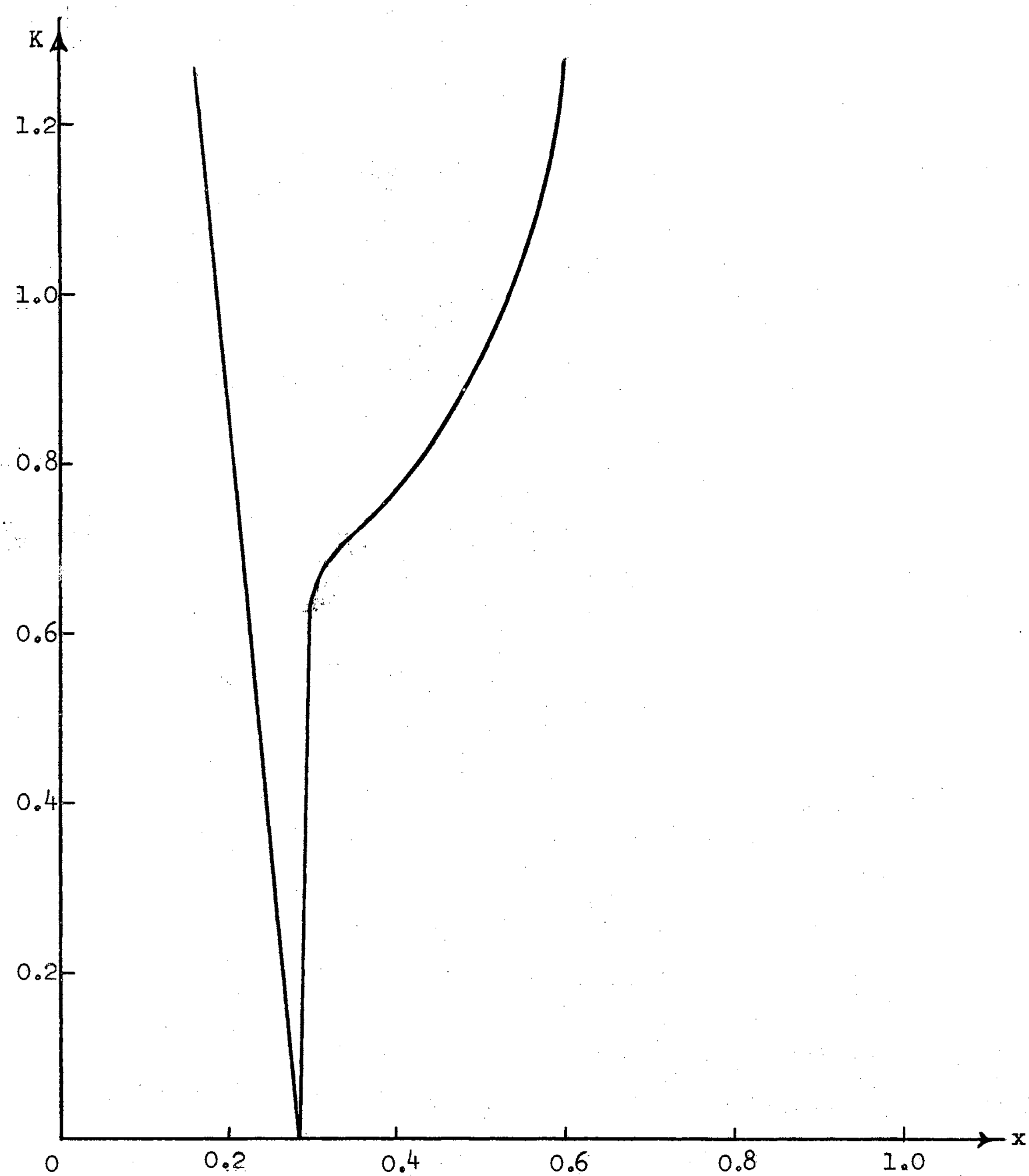


Figure 24. Graphical Presentation of Equation (6-28)

units/period. The range decreases from 1.05 to 0.22. Figure 25 shows the effect of the variance graphically. In general, as the variance increases the K - X curve moves to the right and becomes wider. In other words, as the variance increases, keeping the other variables constant, the minimum expected total system cost service rate, μ_0, decreases and the indifferent range increases.

