A STUDY OF THE CLASSIFICATION OF QUEUES FROM THE ENGINEERING STANDPOINT, WITH PARTICULAR CONSIDERATION TO THE RELATIONSHIP OF SERVICE RATES TO QUEUE LENGTH

By

SCOTT TABOR POAGE

Bachelor of Science
Texas Technological College
Lubbock, Texas
1953

Master of Science
Agricultural and Mechanical College of Texas
College Station, Texas
1957

Submitted to the Faculty of the Graduate School
of the Oklahoma State University
in partial fulfillment of
the requirements for
the degree of
DOCTOR OF PHILOSOPHY
August, 1962

A STUDY OF THE CLASSIFICATION OF QUEUES FROM
THE ENGINEERING STANDPOINT, WITH
PARTICULAR CONSIDERATION TO
THE RELATIONSHIP OF
SERVICE RATES TO
QUEUE LENGTH

Thesis Approved:

Wilson J. Bentley
Thesis Adviser

H. G. Thueson

Solomon Sather

Herbert L. Jones

David L. Weeks

Robert MacVicar
Dean of the Graduate School

PREFACE

The genesis of this work was in the difficulties which I faced as the teacher of introductory courses in operations research in trying to introduce the students to a number of the basic queueing models. While attempting to show the effects on the models of changes in the assumptions made, I found that the students, and often I, could not reasonably handle the differences between two situations without an extremely lengthy, time consuming detailed word description of the situation and the assumptions. Eventually, the search for better means of putting this information over to the student led to the ideas presented in this thesis. Since substantially completing this material, it has been used in note form for this introductory operations research course with, what I believe to be, a great deal of success as an aid to the student of queueing theory. Certainly there is great truth in the "old saw," "Good students are a teacher's sternest master."

While it may be that students are often stern masters to the teacher, the acquisition of knowledge not often is accomplished best through sternness. It is then with great gratitude that I acknowledge my deep debt to Professor Wilson J. Bentley, Head of the School of

Industrial Engineering and Management at the Oklahoma State University for his many kindnesses, constructive criticisms, cheerful encouragement and readily available counsel in his varied capacities as the chairman of my graduate committee, as my supervisor during my tenure as a member of his faculty, as a colleague, but most importantly, as my friend.

To the other members of my graduate committee: Professors Herbert L. Jones, Solomon Sutker, H. G. Thuesen, and David L. Weeks, my debts are hardly fewer or less deep. It would be impossible to enumerate them all, but most prominent are my coming to the Graduate School of Oklahoma State University, sound advice and encouragement since arriving, fair and useful criticism while working on this thesis, and enlightenment on many, many points. For all of these I am grateful.

A special tribute must be paid to Miss Velda D. Davis, who labored long and hard with the preparation of the typescript and, even more, handled, while I was away from the campus, all of those endless details which make completing graduate work while off campus nearly impossible without such a friend and helper.

A final debt which can never be approached is that which I owe to my family for their encouragement, understanding, and the special incentive they offered to me for the completion of this effort.

Scott Tabor Poage

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The study of the theory of "queues" or "waiting lines" refers to the mathematical and physical investigation of a class of systems typified characteristically by five features: (1) The units moving through the system are discrete, (2) A mechanism exists that governs the time at which the units composing the system arrive, i.e., begin to require service, (3) The units which have begun to require service are ordered in some fashion, which may be by a completely determined mechanism or by some probabilistic arrangement, and receive service in that order, (4) A mechanism exists that governs the time at which a unit receiving service has its service terminated, (5) At least one of the two mechanisms, arrival or service, is not completely determined, but can be considered a probabilistic system of some sort. These characteristics and their implications will be considered and extended later, but they, together with some examples, will serve as a basis for a preliminary discussion of the engineering importance and applications of queueing theory. Some of the classic systems to which queueing theory has been applied are the arrival of ships which wait in a harbor until they are unloaded, where the ships are the arriving units, the waiting ships form the queue or waiting line, and the unloading docks are the service mechanisms; the arrival of ticket purchasers who line up to purchase tickets from the service facility or ticket booth; the arrival of telephone calls at an exchange to be serviced by an operator or trunk line; the customers arriving in a barber shop who wait until their turn comes and they have their haircut, etc., by the serving unit, the barber; cafeteria lines; aircraft stacked up waiting to get into a given airport; aircraft on the apron waiting for modification, repair or clean-up; patients waiting to see a doctor, etc. Less obvious and perhaps more important to the industrial engineer are examples of systems to which the theory of queues has been applied through analogy with great success. Some of these applications have been such systems as a group of machines which require repair (arrive) from time-to-time and are idle (waiting in line) until the maintenance crew (service facility) can repair them and set them back to work; the flow of paper work through a production control system; an inventory system where the arrival is the demand by a customer for a unit of stock, the queue or waiting line is the number of units on backorder and the service operation is the process of replenishing the shelf stock by ordering from the wholesale house or manufacturer. (This last system may also be considered to be as a case where the shelf space is the service facility, the customer the departure mechanism and the item of stock the arriving unit.)

Systems such as these are types with which the industrial engineer works each day. Queueing theory attempts to develop answers to such questions as: "What is the average length of the line of waiting units?", "How many machines are needed to do a certain job in such a manner that the facilities for waiting are not overloaded or that so many of the working units are not waiting in line that it is not possible to maintain satisfactory service to the customers?", "How long will an arrival at the service unit have to wait before being served?...Before being completed?", "What is the probability of a unit having to wait longer than a given time?", "What is the probability of instantaneous service?", "What is the utilization efficiency of the machines doing the servicing?", "What is the utilization rate of the machines or things being serviced?". The importance of these questions to the industrial engineer, in such planning functions as plant layout, staffing, machine selection, production control, and inventory control, cannot be overestimated.

Although originally developed in connection with a real world engineering problem by a practicing engineer, queueing theory was largely ignored by the engineering world and most of the developments which were made in the area were undertaken by mathematicians who produced a great number of additions to the body of knowledge in the field, largely in the form of solutions to special cases and situations chosen for their mathematical interest rather than for their engineering importance, even though many of them are quite valuable in everyday engineering work.

It then will be the purpose of this thesis to present the results of an attempt to add to the usefulness to industrial engineers of the theory of queues by developing and testing a system of classification of queues with special emphasis on the testing of certain assumptions regarding the relationship between the service rate and the length of the queue waiting for service.

The classification system which will be presented here was designed to increase the usefulness of queueing theory to the working engineer by presenting a system of classification whereby the engineer may readily and conveniently classify the queueing situation with which he is faced and compare it with those situations for which known solutions exist. This classification system is tested in two ways, by classifying representative queueing situations for which known solutions exist and by a demonstration classification of a real world situation. The classification of the real world system is concerned primarily with the problem of the relationship between the state of the queue and the service rate of the system with the first attention being given to certain common assumptions regarding this relationship when the rate of service is controlled by human operators who act as the service mechanism.

CHAPTER II

HISTORICAL DEVELOPMENT OF QUEUEING THEORY

The historical development of queueing theory is a relatively recent development even in the young history of statistical applications to industrial problems. Generally, this development can be treated in three phases, the pioneering work of A. K. Erlang, the work done in the era between Erlang and the post World War II development of interest in "operations research," and the development of queueing theory since that time. The historical development of queueing theory is considered in this chapter in those phases.

Erlang's Initial Developments

The first person to become interested in the general class of queueing problems was a Dane, F. W. Johannsen (1) of the Copenhagen Telephone Company, who published an article, "Waiting Times and Number of Calls," in the British Post Office Electrical Engineers' Journal in 1907. The first to publish comprehensive theoretical considerations of queueing problems was A. K. Erlang (2), a close friend and co-worker of Johannsen at the Copenhagen Telephone Company. Erlang was a Danish electrical engineer and mathematician who was interested in the problems of waiting lines and waiting times by Johannsen, but went on to become much more deeply involved. By laying the foundations for most of the present work, he can be considered to be the originator of queueing theory as it is known today. In his first publication on the subject, "The Theory of Probabilities and Telephone Conversations," Erlang (2) pointed out that the waiting lines were characterized by an arrival distribution, a queue discipline or rule for determining the order of service of the arriving units and service distribution. He observed that telephone calls could be described as having Poisson arrival distributions, a first come, first served discipline, and an exponential distribution of service times. Later, he studied and built mathematical models for systems characterized by Poisson arrivals, first come, first served discipline, exponential service times, and several service channels; Poisson arrivals first come, first served discipline, regular service times and several service channels; Poisson arrivals, random order service discipline, and exponential service times; and Poisson arrivals, first come, first served discipline, one service channel and Erlangian service distributions (a special case of the Karl Pearson Type III distribution widely used by Erlang because of its ability to form an empirical approximation to a great many distributions in a continum between the exponential and constant). Two points seem to be worth noting about Erlang's work because it has continued to influence to some degree the thinking on queues down to the present time. First, his classification system is limited by the assumption that there are an infinite number of possible arrivals, and second that the discipline mechanism is limited to determining the order in which the arrivals to the line will be served. It is quite possible that Erlang was fully aware of the second point, but proceeded with the simpler cases as they could be solved, but the relationship of finite possible arrivals is not mentioned in any of the earlier writings.

Development Prior to 1950

The queueing problem attracted some attention among telephone engineers in Europe, but was not introduced in American until the publication of "Applications of the Theory of Probability to Telephone Trunking Problems" by E. C. Molina (3) in July of 1927. The work of Erlang and Molina was considerably expanded by Thronton C. Fry (4) in his book, Probability and Its Engineering Uses in 1928. Fry's work served to introduce the subject to American telephone engineers and mathematicians, but did not generate much use of the idea among engineers generally; however, it did attract attention to the subject by a number of European mathematicians. Fry considered the problem primarily in the light of telephone work, as did Erlang, although he pointed out many analogous situations where it was applicable. He regressed from Erlang on the matter of classification in that he did not attempt to classify the situation or solutions which he considered, but merely listed all of the assumptions which he made for each case. From the context, it would seem to be obvious that Fry recognized his solutions as different special cases of the same general problem, but he made no apparent attempt to show any structural relationship between the various solutions.

After the publication of Fry's book, there was a lull in the interest in the subject of queues in America and among engineers. The development of queueing theory then took a strongly theoretical mathematical turn. An extensive mathematical literature began to develop on the subject among European mathematicians with Pollaczek (5) studying the case of Erlangian arrivals, first come, first served discipline, general service distributions (independent) and one service channel in 1930; Khintchine (6), in 1932, investigating the Poisson arrivals, arbitrary service distribution with one servicing channel, and a first come, first served discipline; and in 1932, Pollaczek (7) extending Khintchine's work to the case with many servers generally. During the decade following 1930, the principal interest in the subject of queues remained with mathematicians and telephone engineers with a small, but increasing stream of publications on the subject, generally following the lines of the content of the work of Erlang, but on a much more sophisticated mathematical basis and with the principal interest residing in the mathematical technique involved rather than the application.

During the 1940's, the mathematical interest continued in the general subject of queues as such, but it began to be recognized that this problem was related to the subject of machine down time as a random process in such articles as "How Many Automatics Should a Man Run?" by Phillip Bernstein (8) and R. Kronig's (9) "On Time Losses in Machinery Undergoing Interruptions." All of these papers still followed the method of Erlang in assuming independence between the length of the waiting line or direct proportionality to the length of the waiting line in some form to the service time and in classifying the queues by the word description of the arrival distribution, service distribution and discipline.

Development Since 1950

The 1950's brought a fantastic growth in the literature on queues. In a bibliography of the literature which the author prepared for American Airlines, Incorporated, he found five articles before 1910, two between 1910 and 1920, three between 1920 and 1930, ten between 1930 and 1940, twenty-eight between 1940 and 1950, and over one hundred and seventy-five articles and books between 1950 and 1960. (10). This expansion was largely the result of the general growth of operations research and the mathematical techniques of industrial engineering. A large portion of this literature deals with the practical application of specific models developed by Erlang, although a few deal with the application of models developed later and a sizable number are devoted to the mathematical development of specific cases which are variations on the major theme of Erlang. In general, this massive literature is not germane to the questions considered here except as it emphasizes the growth of knowledge in the area and the need for a ready means of classifying a given situation and relating it to the known solutions.

The most important general developments in the theory of queues after World War II were the application of high speed computers to the numerical computation of queue results, the solution of the "Swedish Machine Problem," the publication of the book, Queues, Inventory and Maintenance, by Phillip Morse (11), and the publication of a system of classification of queues by David G. Kendall (12). Since Kendall's classification of queues is related to the primary subject of this thesis, it will be discussed in a separate section reviewing the development of classifications of queues while the other developments must be chosen here from among all of the recent works in queueing theory for the impact that they have had on the entire range of queueing problems, both directly and indirectly.

The effects of the modern high speed computer upon the theory of queues is in many ways representative of its influence on engineering in general. The computer changed the possible into the practical. The incentive thus given to workers in the area in turn led to whole fields of consideration which have not yet begun to be exploited. Many queueing problems had been solved with awesome mathematical brilliance only to be completely useless to the engineer or other day-to-day workers because either the length of computations required were humanly impossible or they were economically unfeasible. The computer reduced the time required for these calculations to the point of usefulness and made possible the solution of previously insoluble problems through numerical methods. Also under the contributions of the computer to queues must be listed Monte Carlo or simulation methods. These, while not strictly in the field of queueing theory, make possible the approximate solution of many otherwise unsolved problems. These Monte Carlo approximations are often as valuable economically as exact solutions and in other cases one may serve as a direction finder which leads to a rigorous and exact analytical solution. Furthermore, the computer made possible the publication of tables of queueing results which put answers to many problems on the desk of the worker who had not the time or had insufficient calculating facilities to use the known solutions. The availability of these tables and means of computation led to an expanded usefulness and use of queueing theory which in turn led to great attempts to solve new problems in the area.

This process is still going on.

The "Swedish Machine Problem" was so named because its solution was first published by Conny Palm (13), a Swede, in relation to one of the basic problems of industrial engineering, given that each of m machines has a given probability of going down at anytime while it is running and there are n repairmen, each of whom takes a given average time to repair one machine, what is the expected number of machines running, being serviced, and idle waiting for service. This problem because of its great importance in determining the most economic number of machines and/or repairmen had been the subject of a number of approaches by industrial engineers for some years. Professor H. G. Thuesen (14) produced one of the early attacks using the concept of probability in a paper relating to the economical number of repair crews for a given oil field. The importance of Palm's solution rests on three facts: First, it makes explicit the relationship between queueing theory and the machine interference problem which permits an interchange of ideas between the various works on the two problems; second, it provided an exact solution to the problem for a general number of machines and repair facilities which previous solution could economically do only for a reasonably small number; third, it opened up the entire area of queueing systems in which there are a limited number of units which could be related to the earlier works developed after the ideas of Erlang.

Dr. Morse's (11) book holds a unique importance in three aspects. It was the first serious book devoted completely to the theory of queues and its applications. It was an attempt to present a comprehensive introduction which would permit the reader to handle most queueing problems with which he might be faced. Finally, its uniqueness and relative completeness caused it to become a prime reference and popularizer of queueing theory, but its style of exposition (which was relatively obscure where compared to some of the originating publications) served to drive those interested to either search back through the original articles or to redevelop many of the arguments, both of which led to greater interest and ability for research in the field. The systems covered by Morse were essentially those for which the arrival and service distribution could be approximated by combining exponential distributions in such forms as the "Erlang" and the "hyper-exponential". Still the greatest achievement of the book was the widespread demonstration of the power of queueing methods and the exposition of the effects upon queues of various types of changes in the distributions and parameters of queueing situations.

Finally, special mention must be made of the book, Elements of Queueing Theory, by Thomas L. Saaty (15), which appeared as this research was reaching its conclusion. This most notable work, although often leaving the world of interest to the engineer for flights of purely theoretical interest, must be regarded as the complete source book for the state of the science of queueing theory at the time of its publication (September, 1961). While it contains very little original material, Mr. Saaty's book collects practically every known argument regarding queues which is of any interest, many times in greater detail than is available in the original published sources.

