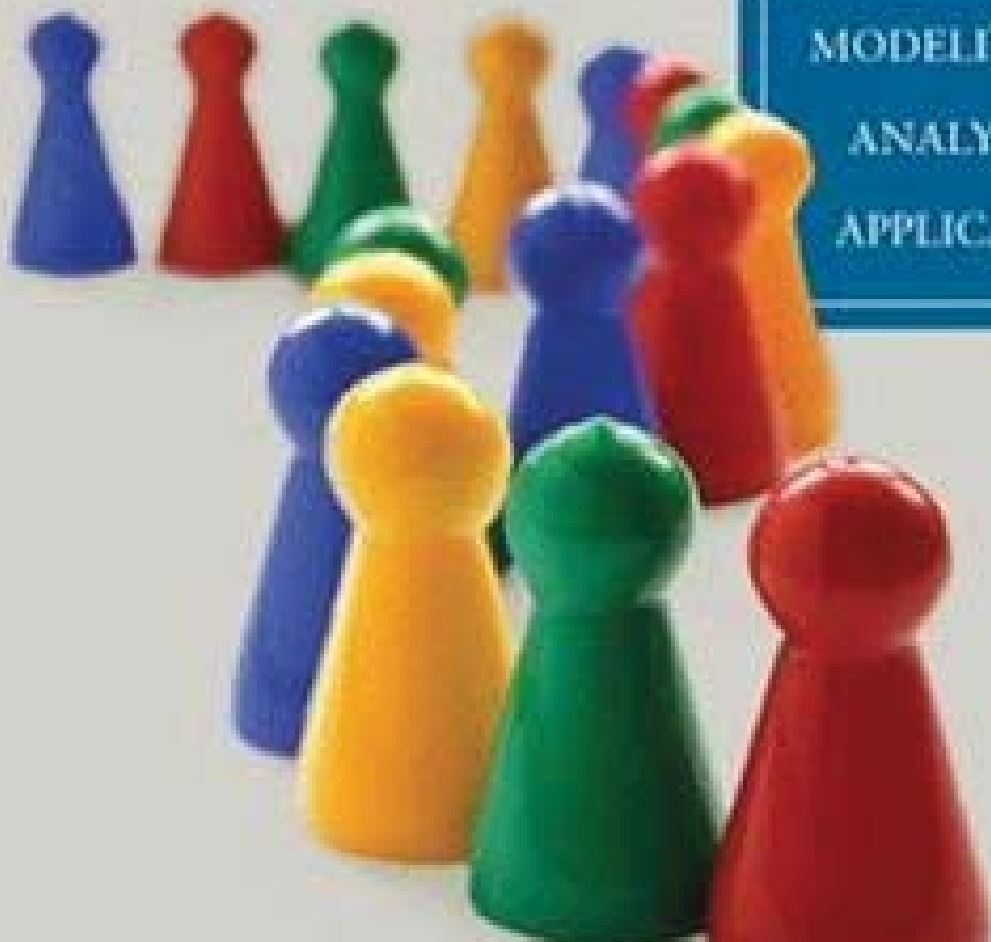
An Introduction to Queueing Theory

MODELING AND ANALYSIS IN APPLICATIONS



U. Narayan Bhat

BIRKHÄUSER

Statistics for Industry and Technology

U. Narayan Bhat

An Introduction to Queueing Theory

Modeling and Analysis in Applications

Birkhäuser
Boston • Basel • Berlin

U. Narayan Bhat
Professor *Emeritus*
Statistical Science & Operations Research
Southern Methodist University
Dallas, TX 75275-0332
USA

ISBN: 978-0-8176-4724-7 e-ISBN: 978-0-8176-4725-4
DOI: 10.1007/978-0-8176-4725-4

Library of Congress Control Number: 2007941114

Mathematics Subject Classification (2000): 60J27, 60K25, 60K30, 68M20, 90B22, 90B36

Cover design: Dutton and Sherman, Hamden, CT.

Printed on acid-free paper.

9 8 7 6 5 4 3 2 1

www.birkhauser.com

In memory of my parents,
Vaidya P. Ishwar and Parvati Bhat

Contents

Appendices

Preface

There are several books on queueing theory available for students as well as researchers. At the low end of mathematical sophistication, some provide usable formulas in a recipe fashion. At the high end there are research monographs on specific topics and books with an emphasis on theoretical analysis. In between there are a few textbooks with one common feature: all of them require an adequate background knowledge of probability and Markov processes that can be acquired normally with a semester-length graduate course. Consequently, most people who deal with the modeling and analysis of queueing systems either do not take a course on the subject because it would require an extra semester, or take a course on queueing systems without the necessary background and learn only how to use the results. This book is addressed to remedy this situation by providing a one-semester foundational introduction to the theory necessary for modeling and analysis of systems while developing the essential Markov process concepts and techniques using queueing processes as examples.