TABLE XVI

THE EFFECT OF VARIANCE

X	$K_{\sigma^2=1}$	$K_{\sigma^2=3}$
0	∞	∞
0.1	1.1579	1.8036
0.2	0.2667	0.8706
0.3	0.0377	0.6837
0.4	0.0062	0.7593
0.5	0.0528	0.9124
0.6	0.2080	1.2715
0.7	0.5133	1.9202
0.8	1.1581	3.2555
0.9	3.1204	
1.0	∞	∞

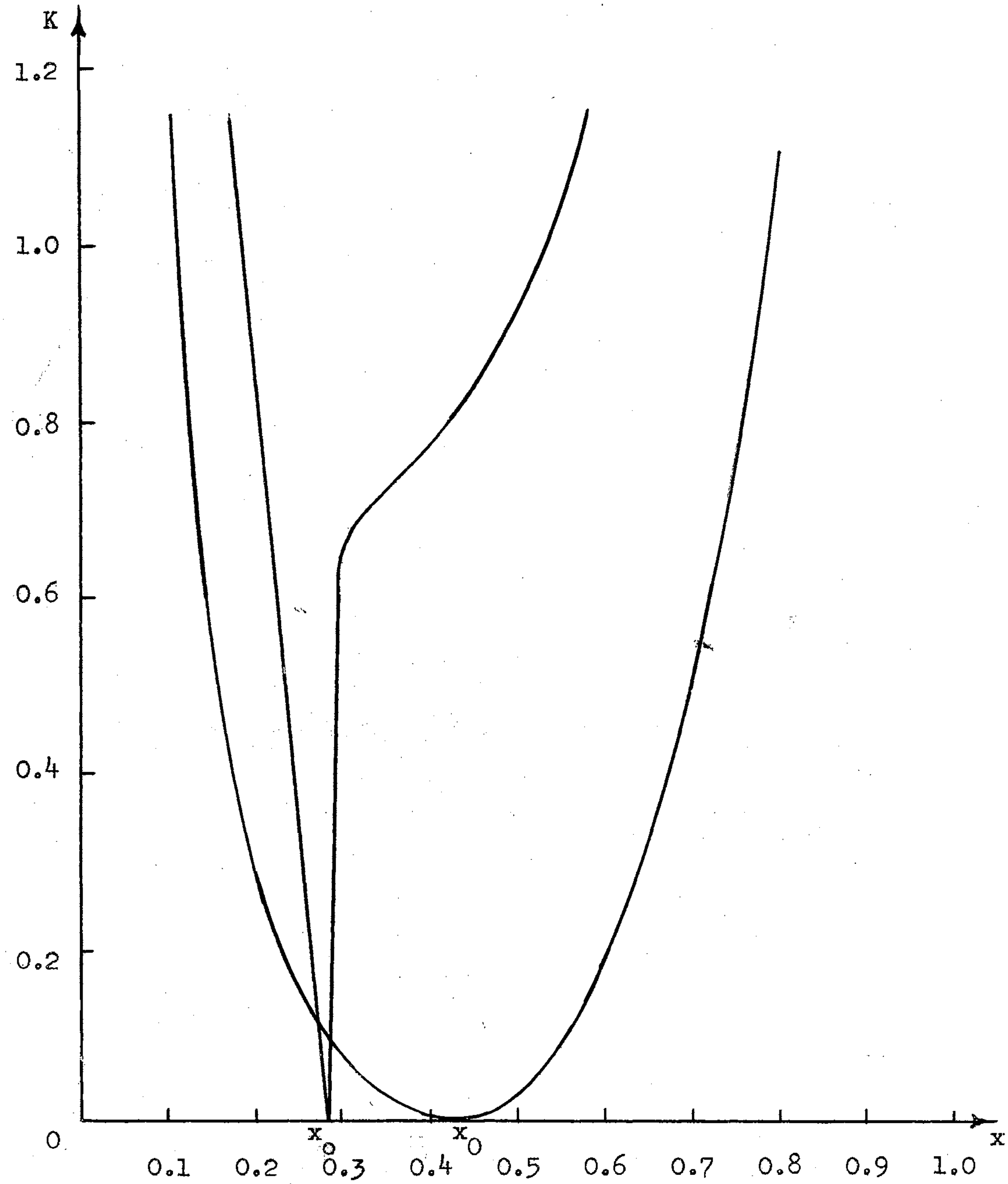


Figure 25. Effect of Variance on x_o and r

CHAPTER VII

SUMMARY AND CONCLUSIONS

The purpose of this concluding chapter is to summarize the research effort, draw conclusions based on results, and make proposals for future study. Hence, this chapter is concerned with three topics: the first will summarize the information presented by reviewing the contributions of each chapter; the second will draw conclusions relative to the results; and the third topic will present proposals for future action and study.

Summary

Chapter I served to introduce the queuing problem. It also involved discussion about framing, defining alternative solutions, and the solution to a queuing problem. In addition, specific definitions are given for some terms used in this treatise. Literature review was cited to indicate the state of development to date.

Chapter II reported the general and specific classifications of queuing theory. It also described the basic structure of the queuing problem. This chapter involves the philosophy of queuing theory in terms of the components of the queuing system and its characteristics drawing upon examples from everyday life. A general discussion of approaches to the classification of queuing situation is also presented.

Chapter III served to introduce the solution of queuing problem and the decision model. In addition, this chapter is to optimize and to give the sensitivity analysis of deterministic queuing models. Single channel-single phase and multiple channel-single phase of deterministic queuing models are presented in this chapter. The insignificant levels and the indifferent range of the service time are presented under sensitivity analysis of the mentioned models. The effect of the time between arrivals and the costs ratio "R" on the insignificant limits and the indifferent range of the service time are presented.