CHAPTER III

EXISTING SYSTEMS OF CLASSIFICATION OF QUEUES

Erlang's Method

Erlang (2), in his pioneering works, recognized that queues differed and believed that they could be completely described by giving the arrival distribution, the discipline, and the service distribution. His method was to give a word description of the queue's three basic characteristics as he understood them. Thus, Erlang would describe a classical queueing situation as "having Poisson arrivals with a mean arrival rate, λ, from a single source; having m service channels, each with identical exponentially distributed service times with mean service rate of μ each; with arrivals forming a single unlimited queue which enters the available (empty) service channels randomly on a first come, first served basis. There are several obvious faults in this method of description, but their existence must not be taken as a fault of Erlang's insight into the problem, but rather to the limited development of queueing theory at the time. For example, the most obvious fault, failure to recognize the difference between queues which have an infinite arriving population and those which have a finite arriving population, was due to the fact that Erlang did not consider finite arrival populations beyond reasoning that the Poisson provided a satisfactory approximation for large populations such as the telephone calls with which he was working. A second weakness, similarly attributable to the state of the development, is the bulkiness and unwieldiness of such descriptions which make the system very poorly adapted to handling a large number of different types of queues. For sometime, these faults were not recognized by workers in the area for much the same reasons that they were not apparent to Erlang and, thus, the practice continued of simply giving a word description to each case as it was considered or solved.

Kendall's Classification

The first recognition that the number of cases and types of problems which had been considered in the literature on queues had grown in quantity and diversity to the point where the bulk and unwieldiness of the word descriptions after Erlang's fashion came in a 1953 article by David G. Kendall (12), "Stochastic Processes Occurring in the Theory of Queues and Their Analysis by the Method of the Imbedded Markov Chain." In this system, Kendall concerns himself only with the steady-state cases (those in which distributions of arrivals, service, and the resulting solutions are independent of time) and ignores the effect of discipline, resulting in his classifying queues by their input or arrival distributions, service distributions, and number of channels, which he indicated by letters, in that order, separated by solidi; thus: input/service/number of servers. For the letter symbols, he used the following code:

Code Letter	Type of Service or Input	Distribution Function
D	Deterministic or regular	$A(u) = 0$ for $u < a$ $A(u) = 1$ for $u \geq a$
	Random or Poisson (exponential)	$A(u) = 1 - e^{-u/a}$
	Erlangian with K phases	$dA(u) = \frac{(k/a)^k}{(k)} e^{-ku/a} u^{k-1} du$
GI	Any distribution in which the intervals are independent of each other.	
G	No assumptions are made about the distribution except that it must exist.	

Thus, under Kendall's system of classification, Erlang's simplest case, that of Poisson arrivals from an infinite population, exponential service with one serving channel, having first come, first served discipline in a single queue allowing infinite length would be classified M/M/1 or the same situation where there were s serving channels with a unit going into the service facility by choosing at random from among the available channels would be classified as M/M/s. Naturally, M/M/s includes M/M/1 as the special case where s = 1, but it would be classified as a separate case where the solution of M/M/s is considerably simplified by setting s = 1. Kendall used this system to indicate the queueing situations for which solutions existed in the literature at that time. This classification system was adopted by some other authors following its presentation and is still used occasionally.

Kendall's system while improving over the word descriptions which had preceded it, did so at the expense of flexibility, exactness, and the ability to be extended to cover queueing situations which are structured and/or disciplined differently from that for which it was originally set up. The most important criticism of Kendall's classification system is its limitation to a single queue structure and discipline.

Galliher's Extension of Kendall's Classification

The next development in the classification of queues was published by Herbert T. Galliher (16) in a collection of notes originally prepared for the Massachusetts Institute of Technology short courses on operations research. This system was an extension of Kendall's work with the addition of a notation or classification for known solutions which could be associated with a particular problem. As does Kendall's system, it makes the assumptions that all parameters and the number of servers are stationary and mutually independent, that arrivals are homogeneous from an infinite population, that servers are identical, and that the queue discipline is single queue, first come, first served, with no arrivals leaving before the completion of service, but differs from Kendall in that it recognizes the existence of queues outside of these rather limiting assumptions which it handles by classifying them as far as is possible by the Kendall notation and then making a word note of the exceptions to the basic assumptions such as classifying the Swedish machine problem as "MMS, finite population". This method of handling the extension of Kendall's system is, of course, better than ignoring the cases outside of the system, but it does have all of the weaknesses of the plain word description used by Erlang and those immediately after him with the two exceptions that it recognizes cases unstudied and/or unrecognized by them and it gains the compactness of Kendall's system whenever and to the extent that Kendall's system applies. The greatest contribution of the Galliher method is that it associates a classification or uniform notation for solutions associated with the problems thus classified.

The classification system for analytic results presented by Galliher (16) uses the following notation:

	to indicate that the probability distribution of the number of units waiting plus those in service is known for any time, t.
W_t	to indicate that the probability distribution of waiting time is known for any time, t.
N_t	to indicate that only the expected value of N_t is known.
W_t	to indicate that only the expected value of W_t is known.

Under this system, $t = \infty$ is taken to mean that only the steady-state or equilibrium case is solved. Where both N_t and N_∞ are shown, Galliher uses this to indicate that for $t = \infty$, the formulae are considerably easier to employ than for finite values of t. This double classification system within the limits imposed by the method used for the classification of the problems serves a number of useful purposes. The most important of these are that it facilitates communications between researchers in the field, permits a ready location of the known theoretical solutions, and facilitates the determination of the means of solving real world problems. In addition to the limitations already noted, all of the systems described so far share a common major weakness which the system whose development is reported here was designed to overcome. This weakness is the failure to show structural relationships between given problems of different classes and between different solutions.

Moore's Presentation

The most recent approach to this problem while not strictly a system of classification was provided by an article, "To Queue or Not to Queue," by James M. Moore (17) in which the structure of queueing problems was presented in the form of an organization chart reproduced in Figure 1. This chart is not strictly a classification structure in the sense that a queue may be placed at any one point

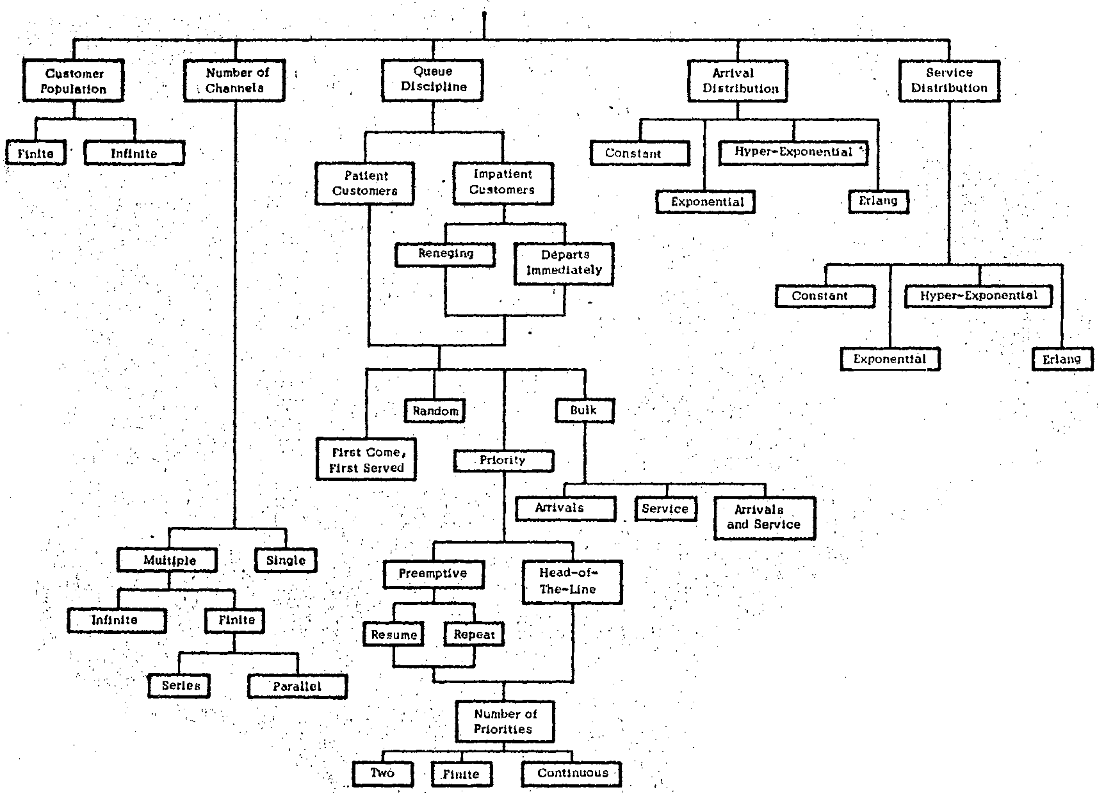

Figure 1. Moore's Organization Chart

on the chart, but is rather a series of five classification systems, one each for customer population, number of channels, disciplines, arrival distributions, and service distributions. Other objections which may be raised to this system are that the classes under the various characteristics are not mutually exclusive nor are they exhaustive. (This last requirement is almost impossible for any classification system; but, unless the classes are exhaustive, the fact that they are not exhaustive should be acknowledged at least by the inclusion of a class "other.") Mr. Moore's (17) presentation which in essence constitutes a dramatic, graphic presentation of Erlang's (2) classification has two very useful characteristics: First, it is expanded to cover most of the cases which have been considered to date and it retains the flexibility to be expanded to cover future cases; secondly, it shows the relationship in structure between some characteristics well although it is not consistent in this aspect. For example, it implies that a queue may have both patient customers and random service order, but does not under the same convention imply that a queue may have a finite customer population and multiple channels. Finally, this system makes no attempt to relate the structure of the queues to the solutions.

Saaty (15) presents a considerably extended version of Moore's chart in his introductory chapter on the description of queues and occasionally uses Kendall's (12) classification, but, in the main, relies on extensive word descriptions.

CHAPTER IV

PROPOSED SYSTEM OF CLASSIFICATION OF QUEUES

The system of classifying queues presented here is an attempt to satisfy six requirements which seem to be necessary for improvement of the existing systems described in the preceding chapter. First and most importantly, this system must be useful to working engineers as a means of describing a situation which faces them in their work, as a means of relating their situation to known solutions and as a means of communicating between workers who are dealing with queueing problems. The other requirements are direct consequences of this first requirement. As nearly as possible, such a system must be capable of covering any situation in its area. It must be as compact as possible. It should relate the structure of the queue to the method of solution as much as possible. It should be as convenient for printing or typing as possible. It should be capable of expansion to meet new types of queueing situations as they develop. Obviously, some of these requirements are, to a degree, contradictory, i.e., compactness versus comprehensiveness, but it is hoped that this strain is minimized by the use of a contracting notation for the more general cases and for the simpler cases.

Organization of Queues

The view taken here is that any queueing situation is in fact composed of fundamental queues and may be described in terms of these fundamental queues. Thus, a production line may be considered to be a number of fundamental queues in series with the arrivals to the first station being governed by the arrival of orders, with each succeeding station's arrivals being the departures of the preceding station and final output of the system being the departures of the last station. For the job shop operation, this group of queues which are arranged in series and in parallel may become as complex an arrangement as any electrical circuit diagram and, to a point, the analogy is rather strong. The structure of the method of solution of certain of these networks of queues was discussed quite ably by James R. Jackson (18), but for the purposes at hand it is sufficient to point out that the description of networks is best accomplished by diagraming the connections from the individual fundamental queues. The classification system here is concerned with the fundamental queue.

The Basic Parts of the Fundamental Queue

The fundamental queue consists of three basic parts, the arrival mechanism, the queues, and the departure mechanism. When these three parts and their interrelationships are defined, the queue and its characteristics are fixed.

Therefore, the classification of the queue will be given in three parts with each part separated by a solidus after the fashion of Kendall (12) in the following order:

Arrival Populations/Queues/Departure Mechanisms .

Arrival Populations

The description of each arrival population will be given in the following order and manner. Each population will be identified by the letter A with a subscript numerically identifying this type of population, a colon, a number indicating the number of identical populations of this type with a capital P indicating a general number, a dash, a number indicating the effective size of the arriving population with the capital letter N indicating a general number, a description of the distribution function of the inter-arrival times of this population in parenthesis and description of the disposition of the arrivals. Each arrival population will be separated from the next by a double colon, ::. The description of the distribution function of the inter-arrival times will be in the following manner, which is derived from the notation of Kendall (12), a code indicating the type of distribution, a colon, and the parameters of the distribution separated by commas. In the general cases, only the parameters will be given; but, in specific cases, each parameter will be followed by an equality sign and its value. A table of the code which is an expansion of that originally set up by Kendall is shown in Table I. Obviously, it is to be expected that from time-to-time, this table will require further additions and/or modifications as the state of the science progresses or as statisticians improve the means of identifying and classifying distributions, but it is suggested that for distributions of rare appearance that the code R_i with the subscript indicating a note completely specifying the distribution to be used and a U be used for an unidentified distribution.

The section describing the disposition of arrivals from a given population should be as follows. Units going to a specific queue would be indicated by the capital letter Q with a double subscript, the first, a number indicating the type, and the second, indicating the queue of that type. Similarly, the units going directly into one of the departure mechanisms would be indicated by the capital letter D and the same subscripting method. Each capital letter would be followed by a comma and a notation of the conditions of the queues and/or departure mechanism would allow this disposition following the usual notation of queues given by Morse (11) to describe the states of the queues and the departure mechanisms. Where more than one queue or departure mechanism can meet the conditions, the selection is presumed to be randomly selected with each mechanism having an equal probability of selection. If the selection is random, but not equal, the relative frequencies will be indicated by enclosing the mechanisms

TABLE I

INTER-ARRIVAL/SERVICE TIME DISTRIBUTION CODES

Code Letter	Type of Arrival or Service	Distribution Function
D	Deterministic or Regular	$A(u) = 0$ for $u < a$ $A(u) = 1$ for $u \geq a$
M	Random or Exponential	$A(u) = 1 - e^{-u/a}$
E_k	Erlangian with k phases	$dA(u) = \frac{(k/a)^k e^{-ku/a} u^{k-1}}{k} du$
GI	Any distribution in which the intervals are independent of each other	
G	No assumptions are made about the distribution except that it exist	
H_l	Hyper-exponential, L branches	
χ^2	Chi Square	
PIII	Karl Pearson Type III[1]	
N	Normal	
K,i,j,...	Any distribution with parameters i, j, ... known	
R_i	Not coded: see note i	
U	Not known	

[1]It should be noted that the χ^2, E_k, M and D distributions may all be obtained from the PIII by the proper selection of parameters, but their greater simplicity and in the case of M and D, their logical basis have made them so widely used that they should be noted separately for convenience.

to which the arrival may be disposed with parenthesis and preceding the expressions by their relative frequencies. Where disposal is to be to a particular mechanism for all units not going to another mechanism, this may be indicated by giving the first followed by a comma, a minus sign and the second. Each arrival population of the elemental queue would be thus described with a double colon, ::, separating each population.

To exemplify this method of classification of the arrival populations, a few of the classic cases will be given and then their arrival mechanisms placed in this classification system.

The original single queue with an unlimited population which arrived randomly at a mean rate of λ and a single queue of unlimited length, having a first come, first served discipline to a single service mechanism with random service completion times, would be:

$$A_1:1 - \infty(M,\lambda)Q_{11},\ n \neq 0;\ D_{11},\ n = 0/ \quad .$$

For the Swedish machine case, the arrival population description would be:

$$A_1:1 - N(M, \tfrac{1}{A})Q_{11},\ n \geq M;\ D_{1i},\ n_{1i} = 0/ \quad .$$

Queues

The description of the queues in the fundamental queue must state four things about the queues: the allowed length, the order of service and the disposition of units which leave the queues. These are classed in this fashion: A capital letter Q with a subscript i numerically identifying this type of queue, a semicolon, a number indicating the number of queues of this type with a Q indicating a general number, a dash indicating the maximum allowed length for this type of queue, a parenthesis containing either the letter F, R, L, or B to indicate the order in which the units in this queue are removed. Where F indicates first come, first served, R indicates random selection, L indicates last come, last served, and B indicates all at once, with the disposition following the parenthesis in the same fashion as the arriving units.

In some cases, it will be found that the units leave a particular queue according to two different disciplines, such as units going into service coming from the ordered queue on a first come, first served basis while units which are leaving the queue to try their luck in another system might be on a first come, last to leave. In this case, the first discipline will be given, followed by the disposition of the units leaving under that discipline followed by the second discipline and the disposition of the units leaving under that discipline.

Following this convention, the original single queue with an unlimited population which arrives randomly at a mean rate of λ and single queue of unlimited length, having a first come, first served discipline to a single service mechanism would have its queue described by:

$$/\ Q_1:\ 1 - \infty(F)D_{11},\ d_{11} = 0/$$

or the Swedish Machine Case by:

$$/\ Q_1:\ 1 - N\text{-}M(F)D_{1i},\ d_{1i} = 0/\ .$$

Departure Mechanism

The description of the departure mechanism follows the pattern of the arrival populations and queues. The identification begins with the capital letter D_i with the subscript i identifying numerically this type of service mechanism followed by a colon followed by a number indicating the number of identical mechanisms of this type with the capital letter M indicating a general number in the same fashion as the arrival time distribution. This is followed by the disposition of units which completed service, using O for out of the fundamental queue, A_{ij} for return to the jth population of type i, Q_{ij} for the jth queue of type i (for units displaced by priority units).

Thus, the original single queue with an unlimited population which arrives randomly at a mean rate of λ and single queue of unlimited length, has a first come, first served discipline to a single service mechanism with random service completion times and a mean service rate of μ would be indicated by:

$$/D_1:1(M,\mu)\ O$$

or for the Swedish Machine case:

$$/D_1:M(M,\mu)\ A_{11}\quad .$$

Identification of Results

Finally, it is highly desirable to be able to compactly describe the results of the queueing system which has been described analytically. This is essentially the matter of adopting a consistent, compact notation which follows the usual practice of statistics and yet is as mnemonic as possible. While it would be desirable to follow the literature in such notation, the literature in the area is remarkably unstandardized even for a mathematical or statistical subject of its relative youth. This makes it relatively difficult to follow the literature, yet whereever there is a fair degree of consistency, such as the use of n to indicate the total number of units in the system or p_n to indicate the probability of there being that number of units in the system, the lead is followed in order that this attempt will add as little additional diversity as possible to an already highly diverse subject. With this in mind, it seems wise simply to adopt the adequate existing system used by Galliher (16) of indicating by various appendages to the symbols for the characteristics of the queueing system the things that are known about the system. Thus, the subscript, t, is used to indicate that the probability distribution of that characteristic is known as a function of time for all values of t, which the subscript ∞ is used to indicate only the steady-state or equilibrium solution is known. Placing a bar over the symbol for the characteristic will indicate that only the expected or average value of that characteristic is known. Similarly, a lower case sigma, σ, subscripted by the characteristic will indicate only that the standard deviation of that characteristic is known.

Since many of the known partial solutions are valid only for a particular range of the parameters, such as n_∞ for the original Erlang model of $A_1{:}1-\infty(M,\mu)Q_{11}$, $n > 0$; D_{11}, $n = 0/Q_1{:}1 - \infty(F)D_{11}$, $n = 0/D_1{:}1(M,\mu)0$ which is valid only for values of $\lambda/\mu < 1$, such limitations should be noted in the usual mathematical manner.

Generally, the notation for the characteristics of interest of queues will follow the most common notation of the literature except in certain cases where it seems advisable to avoid confusion caused by conflicting notation or conflict with other general symbols. The most important symbols adopted here are:

n = number of units in the system

n = number of units in the queues plus the number in the service facilities

q_{ij} = number of units in the queue number j of type i

w = waiting time or time in the queue

s = time in the system or waiting time plus time in service

d_{ij} = number of units in the jth service facility of type i.

CHAPTER V

APPLICATION OF THE SYSTEM TO QUEUES CONSIDERED IN THE LITERATURE

The first demonstration of the usefulness of this system of classification of queues was achieved by classifying under the system a number of the situations described in the literature whose solutions are known and some whose complete solutions are not known. The following list, then, presents the common name applied to them (if a name has been assigned), describes the queue in the usual word manner, the classification of the queue, and the results which are available concerning that queue. Associated with some of the results will be two bibliographical references; the first being the original publication reference and the second being a convenient reference where the same results are available.

The Basic Poisson Queue

The most widely known queueing situation, the Erlang or basic Poisson Queue is usually described as a case where there is an infinite calling population from which units arrive randomly, i.e., that the inter-arrival times are as likely to end at any one instant as any other, with a mean arrival rate of λ, with the arriving units going directly into the service mechanism if it is empty, otherwise entering a queue of units waiting for service. The units which enter the queue leave only to enter the service mechanism on a first come, first served basis. The service times are also random in that the service of a unit in the mechanism is as likely to end at any one instant as another with a mean service rate of μ.

This queue would be classified:

$A_1{:}1 - \infty(M,\lambda)Q_{11}$, $n > 0$; D_{11}, $n = 0/$

$Q_1{:}1 - \infty(F)D_{11}$, $d_{11} = 0/D_1{:}1(M,\mu)0$

n_∞	(1), (15), (11)	n_t	(19), (15)
$\bar{n}_\infty$	(1), (15), (11)	$\bar{q}_\infty$	(11), (15)
w_∞	(15), (11)	$\bar{w}_\infty$	(15)
s_∞	(11)	$\bar{s}_\infty$	(15)
σ_{nt}	(15)	$\sigma_{n\infty}$	(11), (15)

Distribution Variations on the Basic Queue

The queues which follow immediately are a group which have received a great deal of study because they are basically variations on the basic queue given above. Their word descriptions are identical to that for the basic queue except for the distributions, their inter-arrival times, and service completion times. For this reason, only the distribution differences will be noted and the remainder of the word descriptions will be ommitted. The classification will be given in full.

The Deterministic Queue: This queue differs from the basic queue in that the inter-arrival times and the service times are constants.

$A_1 : 1 - \infty(D, \lambda) Q_{11}$, $n = 1$; D_{11}, $n = 0 / Q_1 : 1 - \infty(F) D_{11}$, $d_{11} = 0 /$

$D_1 : 1(D, \mu) 0$.

$\bar{s}_\infty$	(15)	s_∞	(15)
w_t	(15)		

Erlangian Service Times: This queue differs from the basic queue in that the service times follow the Erlangian distribution.

$A_1 : 1 - \infty(M, \lambda) Q_{11}$, $n \geq 1$; D_{11}, $n = 0 / Q_1 : 1 - \infty(F) D_{11}$, $d_{11} = 0 /$

$D_1 : 1(E_k, \mu, k) 0$.

n_∞	(15)	$\bar{\bar{n}}_\infty$	(15)
$\bar{q}_\infty$	(15)		

Erlangian Arrival Times: This queue substitutes the Erlangian distribution for the random distribution of the service times.

$A_1 : 1 - \infty(E_k, \lambda, k) Q_{11}$, $n \geq 1$; D_{11}, $n = 0 / Q_1 : 1 - \infty(F) D_{11}$, $d_{11} = 0 /$

$D_1 : 1(M, \mu) 0$.

n_∞	(20), (11), (15)	(20), (15)
n_∞	(20), (11), (15)	
$\bar{q}_\infty$	(20), (11)	

<u>Erlangian Input, Constant Service Times:</u> In this model, the distribution of the inter-arrival times is Erlangian and the service times are constant.

$A_1:1-\infty(E_k,\lambda,k)Q_{11},\ n\geq 1;\ D_{11},\ n=0/\ Q_1:1-\infty(F)D_{11},\ d_{11}=0/$

$D_1:1(D,\mu)0.$

w_∞	(15)

<u>Arbitrary Service Time Distribution:</u> Here, the service time distribution.

$A_1:1-\infty(M,\lambda)Q_{11},\ n\geq 1;\ D_{11},\ n=0/Q_1:1-\infty(F)D_{11},\ d_{11}=0/$

$D_1:1(G)0.$

n_∞	(15)

<u>Arbitrary Service and Arrival:</u> Here, the only assumptions made about the arrival and service distributions are that they exist and that each service and inter-arrival time is independent of the others.

$A_1:1-\infty(GI)Q_{11},\ n\geq 1;\ D_{11},\ n=0/Q_1:1-\infty(F)D_{11},\ d_{11}=0/$

$D_1:1(GI)0.$

	(21), (15)	(21), (15)

<u>Hyper-Exponential Service Time Distribution:</u> This situation is exactly like the basic queue except that the

service times are distributed according to the hyper-exponential distribution with two phases.

$A_1:1-\infty(M,\lambda)Q_{11},\ n \geq 1; D_{11},\ n=0/Q_1:1-\infty(F)D_{11}, d_{11}=0/$

$D_1:1(H_L,\mu,L=2)0.$

(11) n_∞ (11)

Two Phase Erlangian Arrivals: Here is a specialized version of the general Erlangian arrivals where the Erlangian distribution has the parameter $k=2$.

$A_1:1-\infty(E_k,\lambda,k=2)Q_{11}, n \geq 1; D_{11}, n=0/Q_1:1-\infty(F)D_{11}, d_{11}=0/$

$D_1:1(M,\mu)0.$

(11) (11)

(11) (11)

Erlangian Arrivals and Service, Both Two Phases: This system has both service and arrival time distributions of two phase Erlangian type.

$A_1:1-\infty(E_k,\lambda,k=2)Q_{11}, n \geq 1; D_{11}, n=0/Q_1:1-\infty(F)D_{11}, d_{11}=0/$

$D_1:1(E_k,\mu,k=2)0.$

(11) (11)

q (11)

(11) (11)

(11)

Two Phase Hyper-Exponential Arrivals: This system substitutes a two phase hyper-exponential distribution for the random arrival distribution of the original queueing case. $A_1:1-\infty(H_L,\lambda,L=2)Q_{11},n>1;D_{11},n=0/Q_1:1-\infty(F)D_{11},d_{11}=0/D_1:1(M,\mu)0$.

$$(11) \qquad n_\infty \qquad (11)$$

$$(11) \qquad\qquad (11)$$

$$(11)$$

Multiple Channel Queues

The members of the following group of queues each have more than one service channel. The basic multiple channel model is given first and is described completely in the word description. The remaining queues in this section are not completely described by words, but rather have given only their differences from the basic multiple channel situation given with all other features assumed to be identical. The classifications are given completely.

The Exponential Multiple Channel Queue: This queueing situation is the case where there is an infinite calling population from which units arrive randomly, i. e., the inter-arrival times are as likely to end at any one instant as at any other, with a mean arrival rate of λ, and with the arriving units going directly into an open service mechanism (if there is one) which is chosen at random from among the open service channels. If no service channels are empty when a unit arrives, it joins a queue (which may become infinite) from which it may not depart except to enter a service channel. The units leave queue to enter an empty service channel on a first come, first served basis. The service times of all channels are identically distributed according to the exponential distribution with identical parameters. The units leaving the service mechanism leave the system.

$A_1:1-\infty(M,\lambda)Q_{11},n\geq M;D_{1i},n<M,d_{1i}=0/Q_1:1-\infty(F)D_{1i},d_{1i}=0/$
$D_1:M(M,\mu)0.$

n_∞	(11), (15)	s_∞	(11)
$\bar{n}_\infty$	(11), (15)	$\bar{s}_\infty$	(11)
q_∞	(11), (15)	w_∞	(11)
$\bar{w}_\infty$	(11)	n_t	(15)

Constant Service Times, Multiple Channels: In this situation, all of the service channels have identical constant service times.

$A_1:1-\infty(M,\lambda)Q_{11},n\geq M;D_{1i},d_{1i}=0,n<M/Q_1:1-\infty(F)D_{1i},d_{1i}=0/$
$D_1:M(D,\mu)0.$

$\bar{n}_\infty$	(22), (15)	$\bar{q}_\infty$	(22), (15)
$\bar{w}_\infty$	(22), (15)	w_∞	(23), (15)
s_∞	(23), (15)		

Erlangian Two Phase Arrivals With Two Service Channels: This case substitutes a two phase Erlangian arrival distribution for the random arrivals and limits the number of channels specifically to two.

$A_1:1-\infty(E_k,\lambda,k=2)Q_{11},n\geq 2;D_{1i},d_{1i}=0,n<2/$

$Q_1:1-\infty(F)D_{1i},d_{1i}=0/D_1:2(M,\mu)0.$

n_∞	(11)	$\bar{n}_\infty$	(11)
$\bar{q}_\infty$	(11)		

<u>Two Erlangian Two Phase Service Channels:</u> This case substitutes a two phase Erlangian service distribution for the random distribution and specializes the number of multiple channels to two.

$A_1:1-\infty(M,\lambda)Q_{11},n\geq 2;D_{1i},d_{1i}=0,n<2/Q_1:1-\infty(F)D_{1i},d_{1i}=0/$
$D_1:2(E_k,\mu,k=2)0.$

n_∞	(11)	$\bar{n}_\infty$	(11)
$\bar{q}_\infty$	(11)	w_∞	(11)
s_∞	(11)		

Limited Queues

The following group of queueing situations are those for which the queue is not allowed the possibility of becoming infinite. Since, in most cases, these are variations of other queues, their full word description will not be given, but again the nearest previously described queue will be indicated with the pertinent differences noted.

Basic Single Channel System With No Queue: This case follows the basic model except that, if a unit arrives while another unit is in service, it returns to the calling population.

$$A_1:1-\infty(M,\lambda)A_{11},n\geq 1;D_{11},n=0//D_1:1(M,\mu)0.$$

$$n_t \qquad (11)$$

Basic Multi-Channel Case With No Queue: This case follows the basic multi-channel model except that, if a unit arrives while all of the service channels are filled, it returns to the calling population.

$$A_1:1-\infty(M,\lambda)A_{11},n\geq M;D_{1i},d_{1i}=0,n<M//D_1:M(M,\mu)0.$$

$$n_\infty \qquad (11) \qquad n_\infty \qquad (11)$$

$$(11)$$

Single Channel Erlangian Arrivals With Queue Length Limited: This case is the same as the single channel Erlangian arrivals except that the queue length is limited to Q units with those units which arrive when the service channel is filled and the queue contains Q units will return to the arriving population.

$$A_1:1-\infty(E_k,\lambda,k)Q_{11},Q+1>n\geq 1;\ A_{11},n\geq Q+1;D_{11},n=0/$$

$$Q_1:1-Q(F)D_{11},d_{11}=0/D_1:1(M,\mu)0.$$

$$n_\infty \qquad (11) \qquad n_\infty \qquad (11)$$

Two Phase Hyper-Exponential Arrivals With No Queue: This system corresponds to the single channel two phase hyper-exponential arrivals except that no queue is allowed so that units which arrive when the service channel is filled return to the arriving population.

$A_1 : 1 - \infty(H_L, \lambda, L = 2)A_{11}, n > 0; D_{11}, n = 0 // D_1 : 1(M, \mu).$

(11)

Single Channel Hyper-Exponential Service With No Queue: This system is like the single channel hyper-exponential service system except that units arriving when the service mechanism is filled are returned to the arriving population.

$A_1 : 1 - \infty(M, \lambda)A_{11}, n \geq 0; D_{11}, n = 0 // D_1 : 1(H_L, \mu, L)0.$

(11) (11)

Basic Single Channel Case With a Limited Queue Length: This is the same system as the basic single channel case except that the queue length is limited to Q units with arrivals, when both the service mechanism and the queue are filled, returning to the arriving population.