Chapter IV presented the optimization and sensitivity analysis of the probabilistic queuing models. Ordinarily, both the arrival rate and the service rate are expected values from a specified distribution. The considered distributions of the arrival rate and the service rate in this chapter were random variables from Poisson's distributions. The models are based on the assumption of an infinite population. A review of the expected total system cost derivation is given first. Analytical optimization and sensitivity analysis of the expected total system cost are derived in this chapter. The sensitivity analysis is derived at the lowest point of the expected total system cost surface. This case is named as a special case. In addition, the sensitivity analysis is considered at any arrival rate, not at the lowest point on the expected total system cost, and general case name is assigned to it.

Chapter V involved the optimization and the sensitivity analysis of the expected total system cost model when service is provided automatically by mechanical means, or when the service operation is mechanically paced. The service duration might be a constant. A procedure in six steps is given in this chapter to calculate the minimum expected total system cost service rate. Models of the insignificant limits and indifferent range are developed for the model under consideration: Poisson's arrivals with constant service time expected total cost model.

Chapter VI considered the optimization and the sensitivity analysis of the general model: Poisson arrivals with any service time distribution model. In this chapter, the expected total system cost model is related to the variance and the expected value of the service time distribution and the expected value of the arrival distribution. The insignificant limits and the indifferent range of the service rate are drawn. In addition, the effect of the service distribution variance on the insignificant limits and indifferent range is shown in this chapter.

Appendix A presents a complete derivation of the mean number of units and the mean waiting time in the system models presented in Chapter V and Chapter VI.