$A_1 : 1 - \infty(M, \lambda)A_{11}, n \geq Q + 1; Q_{11}, Q + 1 > n \geq 1; D_{11}, n = 0/$

$Q_1 : 1 - Q(F)D_{11}, d_{11} = 0/D_1 : 1(M, \mu)0.$

w_∞	(15)	w_∞	(15)
n_∞	(11), (15)	q_t	(11)
$\sigma_{n\infty}$	(11), (15)	n_t	(11)
n_∞	(15)	$\bar{q}_\infty$	(11)

Basic Multiple Channel Case Where There Are an Infinite Number of Channels: This case is identical to the basic multiple channel case except that the number of channels is considered to be infinite. Because there is always a service channel available, no queue will ever form.

$A_1 : 1 - \infty(M,\lambda)D_{1n}//D_1 : \infty(M,\mu)O.$

n_t (15)

Epidemic Case: This is a variation on the basic single channel model except that it is assumed that the Poisson arrival rate is proportional to the number present in the system and that the service is exponential with a service rate that is proportional to the number in the system. It is further assumed that once the system is completely empty the operation stops or if the number in the queue reaches a number Q, the operation stops.

$A_1 : 1 - \infty(M,n\lambda)Q_{11}, Q + 1 > n > 0; A_{11}, n = 0/Q_1 : 1 - Q(F)D_{11}, d_{11} = 0;$
$Q_{11}, n = Q + 1/D_1 : 1(M,n\mu)O.$

Distribution of time to reach stop. (24).

Additional Complicated Cases

The following group of cases represent some of the more complex assumptions which may be made regarding queueing situations for which solutions are available.

Balking Arrivals: This case is a modification of the single channel basic queue in which the arrivals look at the queue upon the instant of their arrival and decide either to enter the queue or to return to the population. For a given queue length, the decision is a random one with a fraction $e^{-\alpha n/\mu}$ deciding to enter the queue when there are n units in the system and a fraction $1-e^{-\alpha n/\mu}$ choosing not to enter the queue.

$A_1:1-\infty(M,\lambda)e^{-\alpha n/\mu}(Q_{11},n>0);(1-e^{-\alpha n/\mu})(A_{11},n>0);$

$D_{11},n=0/Q_1:1-\infty(F)D_{11},d_{11}=0/D_1:1(M,\mu)0.$

(11) (11)

(11) $\sigma^2{}_{n\infty}$ (1)

Balking Arrivals With an Arbitrary Arrival Distribution: This case is different from the basic single channel limited queue situation in that the arrival distribution is arbitrary.

$A_1:1-\infty(G)A_{11},n\geq Q+1;Q_{11},Q+1>n\geq 1;D_{11},n=0/$

$Q_1:1-Q(F)D_{11},d_{11}=0/D_1:1(M,\mu)0.$

n_∞ (15)

Balking and Reneging: In this case, the arriving units are constantly evaluating the length of the line. Upon arrival, some fraction which is a function of the length of the line, r(n) decide not to join and the remaining $1-r(n)$ join the line. Similarly, those already in the queue are each making the same evaluation with r(n) deciding to leave the queue and return to the population and the remaining $1-r(n)$ deciding to stay in the queue. The decision of the unit in the queue to renege or not is made on the total number in the queue, not on the individual unit's position in the queue.

$$A_1:1-\infty(M,\lambda)(r(n))A_{11},n\geq 1;(1-r(n))Q_{11},n\geq 1;D_{11},n=0/$$
$$Q_1:1-\infty(F)D_{11},d_{11}=0;\ (r(n))A_{11},n\geq 1/D_1:1(M,\mu)0.$$
$$n_\infty \qquad (15)$$

Differing Queue Disciplines and Priorities

The following group queues are concerned primarily with queue disciplines other than simple first come, first served and with systems in which priorities are granted on some basis other than arrival time to individuals arriving. The cases involving priorities are among the most complex situations found in queueing theory for several reasons. A queueing situation with priorities really means that there are several queues of different classes in front of the service mechanism and that each of these queues must be considered separately since each has different characteristics. Secondly, the basis used for granting priorities is fundamentally arbitrary and can be very difficult to describe or determine.

<u>Single Channel, Random Access to Service From the Queue:</u> This is the variation of the basic model where units in the queue enter the service mechanism on the basis of random selection rather than first come, first served.

$A_1 : 1 - \infty(M,\lambda)Q_{11}, n \geq 1; D_{11}, n = 0/Q_1 : 1 - \infty(R)D_{11}, d_{11} = 0/$
$D_1 : 1(M,\mu)0.$

n_∞	(11)	$\bar{n}_\infty$	(11)
w_∞	(11)	$\bar{w}_\infty$	(11)
s_∞	(11)	$\bar{s}_\infty$	(11)
$\bar{q}_\infty$	(11)		

Multiple Channel, Random Access to Service From the Queue: This case is exactly like the preceding one except that the unit randomly selected from the queue for service goes to a multi-channel service mechanism.

$A_1 : 1 - \infty(M,\lambda)Q_{11}, n \geq M; D_{1j}, d_{1j} = 0, n < M/Q_1 : 1 - \infty(R)D_{1j}, d_{1j} = 0/$
$D_1 : M(M,\mu)0.$

$\bar{w}_\infty$	(15)	w_∞	(15)
$\sigma_{w\infty}$	(15)		

Constant Service Times, Single Channel, Random Access: This system is the same as the single channel, random access to service from the queue with the one difference that the service times are assumed to be constant.

$A_1 : 1 - \infty(M,\lambda)Q_{11}, n \geq 1; D_{11}, n = 0/Q_1 : 1 - \infty(R)D_{11}, d_{11} = 0/$
$D_1 : 1(D,\mu)0.$

w_∞ (25), (15)

Bulk Service With Arbitrary Service Distribution: In this case, the units are served in groups. If the arriving unit finds the service mechanism empty at arrival, it enters and is served alone. If the arrival finds the unit busy, it enters the queue. When the service mechanism completes service, it takes as many as are waiting up to a given number. The service time distribution is arbitrary.

$A_1:1-\infty(M,\lambda)Q_{1i},(1+i)Q \geq n > iQ; D_{11}, d_{11}=0, n<Q/$

$Q_1:\infty - Q(B)D_{11}, Q_{1j<1i}=0, d_{11}=0/D_1:1(G)0.$

q_∞ (26), (15)

w_∞ (26), (15)

Single Channel, Two Priority, Non-Premptive: In this situation, there are two classes of units arriving: high priority units and low priority units. The high priority units have the right to service on a first come, first served basis among themselves ahead of the low priority units which may be waiting, but cannot force a low priority unit out of the service mechanism. The low priority units are served on a first come, first served basis among themselves only when there are no high priority units waiting. A fraction, α, of the arriving units are high priority and $1-\alpha$ are low priority. They are served at the same rate.

$A_1:1-\infty(M,\alpha\lambda)Q_{11}, n>1; D_{11}, n=0::A_2:1-\infty(M,(1-\alpha)\lambda)Q_{21}, -D_{11};$

$D_{11}, d_{11}=0, q_{11}=0/Q_1:1-\infty(F)D_{11}, d_{11}=0::$

$Q_2:1-\infty(F)D_{11}, d_{11}=0, q_{11}=0/D_1:1(M,\mu)0$

(11) (11)

<u>Two Priority, Non-Premptive, Differing Service Rates:</u> This is the same system as that immediately above except that the two classes of units are served at differing service rates with the class two units having a service rate which is some multiple, β, of the service rate of the class one units. $\beta > 0$.

$A_1:1-\infty(M,\alpha\lambda)Q_{11}, n \geq 1; D_{11}, d_{11}=0, d_{21}=0, q_{11}=0::$

$A_2:1-\infty(M,(1-\alpha)\lambda)Q_{21}, n \geq 1; D_{21}, n=0/Q_1:1-\infty(F)D_{11}, d_{11}=0,$

$d_{21}=0::Q_2:1-\infty(F)D_{21}, d_{11}=0, d_{21}=0, q_{11}=0/$

$D_1:1(M,\mu)0::D_2:1(M,\beta\mu)0$

(11) (11)

(11)

<u>R Priority, Non-Premptive, Differing Service Rates:</u> This is a generalization of the two priority case.

$A_1:1-\infty(M, \lambda_1)Q_{11}, n \geq 1; D_{11}, d_{1j}=0, q_{11}=0::$

$A_2:1-\infty(M, \lambda_2)Q_{21}, n \geq 1; D_{21}, d_{11}=0, d_{21}=0, q_{11}=0, q_{21}=0::$

$\ldots::A_j:1-\infty(M,\lambda_j)Q_{j1}, n \geq 1; D_{j1}, n=0/Q_1:1-\infty(F)D_{11}, d_{j1}=0::$

$Q_2:1-\infty(F)D_{21}, d_{j1}=0; q_{11}=0::\ldots::Q_j:1-\infty(F)D_{j1}, d_{j1}=0,$

$q_{k1(k<j)}=0/D_1:1(G,\mu_1)0::D_2:1(G,\mu_2)0::\ldots D_j:1(G,\mu_j)0$

s_∞ (27), (15)

w_∞ (27), (15)

R Priority, Non-Premptive, Multiple Identical Channels: This system is a version of that immediately above except that there are multiple channels which can operate simultaneously and have identical exponential service rates.

$A_1:1-\infty(M,\lambda_1)Q_{11}:n \geq M;D_{1i},d_{1i}=0,n<M::$

$A_2:1-\infty(M,\lambda_2)Q_{21},n \geq M;D_{1i},n<M,q_{11}=0::$

$\ldots::A_j:1-\infty(M,\lambda_j)Q_{j1},n \geq M;D_{1i},n<M,q_{11}=0,q_{21}=0,$

$\ldots q_{j-1,1}=0/Q_1:1-\infty(F)D_{1i},d_{1i}=0::Q_2:1-\infty(F)D_{1i},d_{1i}=0,$

$q_{11}=0::\ldots::Q_j:1-\infty(F)D_{1i},d_{1i}=0,\sum^{i=j-1} q_{i1}=0/D_1:M(M,\mu)0$

s_∞ (15)

Single Channel With a Continuous Number of Priorities: This is the basic single channel case except that units are assigned priorities along a continuous scale according to their required service time and served in that order.

$A_1:1-\infty(M,\lambda)Q_t,n \geq 1;D_{11},n=0/Q_t:\infty-\infty(F)D_{11},d_{11}=0,\sum^{i<t} q_i=0/$

$D_1:1(M,\mu)0$

(28), (15) (28), (15)

Queues From Limited Calling Populations

The following group of queueing situations is characterized by having a finite calling population rather than an infinite one as have the preceding models.

The Swedish Machine Case: This is the case where there is a limited calling population of size N requesting service from a finite number of service facilities, M. The calls from the arriving units are randomly distributed as are the service times. The queue is strictly first come, first served. When a unit's service is completed, it is returned to the calling population.

$A_1:1-N(M,\lambda)Q_{11},n \geq M;D_{1i},d_{1i}=0,n<M/Q_1:1-N-M(F)D_{1i},$
$d_{1i}=0/D_1:M(M,\mu)A_{11}$

n_∞ (13), (15)

$\bar{n}_\infty$ (13), (15)

$\bar{w}_\infty$ (13)

$\bar{q}_\infty$ (13)

The Single Repairman Case: This is the same situation as the Swedish machine case except that there is only one serving facility.

$A_1:1-N(M,\lambda)Q_{11},n>0;D_{11},n=0/Q_1:1-N-1(F)D_{11},d_{11}=0/$
$D_1:1(M,\mu)A_{11}$

n_∞ (13), (15)

$\bar{n}_\infty$ (13)

$\bar{w}_\infty$ (13)

$\bar{q}_\infty$ (13)

CHAPTER VI

PROBLEMS ENCOUNTERED IN THE CLASSIFICATION OF REAL QUEUEING SITUATIONS

The problems associated with the classification of queueing situations may be divided for convenience into four rough types which are not mutually exclusive, but rather overlap a great deal. These types of problems are: (1) the statistical problem of the determination of distributions and parameters from data collected about a given queueing situation, (2) the physical problem of measurement or the collection of the numerical data relating to the system, (3) the problems associated with transientness in the various characteristics of the system, and (4) the problem of a changing informal discipline in the system. It can be seen that these four types of problems in many respects refer to the same things, but this division seems to be useful in considering the particular approaches.

Statistical Determination of Distributions

Since most of the work which has been done in finding solutions for queueing situations depends upon a knowledge of the distributions of the characteristics[1] it follows that the ability to classify a particular situation and to find its related solution depends upon the ability to determine the relevant distributions. The determination of a distribution is an old and common problem of statistics generally. The most common approach to the problem is that which was used in this investigation of hypothesizing various distributions and testing the goodness of the fit provided by the Chi Square test. This "cut and try" approach has many obvious difficulties. The most prominent are that it greatly depends upon the experience and insight of the user for its success, it does not calculate a good fit, it does not converge toward a good fit, and it can be extremely lengthy even with high speed computing facilities and still not be successful.

A second common approach is to fit by the "least squares" method some sort of multiple regression equation in a great many variables and then determine from the regression coefficients the nature of the distribution. This approach is almost prohibitive without the use of a high speed computer and even then is fraught with difficulties and dangers, principally the fact that chance variations when fitted in this manner can completely mask out a real

[1]There are some distribution free or partially distribution free solutions available such as that of Kendall which is independent of the service time distribution, but usually these solutions are incomplete and limited.

distribution. Too, with this approach, there is again no guarantee of a fit or convergence with success depending mainly upon the experience, insight and luck of the user.

A third approach to this problem is that developed by Karl Pearson for the classification of empirical distributions which is discussed in some detail in Fry's (4) Probability Theory and Its Engineering Applications. This approach seems to have the merit of convergence for unimodal distributions, but involves such an extremely large amount of calculation that it is rarely used in practice, even with high speed computers.

Transientness

The problems in the development of solutions to queueing situations associated with transientness fall into two classes; namely, those which result from a situation with stable arrival and service distributions and disciplines in which the probabilities of the system being in the various states have time dependent components which diminish with time to the point where they are inconsiderable, and those wherein one or more of the parameters of the distributions, the distributions themselves or the discipline of the queue, are changing as a function of time.

The first type of transientness corresponds to well-known phenomena in other branches of engineering such as transient currents in electrical engineering or vibrations in mechanical engineering. These problems are well recognized in queueing literature, but are only partially solved. Most of the solutions are for the steady-state situation where sufficient time has elapsed for the transient components to reduce to an insignificant order of magnitude. This time, usually called the relaxation time, is important when its magnitude is large in relation to the period of operation of the system under consideration. From a strict theoretical standpoint, the relaxation time of the system must be known before the steady-state solution is used. In practice, this is not always followed, since the steady-state solutions are usually much easier to obtain than the time dependent solutions. In some cases, this is quite satisfactory because the general nature of the relaxation process can be determined fairly satisfactorily from observation, but in many cases, especially when rapid solutions are required by persons inexperienced in the operation of queueing systems, this practice may lead to extremely poor results.

The second type of transientness has had almost no study and is probably a much greater source of inaccuracy in queueing studies than any other single factor. Problems resulting from this type of transientness have received very little study, beyond a common admonition that the investigator should make sure that the system is stable before beginning a study or to break the times studied down into segments which approximate stability. Both of these suggestions are filled with great difficulties. The first is faced with the fact that many systems simply do not stabilize at all. An example of such a situation is a theatre queue in which the arrival rate has an almost direct relationship to the time of beginning of the attraction. The second suggestion is often made useless by the rate at which the system changes with time. Again, the theatre queue presents an example where the rate of change of the arrival distribution is too great to permit this method to be successful.

Changing Disciplines

In queueing situations involving people, one of the most troublesome problems is lack of discipline or perhaps failure to maintain a given pattern of discipline. A particular characteristic of this problem is that it is very difficult to observe changes in the discipline where people are involved and the discipline is not enforced by mechanical means. Fortunately for industrial engineers, most industrial systems use people as arrival and service mechanisms and have products as the units of the system with a mechanically enforced single discipline system.