Conclusion

A general and specific classification of queuing theory has been developed in this treatise. The general classification of queuing theory depends on the population of individuals requiring service, number of queues in the system, and service facility. The specific classification of the theory depends on the assumption of the previous three properties in queuing system.

Too, the insignificant limits and the indifferent range of service time for deterministic queuing models have also been developed. These limits and indifferent range are developed as functions of R, A, and K. In addition, the effects of changing the values of R, A, and K on the insignificant limits and indifferent range are given.

Thirdly, the insignificant limits and the indifferent range of service rate of the probabilisitic queuing models have been developed. These are given as functions of R, λ, and K. These limits and range are derived for special and general cases. In addition, the effect of variance on the insignificant limits and indifferent range of service rate was shown in Chapter VI.

The analytical method was employed to optimize the queuing models under study. The majority of the models originated in this treatise are optimized.

The primary purpose of this treatise is to furnish the decision maker with the optimum service rate, and its insignificant limits models as a function of cost parameters. Also, it furnishes him with the effect of change of parameters on the total system cost. Through the provision of powerful quantitative tools, this will qualify the decision maker to answer questions which arise in the industrial environment. By calculating the insignificant limits of the service rate, the decision maker can directly decide whether it is worthwhile to change the situation to the optimal one. Additionally, he is qualified to answer questions about the effect of changing parameters under study.

Proposals for Future Study

This section involves the area of future research. It is recommended that further research be devoted to the sensitivity of queuing models. Specific topics that could be investigated are:

1. Specific measures of sensitivity which possibly could be developed and evaluated for other models in queuing theory not considered in this treatise.
2. Specific measures of sensitivity which possibly could be developed and evaluated for the effect of "wrong" values of decision parameters.
3. Investigation could be pursued in

developing a measure of sensitivity of the effect of "wrong" probabilistic arrival distribution.

4. A measure of sensitivity of the "wrong" model could be investigated.
5. A measure of sensitivity of the finite queuing model could be investigated.

SELECTED BIBLIOGRAPHY

(1) Buffa, E. S. Modern Production Management. New York: John Wiley and Sons, Inc., 1965.

(2) Fabrycky, W. J., and P. E. Torgersen. Operations Economy: Industrial Applications of Operations Research. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1966.

(3) Bowman, E. H., and R. B. Fetter. Analysis for Production and Operations Management. Homewood, Illinois: Richard D. Irwin, Inc., 1967.

(4) Saaty, T. L. Elements of Queuing Theory: With Application. New York: McGraw-Hill Book Co., Inc., 1961.

(5) Churchman, C. W., R. L. Ackoff, and E. L. Arnoff. Introduction to Operations Research. New York: John Wiley and Sons, Inc., 1957.

(6) Kaufmann, Arnold. Methods and Models of Operations Research. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1963.

(7) Takács, Lajos. Introduction to the Theory of Queues. New York: Oxford University Press, Inc., 1962.

(8) Prabhu, N. V. Queues and Inventories. New York: John Wiley and Sons, Inc., 1965.

(9) Goetz, Billy E. Quantitative Methods: A Survey and Guide for Managers. New York: McGraw-Hill Book Company, 1965.

(10) Hillier, Frederick S. "Cost Models for the Application of Priority Waiting Line Theory to Industrial Problems," Journal of Industrial Engineering, Volume XVI, No. 3, May-June, 1965.

(11) Hillier, Frederick S. "The Application of Waiting Line Theory to Industrial Problems," The Journal of Industrial Engineering, Volume XV, January-February, 1964.

(12) Bhatia, Arjan, and Anand Gary. "Basic Structure of Queuing Problems," The Journal of Industrial Engineering, Volume XIV, No. 1, January-February, 1963.

(13) Shelton, John R. "Solution Methods for Waiting Line Problems," The Journal of Industrial Engineering, Volume XI, No. 4, July-August, 1960.

(14) Flagle Huggins, and Roy. Operations Research and System Engineering. Baltimore, Maryland: John Hopkins, 1960.

(15) Barish, Norman N. Economic Analysis: for Engineering and Managerial Decision Making. New York: McGraw-Hill Book Company, Inc., 1962.

(16) Mjosund, Arne. "Book Review - Operations Economy by Fabrycky, W. J., and Torgersen, P. E.," Journal of Industrial Engineering, Volume XVIII, No. 2, February, 1967.

(17) Goode, Harry H., and Robert E. Machol. System Engineering. New York: McGraw-Hill Book Company, Inc., 1957.

(18) McMillan and Gonzalez. System Analysis - A Computer Approach to Decision Models. Homewood, Illinois: Richard D. Irwin, Inc., 1965.

(19) Miller and Starr. Executive Decisions and Operations Research. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1960.

APPENDIX A

PROBABILISTIC WAITING LINE MODELS

Arbitrary Arrivals With Arbitrary Service

Assume:

(1) Single-channel waiting line or queue

(2) λ = Expected number of arrivals per period

(3) μ = Expected number of service completions per period

(4) $\mu > \lambda \quad \therefore \frac{\lambda}{\mu} < 1.$

It follows that $\frac{\lambda}{\mu}$ = probability that channel is busy. Consider the instant when a unit C_0 is just leaving the channel. The queue length after its departure is n_0.

If $n_0 = 0$, the next unit to arrive, C_1, will be serviced immediately.

If $n_0 \neq 0$, the next unit, C_1, is just beginning its service time, Δt_1.

In either case, a number of units will arrive during Δt_1, and this number is r_1.

When C_1 leaves, the new queue length is n_1.

$$n_1 = r_1 \qquad \text{if} \qquad n_0 = 0$$

and

$$n_1 = n_0 + r_1 - 1 \qquad \text{if} \qquad n_0 \neq 0.$$

Since this is at the instant of departure, which is of infinitesimal length, the probability of another arrival being generated simultaneously is zero. These two cases are expressed by:

$$n_1 = n_0 + r_1 - 1 + \delta_0 \qquad \text{(A-1)}$$

in which

$$\delta_0 = 1 \qquad \text{if} \qquad n_0 = 0$$

and

$$\delta_0 = 0 \qquad \text{if} \qquad n_0 \neq 0.$$

The quantity δ_0 is a number that takes on only the values 0 and 1 and has an expected value lying between these two. Note that

$$n_0 \, \delta_0 = 0$$

and

$$\delta_0^2 = \delta_0 .$$

The waiting line is assumed stable. Stationarity exists when:

$$\frac{d(p(n))}{dt} = 0.$$

$$\therefore E(n_0) = E(n_1).$$

The expected value of Equation (A-1) relative to r_1 is

$$E(\delta_0) = 1 - E(r_1)$$

$$= 1 - \lambda \cdot \Delta t_1 ,$$

since the expected number of arrivals while C_1 is in the channel is the expected number of arrivals per period multiplied by the service time.

The expected value of δ_o relative to Δt_1 is

$$E(\delta_o) = 1 - \lambda \cdot E(\Delta t_1)$$

$$= 1 - \lambda(\frac{1}{\mu}) = 1 - \frac{\lambda}{\mu} ,$$

since $E(\Delta t) = \frac{1}{\mu}$ by definition.