Measurement

The problems associated with measuring the characteristics of a queueing situation are primarily a function of three characteristics, the physical scope of the system, the complexity of the system, and the magnitude of the arrival and service rates. The measurement problems are all essentially economic except one, instrumentation to detect transientness and changes in discipline.

The effect of the physical scope of the queueing system is directly one of size. In a simple, small system with moderate rates of arrival and departure, it is possible to measure the characteristics of the system by single observer with such simple timing mechanisms as the ordinary stop watch. As the physical size of the system is increased, direct observation becomes less and less effective, and such expedients as artificial vantage points, observation towers, etc., become necessary until the point where direct observation is no longer feasible. At this point, indirect observation becomes necessary. The complexity of the instrumentation for indirect observation may range from a simple trip wire or electrical contact to the most elaborate telemetry. Generally, the physical size of the system governs the expanse of the transmission equipment required. Examples of the increasing complexity of the transmission requirements in studying queues were found in connection with the queues investigated in connection with the classification problem and others. In the measurement of a queueing system involving a movie theatre line, a single observer working with a stop watch and a clip board was quite satisfactory since there was a single line with moderate arrival and service rates, tentatively classified as the type $A_1:1-\infty(U)Q_{11}, n>1;$

$D_{11}, n = 0/Q_1: 1-\infty(F)D_{11}, d_{11} = 0/D_1: 1(U)0$. Another system with a similar basic structure in a "Minute Car Wash" required that the observer using the same equipment take a position on the roof in order to see the arrivals and view the service mechanism through an opening in the roof, while studies of highway traffic queues have required the erection of scaffolding towers.

The problems of measurement presented by the increasing complexity of the system are much more extensive than those of simple size. In the study reported in Chapter VIII of the cafeteria of the type $A_1: 1-N(U)Q_{11}, n \geq 2, q_{12} > q_{11}; Q_{12}, n \geq 2, q_{11} > q_{12}; D_{i1}, d_{i1} = 0/Q_1: 2-\infty(F)D_{i1}, d_{i1} = 0/D_1: 1(U)0::D_2: 1(U)0$, it was found to be very easy to handle each queue under the assumption of independence, $A_1: 1-N(U)Q_{11}, n > 0; D_{11}, n = 0/Q_1: 1-\infty(F)D_{11}, d_{11} = 0/D_1: 1(U)0$, by a single observer with a stop watch, but that it was not possible to handle the combined system without great difficulty by two observers so equipped. Similarly, a system in a supermarket with seven service channels and lines apparently of the type $A_1: 1-N(U)Q_{1i}, n \geq 7, q_{1i} = \min.\ q_1\ ; D_{1i}, n < 7, d_{1j} = 0/Q_1: 7-\infty(F)D_{1i}, d_{1i} = 0/D_1: 7(U)0$ proved to be impossible to measure without elaborate instrumentation which was beyond the limitations imposed on these studies by budgetary restrictions.

The expansion of the complexity of the queueing system requires an expansion of the facilities for recording the inputs from the transmission system. Generally, this expansion is linear with the total number of arrival populations, queues, and departure mechanisms.

The third factor contributing to the problem of measuring the characteristics of queueing systems is the magnitude of arrival rates and departure rates. The measurement of the inter-arrival times, service times and queue states presents an order of difficulty inversely proportional to the magnitude of these quantities. In cases where the arriving units are aircraft arriving for repair and the magnitudes involved are on the order of days, the problem of measurement lies not in measuring a particular unit's times, but in collecting enough data to have a statistical significance before the system changes or in time to make a practical use of the information gained. As the time shortens, the problem shifts and becomes more difficult in two ways. First, the times become sufficiently short that some means of "magnification" must be included in the recording system such as micro-motion filming, high speed recording equipment using punched, smoked or inked paper, magnetic tape, or photographic methods of tracing an electric impulse. Secondly, as the times shorten, the problem of interpreting the data economically becomes acute due to the mass of data which must be translated from the form of the recording mechanism to a numerical form for analysis.

It must be pointed out that the three problems discussed above are not independent, but rather are interacting and that in many cases the magnitude of the difficulties presented by the interaction are of a much larger order of magnitude than those presented by the separate components.

CHAPTER VII

VARIOUS CONSIDERATIONS OF THE PARTICULAR PROBLEM OF THE RELATIONSHIP BETWEEN SERVICE RATES AND QUEUE LENGTH

In studying the problem of classifying queueing systems in the real world, the primary interest was placed on the relationship between service rates and queue length. This chapter is devoted to a brief review and examination of the principal hypotheses to be found in the literature regarding this relationship.

Erlang's Assumptions

In his development of the classic queue with exponential arrivals and service to and from a single first come, first served queue which is not limited in length or $A_1:1-\infty(M:\lambda)Q_{11},n>0;D_{11},n=0/Q_1:1-\infty(F)D_{11},d_{11}=0/D_1:1(M,\mu)0$, Erlang assumed that the service rate was independent of the queue length and, in fact, had the time of service end to be completely random. (1). In the queue, $A_1:1-\infty(M,\lambda)Q_{11},n\geq M;D_{1j},n<M,d_{1j}=0/Q_1:1-\infty(F)D_{1j},d_{1j}=0/D_1:M(M,\mu)0$, he assumed that the service rates of the individual channels were independent of the queue length, but that the total service rate of all of the channels was

directly proportional to the number in service, up to the point where the number in service is equal to the number of service channels and is independent beyond that point, thus:

μ_c = the stationary service rate of a single channel when serving

μ_n = the total service rate of the system when it contains n units

$\mu_n = n\mu_c \qquad n < m$

$\mu_n = m\mu_c \qquad n > m$

m = number of service channels available.

There is considerable logical support for the validity of these models in certain cases which follows the same general line of reasoning that supports the Poisson distribution in most other situations where the probability of the event is very small, but the number of opportunities for the event to occur is very large so that the average number of occurrences is constant.

Feller's Birth and Death Process

Feller (29) describes the solution for the case $A_1:1-\infty(M:n\lambda)Q_{11},n>0;D_{11},n=0/Q_1:1-\infty(F)D_{11},d_{11}=0/D_1:1(M,n\mu)0$, where the number in the system is allowed to become infinite and the number of service channels, each with an identical exponential service time distribution, is always equal to the number of units in the system and, thus, is allowed to become infinite also. The infinite service rates which this system indicates are rare in industrial practice and situations for which this system forms a reasonable approximation have not been reported, although it forms a useful tool in the theoretical consideration of population studies.

Romani's Variable Channels

The solution to a variation in Feller's birth and death process was developed by Romani (30) in which the number of service channels was held constant until the queue reached a given length when another channel is opened. The number of channels is allowed to become infinite. Channels are removed from service when the queue reaches zero and the unit in the channel completes service. Again, it is difficult to find industrial operations which are approximated by this type of service variation. This system under the proposed system of classification would be described by $A_1:1-\infty(M:\lambda)D_{1i}, n=i-1//D_1:i(M,\mu)0$.

Phillips' Limited Variable Channels

A more practical case was studied by Phillips (31) in which the number of service channels is limited to some finite number with the queue being allowed to become infinite after this finite number of channels is put into service. Under Phillips model, the number of channels is reduced by one whenever a channel completes the service of its unit and the queue length is zero. This case was shown to form a good approximation to the situation where a reserve of identical workers or service mechanisms is available, but does not handle a varying service rate by the service channels. This system would be classified as $A_1:1-\infty(M:\lambda)Q_{11},n \geq M, D_{1i}, n = i < M/Q_1:1-\infty(F)D_{1i}; d_{1i} = 0/$ $D_1:M(M,\mu)0$.

Conway and Maxwell's Case

The most recent hypothesis regarding the relationship between service rate and the length of the queue was described together with its solutions by Conway and Maxwell (32). This system operates for the queue $A_1:1-\infty(M,\lambda_n = \lambda n^c)Q_{11},n > 0; D_{11},n = 0/Q_1:1-\infty(F)D_{11},d_{11} = 0/$ $D_1;1(M,\mu_n = \mu n^c)0$ where the service rate of the single channel is proportional to some power of the queue length. Thus:

μ_n = the total service rate of the system when it contains n units.

μ = "average" service rate when there is only one unit in the system.

$\mu n = n^c$.

This system then can be reduced to the queue, $A_1:1-\infty(M:\lambda)Q_{11},n > 0; D_{11},n = 0/Q_1:1-\infty(F)D_{11},d_{11} = 0/$ $D_1:1(M,\mu)0$, by the selection of $c = 0$, or to the birth and death system of Feller, $A_1:1-\infty(M:n\lambda)Q_1,n > 0; D_{1i},n = 0/$ $Q_1:1-\infty(F)D_{11},d_{11} = 0/D_1:1(M,n\mu)0$, by the selection of $c = 1$. For values of $c > 0$, the situation is modeled where the units in the queue create a pressure which tends to force the units through the service mechanism at an increasing rate. For values of $c < 0$, the service mechanism is increasingly clogged by the units in the line.

There are reported in the literature tests of the validity of only three of these various assumptions. The assumption of independence in the models of the general type $A_1:1-\infty(M:\lambda)Q_{11}, n > 0; D_{1j}, n = 0/Q_1:1-\infty(F)D_{11}, d_{11} = 0/D_1:1(M,\mu)0$ has been widely tested. The first such examination was reported by Molina (3) where he found the results of a 1925 New Jersey study of 7837 local telephone calls which were found to fit Erlang's assumptions well. This assumption seems in general to fit well those cases where the service time is determined by the needs of the calling population and these are in turn determined by a possible very large number of unlikely small needs in combination. Phillips (31) reports testing his hypothesis in the reception office of a large hospital and found a good fit. In addition, there is reason to believe that the assumption of constant service times exists or is very closely approximated by many situations where there is only one service to be performed by a mechanical device.

The following chapter reports on the investigation of the relationship between the service rate and the line length in a system wherein human beings act as the service mechanism and have a degree of control over the time required to perform the service by varying their effort levels.

CHAPTER VIII

REAL QUEUES WITH LENGTH DEPENDENT SERVICE RATES

In studying the problems associated with the classification of queueing systems, it was decided that a most useful approach would be to investigate in an actual case the classification of a particular type of service distribution. This approach recommended itself for two reasons. First, it would provide an insight in the work-a-day problems faced by the engineer generally attempting to classify distributions for this purpose. Second, by choosing to investigate the particular class where the service mechanism was a human operator, it was hoped to shed a little light on certain long time problems of industrial engineering with regard to the effect of a back log of work on the operator. For these reasons, one of the principal points in the investigation of queues was made by the problem of the relationship between the length of the queue and the service rate of the system when human beings were acting as the service mechanism, when these operators were capable of observing the line length, when changes in their effort level were capable of influencing the service rate and when their reaction to the line length was uninfluenced by <u>special</u> incentives to either maintain a given length of the queue or to reduce the queue. This chapter reports in detail on the investigation of a representative system of this sort and comments on another case which bears on the problem.

Description of a Typical Situation Investigated in Detail

A queueing situation which presents a typical case of the human operator acting as the service mechanism, with a partial ability to control the service rate by varying, either consciously or unconsciously, the effort level in response to an observable change in the queue length was the operation of the cashier's booth in the cafeteria of the Student Union Building of Oklahoma State University at Stillwater, Oklahoma. This cafeteria is a large commercial type cafeteria operated by the University for the convenience of the faculty, staff, students and visitors of the University that offers a wide selection of foods at each meal. The patron enters a single queue which feeds two service lines. After being served with food, the patron is given a ticket for the price of his food at the end of the food service line after which he repairs to the tables to consume his meal. After finishing his meal, the customer enters one of two queues to the cashier's booth, each queue leading to a separate cashier, where he pays the bill for his meal and leaves the system. Observation showed that there was very little crossing from one line to the other, i.e., that once a customer had entered a given line, he usually stayed in it. It was the belief of the observer, both by personal experience and by contact with various other customers, that the patron usually chose the line which appeared to be the shortest and that by this mechanism, the lines generally maintained approximately the same length. According to the manager of the cafeteria, the operators of the service facility (the cashiers) were of approximately the same skill level and to further equalize matters were shifted frequently during the study. Even so, at the beginning of the study, there were grounds for believing that the service rate of the North Line would be significantly smaller than that of the South Line due to the South Line's service station being equipped with a modern electrically driven change machine which delivered the correct change to the customer by a chute while the North service station was equipped with a manual machine from which the change had to be handed to the customer manually. The apparent classification of this queue was $A_1{:}1-N(U)Q_{11}, n>2, q_{12}>q_{11}; Q_{12}, n>2, q_{11}>q_{12}; D_{ij}, d_{ij}=0/Q_1{:}2-\infty(F)D_{ij}, d_{ij}=0/D_1{:}1(U)0{::}D_2{:}1(U)0$.

Method of Measurement

Since preliminary observation showed that the times involved in the measurement of the queue length-service time relationship in this particular case would be well within the capabilities of an observer using an ordinary industrial engineering decimal minute stop watch, this method was selected for the study. Because the two cashier queues acted independently except for the joint variation in their arrival rates which was obviously a time dependent function of the class schedules and of the propensity of the patrons for having their meals at or near noon and six p.m., it was decided for convenience to study the queues independently. The validity of this procedure was indicated by the fact that the relationship in question, that of the service rate of a particular service facility and the queue length made the service rate depend, if at all, on its own queue. The observer was placed at the position midway between the two cashiers at the head of the lines. From this vantage point, he could observe both departures from the system and entries into it. The times of each entry and each departure were recorded on columnar data paper as the occurred. A sample data sheet is included in Appendix A. At the end of each day, the observer calculated on the data sheet, the inter-arrival interval, the service time, and the number in the system (including the occupant of the service facility) at the time of each arrival. These facts then served as the raw data for the study.

The determination of the inter-arrival time, the service time and the number in the system at the time of arrival for each arrival was made in the following manner.

The inter-arrival time for the nth arrival was simply the time of arrival of the nth arriving unit minus the arrival time of the n-1th arriving unit. The service time for the nth unit was determined by comparing the arrival time of the nth unit with the departure or service completion time of the n-1th unit. If the arrival time of the nth unit was greater than or equal to the departure time of n-1th unit, the service time of the nth unit was the arrival time of the nth unit minus the departure time of the nth unit since that unit spent all of its time in being served with no waiting. If the arrival time of the nth unit was less than the departure time of the n-1th unit, the service time of the nth unit was taken to be the departure time of nth unit minus the departure time of the n-1th unit. This assumes that the departure from the service window of a unit and its subsequent replacement occurred instantaneously which was not strictly true, but the time involved was of such a much smaller order of magnitude compared to the other times involved as to be immeasurable by the techniques employed and, thus, it was felt this assumption was a reasonable one. The number in the system at the time of the arrival of the nth unit was determined by locating the arrival time of the nth unit between the departure times of the previous units and counting back. In the very rare cases where the departure time and arrival time coincided, for this purpose, it was assumed arbitrarily that the departure had actually occurred before the arrival by a minute amount. This method of measurement appeared to be entirely satisfactory for the purpose of this study, although it has severe limitations which prevent its successful use on many systems. These limitations and other problems involved in the measuring process are discussed in the section on measurement in Chapter VI.

Determination of Stability

The first question to be determined from the data was the meal-to-meal stability of the variation of the service times. A σ control chart was prepared. This chart, shown in Figure 2, indicates that a number of points (meals 11, 17, 20, 21, and 24) were "out of control" or not members of the same system of chance variations from the average unless an extremely improbable combination of mischances had occurred in the sampling. Although it is not possible to consider this a stable system, probably because of variations in the menu offering, it was decided that the results should be analyzed in three ways for the sake of comparison. First and correctly, a day-by-day analysis of the independent results of each day. Secondly, by assuming that there existed a reasonable assignable cause, the change in the menu, for the "out of control" days and removing them. This selected data was then tested for stability of the standard deviation from day-to-day and found to be "in control". This control chart is shown in Figure 3. The means of the selected stable samples were then

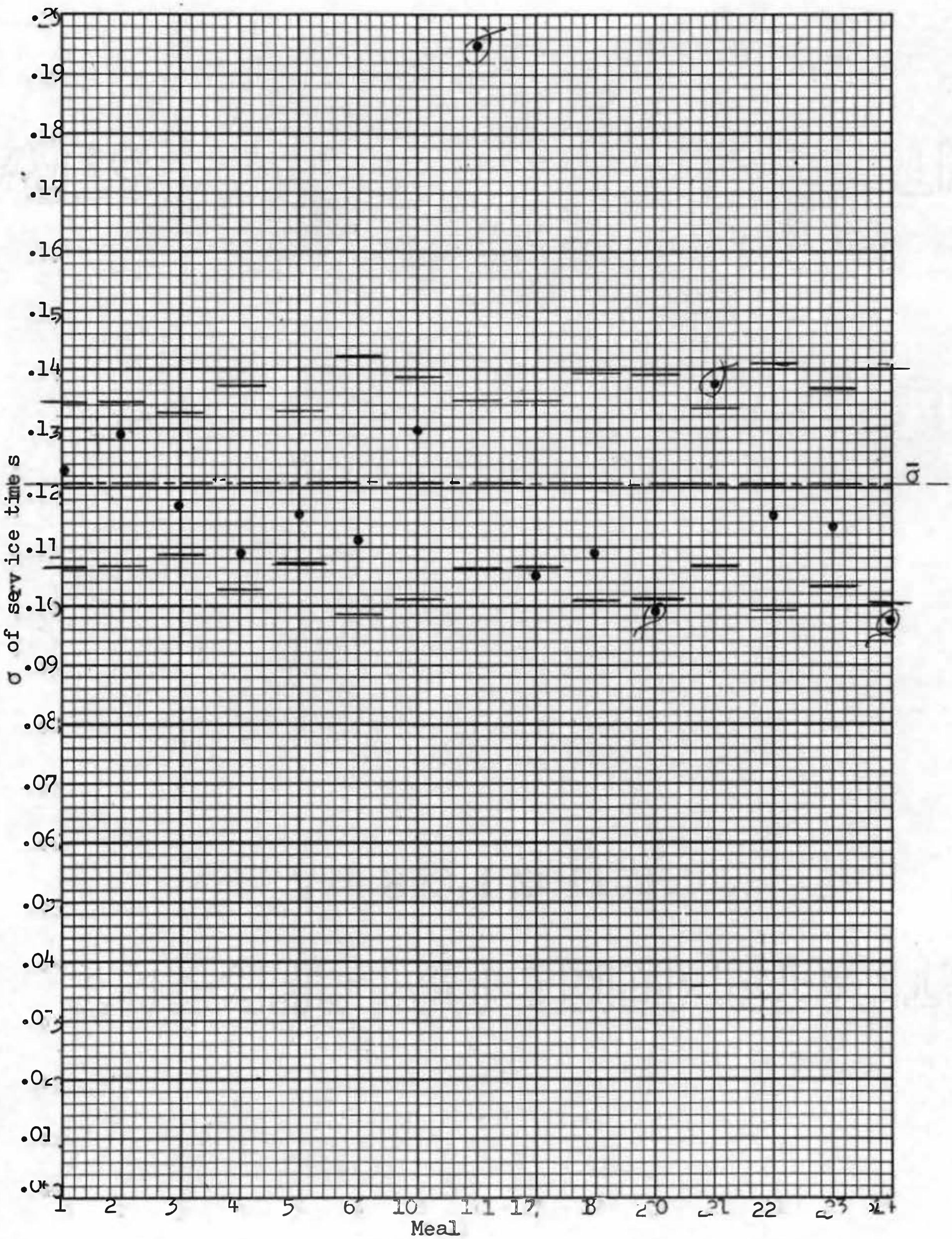

Figure 2. Service Time σ Control Chart — Total Study, Cafeteria

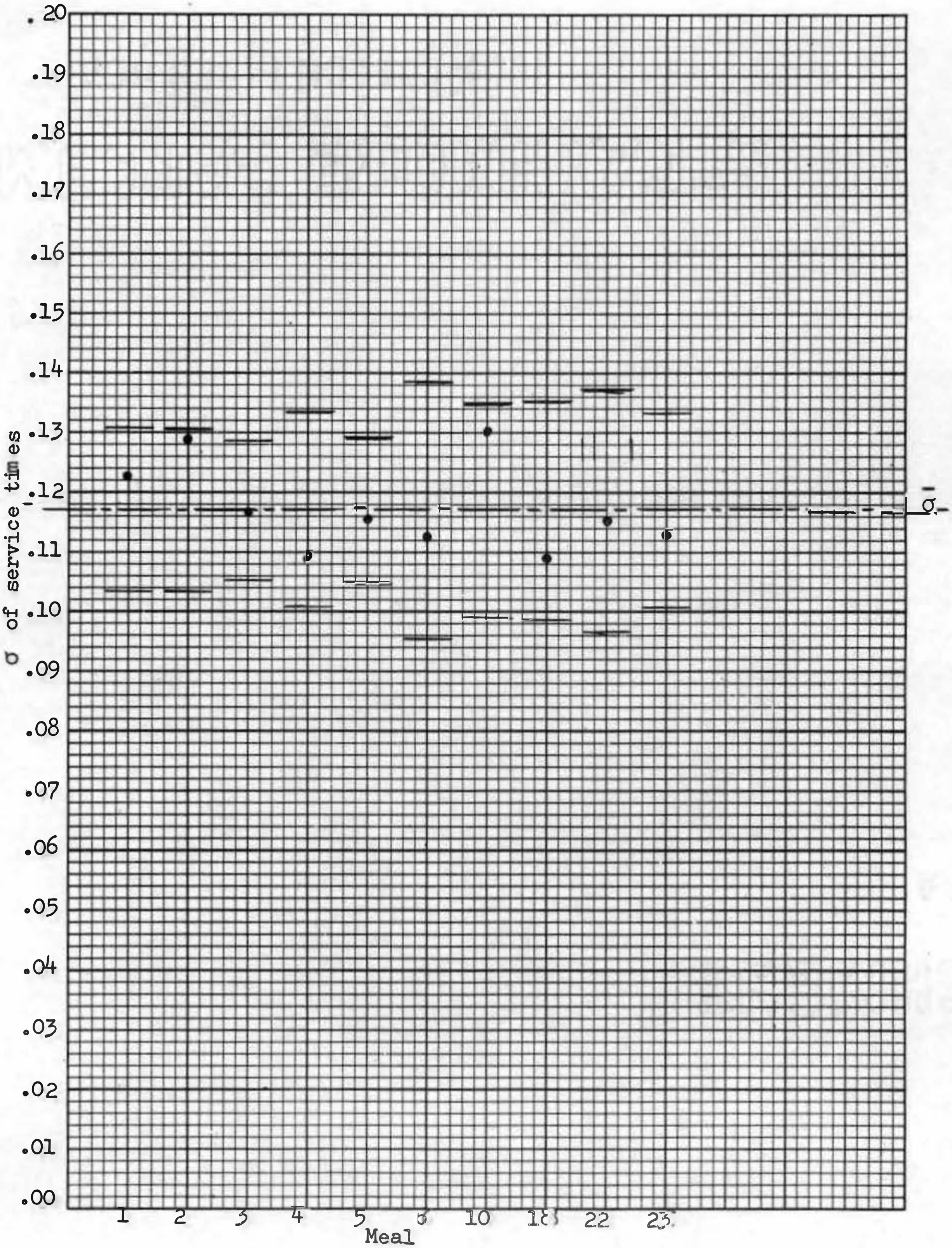

Figure 3. Service Time σ Control Stable Data, Cafeteria

formed into an $\overline{X}$ chart in Figure 4 and the means were found also to be stable from the same universe. This selected stabilized data was then grouped and examined as whole on the same basis as the individual days'. Thirdly, this lack of control was ignored to consider the complete data for the study from the standpoint of the relationship between the service time and number in line only.

Test of the Randomness Hypothesis

The first hypothesis to be tested in this system was the common one that the service times ended randomly which implies that the service times would be distributed according to the exponential distribution.[1] This hypothesis was tested first because it forms the underlying assumption of almost all of the queueing studies reported in the literature as discussed in Chapter VII. This hypothesis was checked day-by-day, for the stabilized data, and for the total study, in two ways each, by a visual check of the histograms of the data and by the "Chi Squared" test.

Each day, the total study and the stabilized data each showed roughly the same pattern when plotted as a histogram. To the eye, it seems apparent that this distribution is not exponential. Figure 5 shows the histogram of the data of a typical day with the exponential

[1]The implication of the exponential distribution from random service termination is fully discussed by Fry (4, pp. 221-223).

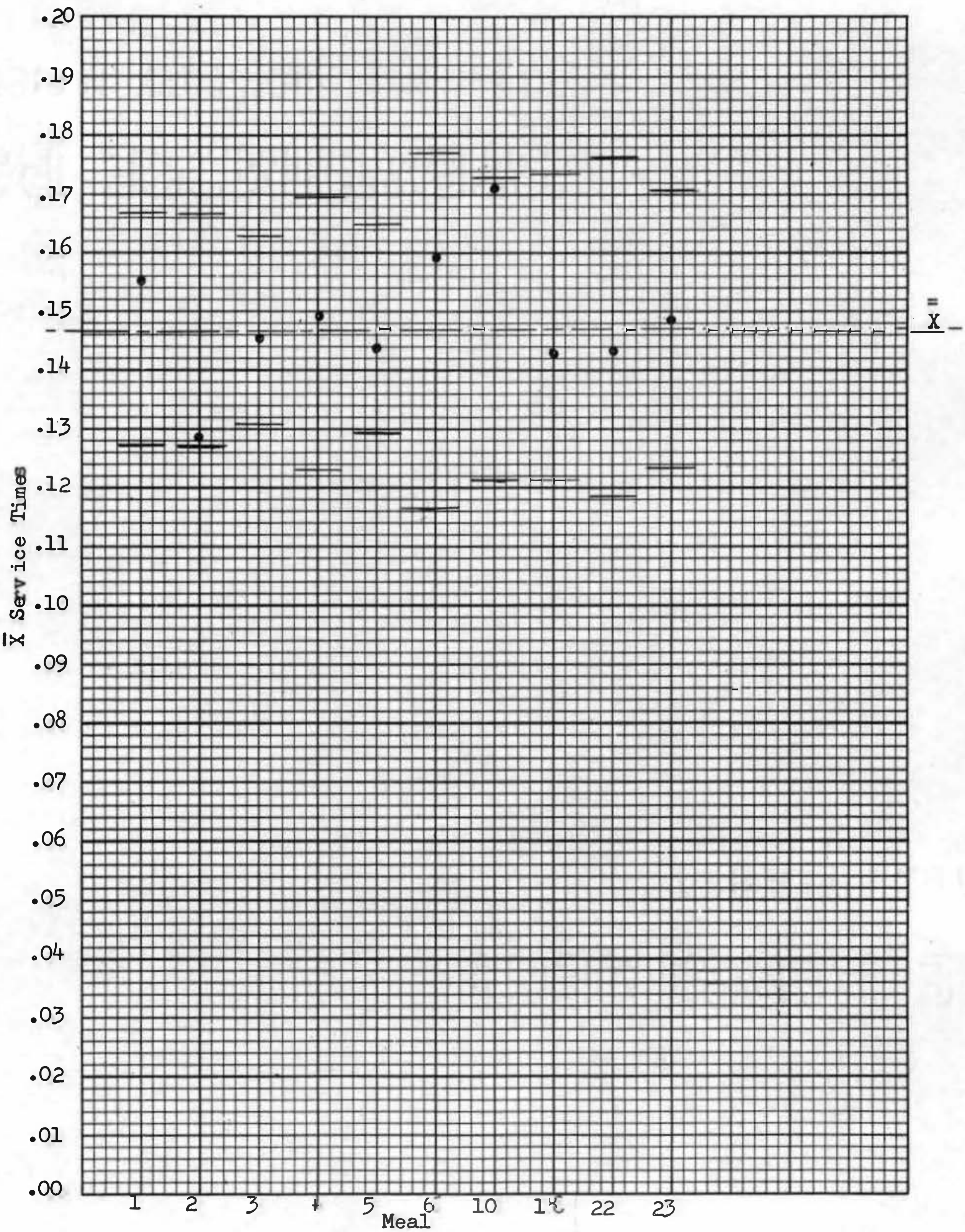

Figure 4. Service Time $\bar{X}$ Chart - Stable Data, Cafeteria

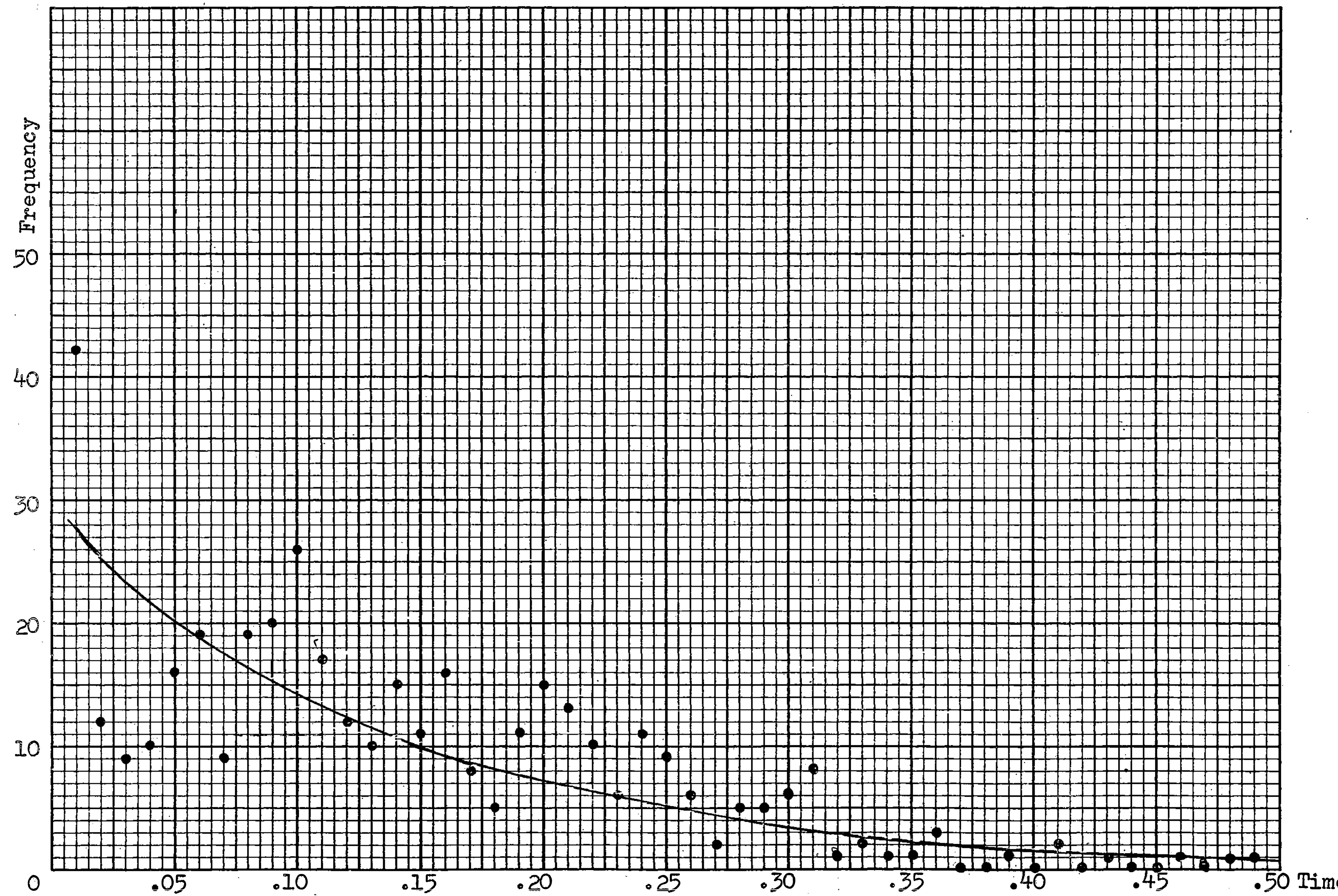

Figure 5. Service Times Histogram With Plotted Exponential for a Typical Meal (Meal 5)

distribution fitted to the same mean rate superimposed. The impression given by the histograms was confirmed by the "Chi Square goodness of fit" test. The results of the "Chi Square" tests are shown in Table II. Here, it is seen that with one exception for no one day or for the total study or for the stabilized data was the probability of getting a greater value of "Chi Square" by a random sampling from an exponential distribution with the appropriate mean service rate greater than 0.001. The one exception was meal 2 where the probability of a greater value of χ^2 was still only 0.20. This appeared to provide sufficient evidence that the service times did not conform to the random hypothesis to reject the random hypothesis and to continue the search elsewhere.