Squaring Equation (A-1), one obtains

$$n_1^2 = n_o^2 + (r_1 - 1)^2 + \delta_o^2 + 2n_o(r_1 - 1) + 2\delta_o(r_1 - 1) + 2n_o \delta_o .$$

The last term is equal to zero. By substituting $\delta_o^2 = \delta_o$ and taking expected values, one has

$$E(n_1^2)=E(n_o^2)+E[(r_1-1)^2]+E(\delta_o)+2E[n_o(r_1-1)]+2E[\delta_o(r_1-1)]$$

Again, because of stationarity, $E(n_1^2) = E(n_o^2)$. Furthermore, the expected value operators are passed through the two product terms, and since r_1, the number arriving, is assumed independent of n_o, the number on line, it must also be independent of δ_o which depends only on n_o.

$$0 = E[(r_1-1)^2] + E(\delta_o) + 2E(n_o)E(r_1 - 1) + 2E(\delta_o)E(r_1 - 1).$$

Taking expected value relative to both r_1 and Δt_1, and substituting $E(\delta) = 1 - \frac{\lambda}{\mu}$ and $E(r) = \frac{\lambda}{\mu}$

$$0 = E(r^2) - 2\frac{\lambda}{\mu} + 1 + (1 - \frac{\lambda}{\mu}) + 2E(n)(\frac{\lambda}{\mu} - 1) + 2(1 - \frac{\lambda}{\mu})(\frac{\lambda}{\mu} - 1)$$

$$2E(n)(1-\frac{\lambda}{\mu}) = E(r^2) - \frac{3\lambda}{\mu} + 2 - 2 + \frac{4\lambda}{\mu} - 2(\frac{\lambda}{\mu})^2$$

$$= E(r^2) - 2(\frac{\lambda}{\mu})^2 + \frac{\lambda}{\mu}$$

$$E(n) = \frac{E(r^2) - 2(\frac{\lambda}{\mu})^2 + \frac{\lambda}{\mu}}{2(1-\frac{\lambda}{\mu})}$$

$$E(n) = \frac{E(r^2) - \frac{\lambda}{\mu}}{2(1-\frac{\lambda}{\mu})} + \frac{\lambda}{\mu} \qquad \text{(A-2)}$$

$E(n)$ = Expected number of units in the line

n_m = Mean number of units in the line

$$\therefore E(n) = n_m$$

$$n_m = \frac{E(r^2) - \frac{\lambda}{\mu}}{2(1-\frac{\lambda}{\mu})} + \frac{\lambda}{\mu}. \qquad \text{(A-3)}$$

This equation holds for arbitrary arrivals and arbitrary service, provided only that these distributions are independent of n and t and that $\frac{\lambda}{\mu} < 1$ for single-channel service. It may be noted that, although $E(r) = \frac{\lambda}{\mu}$, $E(r^2)$ is, in general, not equal to $\frac{\lambda}{\mu}$, and so Equation (A-3) says that as $\frac{\lambda}{\mu} \rightarrow 1$, $n_m \rightarrow \infty$.

If something is known about the distribution of the arrival-time intervals, then the $E(r^2)$ can be found, and, therefore, n_m can be evaluated.

Poisson Arrivals With Arbitrary
Service Time Distribution

Poisson probability distribution

$$p(k) = \frac{e^{-\mu t} (\mu t)^k}{k!}$$

where

p(k) = Probability of exactly k occurrences in a time interval t.

λ = Expected number of occurrences per unit time.

λ = The mean and the variance of the distribution.

Using this general equation in terms of the assumptions of the completely arbitrary model, one has

$$p(r_1) = \frac{e^{-\lambda \cdot \Delta t_1} (\lambda \cdot \Delta t_1)^{r_1}}{r_1 !}$$

where

r_1 = The number of units arriving during Δt_1

$\lambda \cdot \Delta t_1$ = The constant mean of the distribution and its variance.