The Hypothesis of Normality

The second hypothesis to be considered for this data was that the service times were normally distributed. This hypothesis was tested by visual inspection of the histograms and by the Chi Squared test for goodness of fit. The visual inspection showed that again the "fit" was a poor one, and this was confirmed by the Chi Squared test which showed probabilities of a greater value of χ^2 for a random sample from the fitted distribution were less than 0.001 for the total study, the stabilized data and a representative day. This was taken as sufficient evidence to justify rejection of the hypothesis of normality for the service times.

TABLE II

RESULTS OF TESTS OF EXPONENTIAL DISTRIBUTION - CAFETERIA

Meal	Number Served	Avg. Service Time (Min.)	χ^2	$P(>\chi^2)$	d.f.
1	327	.15508	77.57	< .001	24
2	332	.12825	30.07	.200	24
3	463	.14481	140.59	< .001	24
4	222	.14909	73.77	< .001	22
5	403	.14389	89.26	< .001	24
6	137	.15930	62.25	< .001	17
10	188	.17140	78.11	< .001	20
11	320	.15520	98.10	< .001	24
17	323	.15180	92.37	< .001	24
18	176	.14280	48.49	< .001	20
20	183	.13830	54.18	< .001	18
21	373	.14720	113.87	< .001	24
22	148	.14720	63.38	< .001	18
23	227	.14830	53.05	< .001	21
24	160	.13310	58.99	< .001	18
Total	3,982	.14732	740.00	< .001	24
Stable	2,623	.14717	440.85	< .001	24

Test of Linear Correlation

Of particular importance to industrial engineers is the influence of the operator on the service rate in this case. This importance rests on two questions: Does the operator when not influenced by special incentives to either maintain a given length of the queue or to reduce the queue tend to increase or decrease the service rate naturally to obtain one of these ends? .. If the operators do exhibit one of these tendencies, what is the nature of this control and to what extent does it affect the performance of the operator? A tendency to maintain the queue length might be exploited by formal work groups such as unions or by informal work groups in the effort to "save the work" or "make the job last" while a tendency to reduce the line might be exploited by the industrial engineer in the design of incentive systems, the design of production control systems and their operation.

This hypothesis was tested by determining for each day for the total study, and for the stabilized data the linear regression coefficients, the coefficient of correlation, and the probability that the correlation is real. These values are tabulated in Table III. Here it is seen that generally there is a negative regression of between five and ten per cent of the average service rate for each additional unit in the system. The linear correlation between the service rate and the number in the system was

TABLE III

REGRESSION AND CORRELATION ANALYSIS - CAFETERIA

Meal	a_{12}	b_{12}	$\hat{r}_{12}^{2}$	$p(H_o : r_{12}^{1} = 0)$	t	Sig.
1	0.1737	-.007170	.008600	< .001	1.9400	V.H.S.
2	0.1619	-.014740	.024880	< .010 > .001	3.2900	H.S.
3	0.1670	-.007870	.019790	< .001	2.8790	V.H.S.
4	0.1846	-.016700	.023780	< .001	2.5270	V.H.S.
5	0.1736	-.009550	.018620	> .010 < .001	2.9360	H.S.
6	0.1428	.008140	.000630	> .100 < .200	1.0560	
10	0.1965	-.014540	.005940	> .050 < .100	1.4540	H.S.
11	0.1412	-.005130	-.021900	> .300 < .400	0.9580	
17	0.1499	.000786	-.002664	> .700 < .800	0.3832	
18	0.1935	-.024820	.071200	< .001	3.862	V.H.S.
20	0.1850	-.021780	.023980	> .010 < .020	2.340	S.
21	0.1724	-.008050	.012900	> .010 < .020	2.422	S.
22	0.1777	-.014880	.020150	> .020 < .050	2.021	S.
23	0.1750	-.010300	.021700	> .010 < .020	2.448	S.
24	0.1106	.012600	.010600	> .100 < .200	1.626	
Total	0.1656	-.007230	.012060	> .020 < .050	2.227	S.
Stable	0.1742	-.010890	.021410	< .001	7.620	V.H.S.

generally on the order of one or two per cent and in every case in the stable system except one was significant at either the 0.05 level, the 0.01 level or the 0.001 level. In the one exception, the probability of a greater value of t was between 0.2 and 0.1. For the data of the study as a whole, the regression equation for the system was:

(1) Service time = 0.1656 - 0.00723 (number in the system) with a correlation of 1.2% which was significant at the 0.05 level. For the data during the stable portion of the study, only the regression equation for the system was:

(2) Service time = 0.1742 - 0.0189 (number in the system) with a correlation of 2.14% which was significant at the 0.001 level.

It seems to be a point worth noting that the points outside of the stable system contained three points out of five for which the regression and/or correlation was not significant. Similarly, it is worth noting that three out of the four points in the whole study in which the correlation and/or regression were not significant had very small positive regressions.

The inference then is drawn that in this situation, the worker apparently has a very real ability to judge the length of the line and do, either consciously or unconsciously, increase their service rate as the line length increases. It would seem from the correlation coefficients that the model of linear regression accounts for only a small part of the variation in the individual service times. This would then seem to support the hypothesis that the operator does influence the average service rate in a basically random situation. While this information is of very little use in predicting an individual service time, it is most useful in the consideration of the system. This is particularly important in two aspects, in that the natural tendency, if this hypothesis is accepted, is for the worker to act to stabilize the line length both by some shortening of the average number in the system and by a reduction in the variation of the mean number in the system. The consequences of these effects will be considered in detail in the conclusions.

Another Situation

Another case, a "Minute Car Wash", presented a considerably dissimilar situation with some similar results. The car wash had one line with perfect first come, first served discipline with only one service mechanism, but that service mechanism was composed of six operators working as a team or apparently would be classed $A_1:1-\infty(U)Q_{11}, n > D, D_{11}, n = 0/Q_1:1-\infty(F)D_{11}, n = 0/D_1:1(U)0$. Here, the same tests were applied to the data and it was found that the mean service rate was extremely unstable from day-to-day. The primary cause of this unstability was believed to be the fact that the car wash used "pick-up" labor which varied tremendously in ability and energy from man to man. Here the Chi Square test for the goodness of fit of the exponential distribution showed on each day a probability of less than 0.001 of getting as great a difference from an exponentially distributed universe with the same mean. Table IV presents the summary of the data from this study. Using the Chi Square test for fit of the service times to the normal distribution showed that on each day the reduction in the Chi Square value by using a random sample from a normal universe had a probability for each day between 0.10 and 0.20 which provides no evidence to reject the hypothesis that the service times in this situation were distinctly different from normal, but at the same time it was found that on the first day the linear regression and correlation were significant on the 0.05 level and on the second day had a probability of approximately 0.2 of getting this great a value of t where no real correlation existed. While this evidence is far from conclusive, it would seem to indicate that the fundamental distribution is normal with the number in the line having some affect on the mean rate.

Summary

The investigations reported in this chapter seem to indicate that the worker does have some tendency to increase his service rate when the line increases, but that this is not usually the primary relationship. In the cafeteria study, it was felt that, based on the

TABLE IV

SUMMARY OF CAR WASH DATA

Test of Exponentiality

Day	χ^2	P($> \chi^2$ from Exp.)	$\bar{X}$
1	172.794	< 0.001	7.7582
2	178.938	< 0.001	9.3428

Test of Normality

Day	χ^2	P($> \chi^2$ from Normal)
1	6.851	$0.20 > P > 0.10$
2	7.325	$0.20 > P > 0.10$

Test of Linear Regression

Day	a_{12}	b_{12}	r^2	t	Prob(>t)
1	8.5953	-0.1266	0.07411	2.2207	< 0.05
2	10.0475	-0.0835	0.01186	1.2926	$\cong 0.20$

observation, the service time might be strongly influenced by the pattern of price distributions and the distributions of the persons carrying exact change and those carrying only large bills which required changing.

CHAPTER IX

CONCLUSIONS

The study covered in the prior chapters was successful in several aspects and unsuccessful in others. Possibly the most important unexpected development was the exploration of numerous problem areas which are important to industrial engineers. The most important of these areas and recommendations for their investigation are presented in the next chapter. The principal conclusions of this study regarding the classification of queueing systems are presented here.

Classification System

The classification system developed and described in Chapter IV was tested by classifying a number of representative systems considered in the literature and several real world cases without meeting a situation which it could not handle. While this is not evidence that it will be able to handle all cases presented to it, it does seem to indicate that the system is widely applicable. Further, it was found to offer definite advantages in compactness and completeness. Particularly, it is important as means of recognizing similar systems to which similar analysis is likely to prove to be profitable and, thus, extend the value of existing solutions for systems which have already been successfully analyzed. The system developed should prove with continued experience to provide a material contribution to working engineers in handling queueing problems and to students beginning the study of the subject in organizing the existing knowledge and fitting in new ideas as they are mastered.

Service Rates

In every situation tested where the service mechanisms of the queueing system were human beings whose pace of activity governed the service rate and who were in a position to observe the length of the queues backing up from their service facility, it was found that there was a statistically real influence of the line length on the service rate with greater line lengths tending to be associated with faster service. The correlation provided by the linear model was usually found to be extremely small, on the order of 0.01.

The conclusions reached from this portion of the investigation were that while the effect of line length in the cases studied was real, it was so small as to be negligible in day-to-day work, but of significance in the investigation and development of incentive systems for such workers and the planning of such systems including incentives. Even relationships of the magnitude found in this study would indicate two facts of considerable utility. First, since this tendency is real and since it is almost always an advantageous one in that it stabilizes the system and increases the average rate of work, it should be reinforced by incentives where it is possible to do so. Secondly, the economic effects of this incentive should be considered in choosing the economic average line length to be allowed as the effect is most apparent in the economically most unfavorable situation where the line length is longest.

would not economic incentives in this direction be expected to _increase_ the correlation???

CHAPTER X

RECOMMENDATIONS FOR FURTHER INVESTIGATIONS

The problem areas encountered in this study which recommend themselves and likely prospects for future investigations are concerned with three major areas: the problems of identifying or classifying the statistical distributions underlying the arrival and service times, the problems of measurement and the management problem 2 of designing incentive systems to take a maximum advantage of the characteristics of the implications of the ability 3 of human operators to control their work rate.

Mathematical and Statistical Problems

The first of these problems falls primarily in the areas of mathematical and statistical research, but the location of the required progress and developments within these areas and their application to the problems of industrial engineering is the responsibility of the industrial engineer. The two areas of development have almost immediate application to queueing problems.

Development of Measurement Techniques

The problems associated with the development of better methods of measuring the characteristics of queueing systems are primarily industrial engineering problems. In many ways, these problems are associated with the traditional reliance of the industrial engineer on certain instruments and techniques such as the stop watch, the motion picture camera micro-motion study, and charting methods in measuring work situations. The extension of the analytical tools, such as queueing theory, which are available to the industrial engineer calls for the development of new measurement techniques which take greater advantage of the sophisticated instrumentation which is becoming available through technological progress in other areas of engineering.

Effects of Incentives

Finally, this study would seem to point the need for more intensive quantitative research on the effects of incentives on both a very short term basis and over the longer run. This problem has long been one of interest to industrial engineers, but now seems to offer new fruitfulness of opportunity in the light of the current state of technological progress in instrumentation and analytical tools for the analysis of such situations.

A SELECTED BIBLIOGRAPHY

(1) Johannsen, F. W. "Waiting Times and Number of Calls," Post Office Electrical Engineers Journal, 1907.

(2) Brockmeyer, E., H. L. Halstrom, and Arne Jensen. The Life and Works of A. K. Erlang (All of Erlang's papers translated into English.). Copenhagen: Copenhagen Telephone Company, 1948.

(3) Molina, Edward C. "Application of the Theory of Probability to Telephone-Trunking Problems," Bell System Technical Journal, Vol. 6, July, 1927, pp. 461-494.

(4) Fry, Thornton C. Probability and Its Engineering Uses. Princeton, New Jersey: D. Van Nostrand Company, Inc., 1928.

(5) Pollaczek, Felix. "Uber eine Aufgabe der Wahrscheinlichkeitstheorie," Mathematische Zeitschrift, Vol. 32, June, 1930, pp. 64-100; December, 1930, pp. 729-50 (Reported in English by Kendall).

(6) Khintchine, A. "Mathematisches uber die Erwartung vor einemoffentlichen Schalter," Mat. Sbornik, Vol. 39, 1932, pp. 73-84 (Russian; German summary, results available in English by Saaty).

(7) Pollaczek, Felix. "Uber das Warteproblem," Mathematische Zeitschrift, Vol. 38, May, 1934, pp. 492-537 (Reported in English by Kendall).

(8) Bernstein, Phillip. "How Many Automatics Should a Man Run?" Factory Management and Maintenance, Vol. 99, March, 1941, pp. 85-164.

(9) Kronig, R. "On Time Losses in Machinery Undergoing Interruptions," (in Dutch) Physica, Vol. 10, No. 4, April, 1943, pp. 215-224. Part II, Ibid, Vol. 10, No. 5, May, 1943, pp. 331-336.

(10) Poage, Scott T. Management Training Course Queueing Bibliography, American Airlines, Inc., Tulsa, Oklahoma, (privately published for company use), 1960.

(11) Morse, Phillip M. Queues, Inventory and Maintenance. New York: John Wiley and Sons, Inc., 1958.

(12) Kendall, David G. "Stochastic Processes Occurring in the Theory of Queues and Their Analysis by the Method of the Imbedded Markov Chain," Annals of Mathematical Statistics, Vol. 24, 1953, pp. 338-354.

(13) Palm, Conny. "The Distribution of Repairmen in Servicing Automatic Machines," Industritidningen Norden, Vol. 75, 1947, pp. 75-80, 90-94, 119-123, (Swedish, a translation is given in the Journal of Industrial Engineering, Vol. I, No. 1, Jan.-Feb. 1958, pp. 28-41.

(14) Thuesen, H. G. "The Application of Job Design to Petroleum Production Operations," Drilling and Production Practice, May, 1944.

(15) Saaty, Thomas L. Elements of Queueing Theory. New York: McGraw-Hill Book Company, 1961.

(16) Galiher, H. P. "Queueing Theory," Notes on Operations Research 1959. The Technology Press, Massachusetts Institute of Technology, Cambridge 39, Massachusetts.

(17) Moore, James M. "To Queue or Not to Queue," Journal of Industrial Engineering, Vol. XII, No. 2, March-April, 1961, pp. 119-121.

(18) Jackson, James R. "Newworks of Waiting Lines," Journal of Operations Research Society of America, Vol. 5, No. 4, August, 1957, p. 518.

(19) Clarke, Allen Bruce. The Time Dependent Waiting Line Problem, Engineering Research Institute, University of Michigan, Report Number R-39, March, 1953.

(20) Jackson, R. R. P. "Random Queueing Process With Phase Type Service," Journal of the Royal Statistical Society, Ser. B, Vol. 18, No. 1, 1956, pp. 129-132.

(21) Lindley, D. V. "The Theory of Queues With a Single Server," *Proceedings of the Cambridge Philosophical Society,* Vol. 48, 1952, pp. 277-289.

(22) Crommelin, C. D. "Delay Probability Formulae When the Holding Times are Constant," *Post Office Electrical Engineers' Journal, Vol. 25,* Pt. 1, April, 1932, pp. 41-50.

(23) Crommelin, C. D. "Delay Probability Formulae," *Post Office Electrical Engineers' Journal,* Vol. 26, Pt. 4, January, 1934, pp. 266-274.

(24) Reuter, G. E. H. "Denumerable Markov Processes and the Associated Contraction Semigroups on 1," *Acta Math.,* Vol. 97-98, 1957, pp. 1-46.

(25) Burke, P. J. "Equilibrium Delay Distribution for One Channel With Constant Holding Time, Poisson Input and Random Service," *Bell System Technical Journal,* July, 1959, pp. 1021-1031.

(26) Bailey, Norman T. J. "On Queueing Process With Bulk Service," *Journal of the Royal Statistical Society,* Ser. B, Vol. 16, No. 1, 1954, pp. 80-87.

(27) Cobham, Alan. "Priority Assignment in Waiting Line Problems," *Journal of Operations Research Society of America,* Vol. 2, No. 1, February, 1954, pp. 70-76. "A Correction," Ibid, November, 1955, p. 547.

(28) Phipps, Thomas E., Jr. "Machine Repair as a Priority Waiting Line Problem," (with comments by W. R. Van Voorhis), *Operations Research,* Vol. 4, No. 1, February, 1956, pp. 76-86.

(29) Feller, W. *An Introduction to Probability Theory and Its Applications.* New York: John Wiley and Sons, Inc., 1950.

(30) Romani, Jose. "Un Modelo de la Teoría de Colas con Número Variable de Canales," *Trabajos Estadistica,* Vol. 8, No. 3, 1957, pp. 175-189.

(31) Phillips, Cecil R., Jr. "A Variable-Channel Queueing Model With a Limited Number of Channels (unpub. Thesis, Georgia Institute of Technology, 1960).

(32) Conway, R. W., and W. L. Maxwell. "A Queueing Model With State Dependent Service Rate," The Journal of Industrial Engineering, Vol. XII, No. 2, March-April, 1961, pp. 132-136.

(33) Duncan, A. J. Quality Control and Industrial Statistics, Rev. Ed., Homewood, Illinois: Richard D. Irwin Co., 1959.

APPENDIX A

SAMPLE DATA SHEET

7-17-61
Evening Meal
5:30 P. M.

Departure	Arrival	No. in Line	Arrival	Service
5:30:39	5:30:09	1		:30
1:04	:95	1	:86	:09
2:58	2:38	1	:43	:20
:76	:39	2	:01	:18
:91	:41	3	:02	:15
3:26	3:15	1	:64	:11
:39	:27	1	:12	:12
:50	:29	2	:02	:11
:54	:32	3	:03	:04
:90	:72	1	:40	:18
4:07	:74	2	:02	:17
:20	:78	3	:04	:13
:68	4:56	1	:78	:12
8:22	8:07	1	3:51	:15
:40	:13	2	:06	:18
:75	:50	1	:37	:25
:93	:85	1	:35	:08
9:14	:88	2	:03	:21
:16	:90	3	:02	:02
:51	9:45	1	:55	:06
:78	:57	1	:12	:21
40:53	40:25	1	:68	:28
:62	:30	2	:05	:09
:68	:50	3	:20	:06
:78	:64	2	:14	:10
:88	:80	1	:16	:08
:89	:83	2	:03	:09
1:07	:84	3	:01	:18
:53	1:41	1	:57	:12
2:46	2:35	1	:94	:11
3:81	3:72	1	1:37	:09
:94	:78	2	:06	:07
4:45	4:26	1	:48	:19
:56	:28	2	:02	:11
:64	:54	2	:26	:08
:70	:55	3	:01	:06
5:28	:80	1	:25	:48
:49	5:32	1	:52	:17

APPENDIX B

SAMPLE CALCULATIONS

STABILIZED DATA σ CONTROL CHART CALCULATIONS - CAFETERIA

Day	Σx_1	N	$\sigma^2 = \Sigma x_1/N$	σ	$\sqrt{2N}$	$\frac{3}{\sqrt{2N}}$	$B_3 = 1 + \frac{3}{\sqrt{2N}}$	$B_4 = 1 - \frac{3}{\sqrt{2N}}$	$B_3\bar{\sigma}$ = U.C.L.	$B_4\bar{\sigma}$ = L.C.L.	Control?
1	4.89600	327	.01497	.12235	25.573	.11731	1.11731	0.88269	.13083	.10335	Yes
2	5.49440	322	.01654	.12860	25.768	.11642	1.11642	0.88358	.13072	.10346	Yes
3	6.30162	463	.01361	.11666	30.430	.09858	1.09858	0.90142	.12863	.10555	Yes
4	2.63640	222	.01187	.10895	21.071	.14237	1.14237	0.85763	.13376	.10042	Yes
5	5.39125	403	.01337	.11563	28.390	.10567	1.10567	0.89433	.12946	.10472	Yes
6	1.69190	137	.01234	.11109	16.553	.18123	1.18123	0.81877	.13831	.09587	Yes
10	3.18080	188	.01691	.13000	19.391	.15471	1.15471	0.84529	.13520	.09898	Yes
18	2.07220	176	.01177	.10849	18.762	.15989	1.15989	0.84011	.13581	.09837	Yes
22	1.96900	148	.01330	.11533	17.205	.17436	1.17436	0.82564	.13751	.09667	Yes
23	2.94260	227	.01296	.11384	21.307	.14079	1.14079	0.85920	.13357	.10060	Yes
Σ				1.17094							

$\bar{\sigma} = \frac{\Sigma\sigma_i}{N} = \frac{1.17094}{10} = 0.117094$ U.C.L. $= B_4\bar{\sigma}$ L.C.L. $= B_3\bar{\sigma}$

(Factors taken from Duncan (33, p. 886.)

STABILIZED DATA $\bar{X}$ CONTROL CHART CALCULATIONS - CAFETERIA

$\bar{\bar{X}} = 0.14717$

$\hat{\sigma} = 0.12471$

Day	N	$\frac{1}{\sqrt{N}}$	$\frac{1}{\sqrt{N}}\sigma = \sigma_{\bar{x}}$	$\bar{\bar{X}} + 3\sigma_{\bar{x}}$	$\bar{\bar{X}} - 3\sigma_{\bar{x}}$	$\bar{X}$	Control?
1	327	.05530	.006475	.16661	.12773	.15508	Yes
2	332	.05488	.006426	.16646	.12788	.12825	Yes
3	463	.04647	.005441	.16349	.13085	.14481	Yes
4	222	.06711	.007858	.16985	.12359	.14909	Yes
5	403	.04981	.005832	.16466	.12968	.14389	Yes
6	137	.08543	.010000	.17717	.11717	.15930	Yes
10	188	.07293	.008539	.17279	.12155	.17140	Yes
18	176	.07537	.008825	.17366	.12068	.14280	Yes
22	148	.08219	.009624	.17603	.11831	.14270	Yes
23	227	.06637	.007771	.17048	.12386	.1483	Yes

7-16-61
Service Times
Noon

Meal 2
332
Cafeteria Study

Regression and Correlation Analysis

$N = 332$

$\Sigma X_1 = 42.58$ $\qquad \Sigma X_1^2 = 10.9552$ $\qquad 2\Sigma X_1 X_2 = 173.42$

$\Sigma X_2 = 757$ $\qquad \Sigma X_2^2 = 2429$ $\qquad \Sigma X_1 X_2 = 86.71$

$$\bar{X}_1 = \frac{\Sigma X_1}{N} = 0.12825$$

$$\bar{X}_2 = \frac{\Sigma X_2}{N} = 2.28$$

$$\Sigma_{x_1 x_2} = \Sigma X_1 X_2 - N\bar{X}_1\bar{X}_2 = -10.3702$$

$$\Sigma x_2^2 = \Sigma X_2^2 - N\bar{X}_2^2 = 703.1312$$

$$\Sigma x_1^2 = \Sigma X_1^2 - N\bar{X}_1^2 = 5.49444325$$

$$b_{12} = \frac{\Sigma x_1 x_2}{\Sigma x_2^2} = \frac{-10.3702}{703.1312} = -0.01474$$

$$a_{12} = \bar{X}_1 - b_{12}\bar{X}_2 = 0.1618572$$

$$X_{1r} = 0.1618572 - 0.01474\, X_2$$

$$\Sigma v_{1.2}^2 = \Sigma x_1^2 - b\,\Sigma x_1 x_2 = 5.341583$$

$$s_{1.2}^2 = \frac{\Sigma v_{1.2}^2}{N-2} = \frac{5.341583}{330} = 0.01618$$

$$\hat{r}_{12}^2 = 1 - \frac{(N-1)\ \Sigma v_{1.2}^2}{(N-2)\ \Sigma x_1^2} = 1 - \frac{(331)(5.34158)}{(330)(5.49444)} = 0.02488.$$

Testing $H: v_{12}^{1} = 0$:

$$t = \frac{b_{12} - 0}{s_{b_{12}}} = \frac{-0.01474}{0.004797} = 3.07275$$

$$s_{b_{12}}^{2} = \frac{s_{1.2}^{2}}{\Sigma x_2^{2}} = \frac{0.01618}{703.1312} = 0.0000230113$$

$$s_{b_{12}} = 0.004797$$

$t_{0.01} = 2.576$ ($df = \infty$) $\qquad t_{0.001} = 3.291$ ($df = \infty$)

$\therefore$ it is concluded that this small correlation is highly significant.

χ^2 Test for Exponential Distribution - Cafeteria Meal 3

$\bar{X} = 0.14481$ $\mu = 6.9056$ $N = 463$

Class	f	μt	$e^{-\mu t}$	p(n)	np(n)	$\frac{[f - np(n)]^2}{np(n)}$
0	-		1.0000			
1	27	.0691	.9324	.1306	60.468	18.52398
2	15	.1380	.8694	.0588	27.224	5.48878
3	11	.2070	.8106	.0548	25.372	8.14103
4	14	.2760	.7558	.0511	23.659	3.94337
5	21	.3450	.7047	.0411	19.029	.20415
6	18	.4140	.6637	.0449	20.789	.37416
7	11	.4830	.6188	.0418	19.353	3.60528
8	20	.5520	.5770	.0391	18.103	.19878
9	22	.6220	.5379	.0363	16.807	1.60452
10	42	.6910	.5016	.0339	15.695	44.08132
11	21	.7600	.4677	.0317	14.677	2.72401
12	22	.8290	.4360	.0294	13.612	5.16886
13	15	.8980	.4066	.0275	12.733	.40361
14	20	.9670	.3790	.0256	11.853	6.59973
15	22	1.0360	.3535	.0239	11.066	10.80357
16	14	1.1050	.3296	.0192	8.890	2.93724
17	10	1.1740	.3104	.0210	9.723	.00789
18	14	1.2430	.2894	.0196	9.075	2.67279
19	14	1.3120	.2698	.0182	8.427	3.68557
20-	50	1.3810	.2516	.0743	34.401	7.07330
25-	28	1.7260	.1773	.0511	23.659	.79649
30-	11	2.0720	.1262	.0373	17.270	2.27636
35-	5	2.4170	.0889	.0256	11.853	4.96217
40-	6	2.7620	.0633	.0187	8.658	.81600
45-	3	3.108	.0446	.0129	5.973	1.47978
50-	7	3.453	.0317	.0317	14.677	4.01555

$df = N - 2 = 26 - 2 = 24$ $\chi^2 = 140.58823$

$p(>\chi^2) < 0.001$

Car Wash
Day 1

TEST OF NORMALITY

$\bar{X} = 7.7582$

$\sigma = 1.2024$

t	fo		Z	$\Sigma\, p(>n)$	p(n)	np(n)	group	$\frac{[f_o - np(n)]^2}{np(n)}$
-0.0			-6.4523	.5000	.0000	0.000		
0.5			-6.0364	.5000	.0000	0.000		
1.0			-5.6205	.5000	.0000	0.000		
1.5			-5.2047	.5000	.0000	0.000		
2.0			-4.7889	.5000	.0000	0.000		
2.5			-4.3731	.5000	.0002	0.000		
3.0			-3.9572	.5000	.0007	0.010	7.830	2.9794
3.5			-3.5414	.4998	.0015	0.035		
4.0			-3.1256	.4991	.0176	0.075		
4.5	1		-2.7097	.4966	.0190	0.880		
5.0		3	-2.2939	.4890	.0421	0.905		
5.5			-1.8780	.4700	.0755	2.105		
6.0	2		-1.4622	.4279	.1167	3.775		
6.5	9		-1.0464	.3524	.1505	5.835	5.835	1.7167
7.0	11		-0.6305	.2357	.1646	7.525	7.525	1.6047
7.5	9		-0.2147	.0852	.1523	8.230	8.230	0.07204
8.0	7		0.2011	.0794	.1172	7.615	7.615	0.04966
8.5	5		0.6164	.2317	.0776	5.860	5.860	0.12621
9.0	3		1.0327	.3489	.0421	3.880		
9.5	2		1.4486	.4265	.0201	2.105		
10.0			1.8644	.4686	.0077	1.005		
10.5			2.2802	.4887	.0017	0.385		
11.0			2.6961	.4964	.0007	0.085		
11.5		6	3.1119	.4991	.0002	0.035	7.505	0.30180
12.0	1		3.5277	.4998	.0000	0.010		
12.5			3.9436	.5000	.0000	0.000		
13.0			4.3594	.5000	.0000	0.000		
13.5			4.7752	.5000	.0000	0.000		
14.0			5.1911	.5000	.0000	0.000		
14.5-			5.6069	.5000	.0000	0.000		

$\chi^2 = 6.8505$

$df = N - 3 = 7 - 3 = 4$

$p(>\chi^2) = 0.10 \qquad \chi^2 = 7.78$

$p(>\chi^2) = 0.20 \qquad \chi^2 = 5.99$

VITA

Scott Tabor Poage

Candidate for the Degree of

Doctor of Philosophy

Thesis: A STUDY OF THE CLASSIFICATION OF QUEUES FROM THE ENGINEERING STANDPOINT, WITH PARTICULAR CONSIDERATION TO THE RELATIONSHIP OF SERVICE RATES TO QUEUE LENGTH

Major Field: Industrial Engineering and Management

Biographical:

Personal Data: Born in Waco, Texas, December 5, 1931, the son of Scott Allen and Robbie Lee Tabor Poage.

Education: Attended grade school and junior high school in Waco, Texas; graduated from Waco High School in 1949; received the Bachelor of Science degree from Texas Technological College, with a major in Industrial Engineering, in May, 1953; received the Master of Science degree from the Agricultural and Mechanical College of Texas, with a major in Industrial Engineering, in January, 1957; completed the requirements for the Doctor of Philosophy degree in August, 1962.

Professional experience: Served as assistant plant industrial engineer at the Phillips Petroleum Company's McGregor Rocket Fuels Plant from June, 1953 to November, 1953; was a production control officer with the United States Air Force Air Materiel Command from 1953 to 1955; was a teaching fellow in Mathematics at the Agricultural and Mechanical College of Texas during 1956; was an instructor in the Department of Industrial Engineering at the Agricultural and Mechanical College of Texas from 1957 to 1959; was an assistant professor in the School of Industrial Engineering and Management at Oklahoma State University from 1959 to 1961; since September, 1961 has been head of the Department of Industrial Engineering at Arlington State College; during the period of 1957 to the present, has acted as a consultant in the areas of industrial engineering, production control, and operations research.