From the general laws of probability

$$\sigma_{r_1}^2 = E(r_1^2) - [E(r_1)]^2 .$$

Hence,

$$E(r_1^2) = \sigma_{r_1}^2 + [E(r_1)]^2 = \lambda \cdot \Delta t_1 + (\lambda \cdot \Delta t_1)^2 .$$

Now, take the expected value of r_1 over all Δt_1's and obtain

$$E(r_1^2) = \lambda E(\Delta t_1) + \lambda^2 E(\Delta t_1^2) \qquad \text{(A-4)}$$

Again, from the general laws of probability

$$E(\Delta t_1^2) = \sigma_{\Delta t_1}^2 + [E(\Delta t_1)]^2 = \sigma_{\Delta t_1}^2 + \frac{1}{\mu^2}. \qquad (A-5)$$

Hence, from Equation (A-4) and Equation (A-5):

$$E(r_1^2) = \frac{\lambda}{\mu} + \lambda^2 [\sigma_{\Delta t_1}^2 + \frac{1}{\mu^2}]$$

$$E(r_1^2) = \frac{\lambda}{\mu} + \lambda^2 \ \sigma_{\Delta t_1}^2 + \frac{\lambda^2}{\mu^2}. \qquad (A-6)$$

Substituting Equation (A-6) into Equation (A-3)

$$n_m = \frac{\left[\frac{\lambda}{\mu} + \lambda^2 \ \sigma_{\Delta t_1}^2 + \frac{\lambda^2}{\mu^2}\right] - \frac{\lambda}{\mu}}{2(1 - \frac{\lambda}{\mu})} + \frac{\lambda}{\mu}$$

$$n_m = \frac{(\frac{\lambda}{\mu})^2 + \lambda^2 \ \sigma_{\Delta t_1}^2}{2(1 - \frac{\lambda}{\mu})} + \frac{\lambda}{\mu}. \qquad (A-7)$$

Equation (A-7) is Equation (6-1) in Chapter VI. Since

$$W_m = \frac{n_m}{\lambda}$$

$$\therefore W_m = \frac{1}{\lambda}\left[\frac{(\frac{\lambda}{\mu})^2 + \lambda^2 \ \sigma^2}{2(1 - \frac{\lambda}{\mu})} + \frac{\lambda}{\mu}\right]$$

$$W_m = \frac{(\frac{\lambda}{\mu^2}) + \lambda\sigma^2}{2(1 - \frac{\lambda}{\mu})} + \frac{1}{\mu}. \qquad (A-8)$$

Equation (A-8) is Equation (6-2) in Chapter VI. By putting $\sigma^2 = 0$ in Equation (A-7) and (A-8), they would be Equations (5-1) and (5-2) in Chapter V, respectively.

VITA

Ali Mohamed A. Ali

Candidate for the Degree of

Doctor of Philosophy

Thesis: SENSITIVITY ANALYSIS OF QUEUING MODELS

Major Field: Engineering

Biographical:

Personal Data: Born in Alexandria, U.A.R., November 7, 1936, the son of the late Mohamed A. Ali and the late G. M. Abdel-Kerim.

Education: Attended said EL-Awal primary school in Alexandria, U.A.R., in 1946; attended Alexandria secondary school in Alexandria, U.A.R., in 1951; received Bachelor of Engineering degree in Mechanical Engineering - Production special course - from Alexandria University, Alexandria, U.A.R., in June, 1961; and received the Master of Science degree, with major in Industrial Engineering and Management, from Oklahoma State University, Stillwater, Oklahoma, U.S.A., in May, 1965; re-entered Oklahoma State University in June, 1965, and completed requirements for the Degree of Doctor of Philosophy in May, 1968.

Professional Experience: Worked as an Instructor in Mechanical Engineering Department, Alexandria University, Alexandria, U.A.R., during the period October 1961-August 1963. Worked also as Instructor, part-time, for Naval Academy, Alexandria, U.A.R., during the period October 1962-August 1963. Employed as Graduate Assistant for School of Industrial Engineering and Management, Oklahoma State University, Stillwater, Oklahoma, from February 1965 to June 1966; employed by Lamar State College of Technology, Beaumont, Texas, as Assistant Professor of Industrial Engineering from September 1966 to July 1967

Professional Activity: Member of American Institute of Industrial Engineers, Alpha Pi Mu (honorary Industrial Engineering fraternity), and recipient of American Society for Testing and Materials Membership Award.