Some of the key features of the book also distinguish it from others. Its introductory chapter includes a historical perspective on the growth of queueing theory in the last 100 years. With an emphasis on modeling and analysis it deals with topics such as identification of models, collection of data, and tests for stationarity and independence of observations. It provides a rigorous treatment of basic models commonly used in applications with references for advanced topics. It gives a comprehensive discussion of statistical inference techniques usable in the modeling of queueing systems and an introduction to decision problems in their management. The book also includes a chapter, written by computer scientists, on the use of computational tools and simulation in solving queueing theory problems.

The book can be used as a text for first-year graduate students in applied science areas such as computer science, operations research, and industrial and/or systems engineering, and allied fields such as manufacturing and communication engineering. It can also serve as a text for upper-level undergraduate students in mathematics, statistics, and engineering who have a reasonable background in calculus and basic probability theory. This book is the product of the author's experience in teaching queueing theory for 40 years at various levels to students with or without the necessary background in stochastic processes.

The mathematical background assumed here is a two- or three-semester course in calculus, some exposure to transforms and matrices, and an introductory course in probability and statistics—all at the undergraduate level. An appendix on mathematical results provides some of the essential theorems for reference. Instructors may request a guide to the solutions of exercises via the Birkhäuser website at www.birkhauser.com/978-0-8176-4724-7.

The book does not advocate any specific software for the numerical analysis of queueing problems. The one chapter on modeling and analysis using computational tools employs MATLAB® for the purpose, and we believe students can benefit more by using mathematical software such as MATLAB and Mathematica® rather than system-specific software because of their limited scope.

For this author, writing the book has been a retirement project. He is indebted to Southern Methodist University and the Institute for the Study of Earth and Man for providing necessary resources and facilities even after his retirement. He acknowledges his gratitude to Professors Krishna Kavi and Robert Akl of the University of North Texas for contributing a chapter on the numerical analysis of queueing systems (in which the author's expertise is limited). Special acknowledgement of indebtedness is also made to the reviewers' comments, which have helped to improve the organization and contents of the book. The author also wishes to thank Professor N. Balakrishnan for recommending this book for inclusion in the *Statistics for Industry and Technology* series of Birkhäuser. Thanks are due to Professor Junfang Yu of the Department of Engineering Management, Information, and Systems of Southern Methodist University for using the prepublication copy of this book in his class and pointing out some of the typographical errors in it. Thanks are also due to Ms. Sheila Crain of the Department of Statistical Science for setting the manuscript in LaTeX with care and perserverance.

The author's wife, Girija, son Girish, and daughter Gouri have supported and encouraged him throughout his academic career. They deserve all the credit for his success.

U. Narayan Bhat
Dallas, TX
January 2008

1

Introduction

1.1 Basic System Elements

Queues (or waiting lines) help facilities or businesses provide service in an orderly fashion. Forming a queue being a social phenomenon, it is beneficial to the society if it can be managed so that both the unit that waits and the one that serves get the most benefit. For instance, there was a time when in airline terminals passengers formed separate queues in front of check-in counters. But now we see invariably only one line feeding into several counters. This is the result of the realization that a single line policy serves better for the passengers as well as the airline management. Such a conclusion has come from analyzing the mode by which a queue is formed and the service is provided. The analysis is based on building a mathematical model representing the process of arrival of passengers who join the queue, the rules by which they are allowed into service, and the time it takes to serve the passengers. Queueing theory embodies the full gamut of such models covering all perceivable systems that incorporate characteristics of a queue.

We identify the unit demanding service, whether it is human or otherwise, as the *customer*. The unit providing service is known as the *server*. This terminology of customers and servers is used in a generic sense regardless of the nature of the physical context. Some examples are given below:

(a) In communication systems, voice or data traffic queue up for lines for transmission. A simple example is the telephone exchange.
(b) In a manufacturing system with several work stations, units completing work in one station wait for access to the next.
(c) Vehicles requiring service wait for their turn in a garage.
(d) Patients arrive at a doctor's clinic for treatment.

Numerous examples of this type are of everyday occurrence. While analyzing them we can identify some basic elements of the systems.

Input process. If the occurrence of arrivals and the offer of service proceed strictly according to schedule, a queue can be avoided. But in practice this does not happen.

U.N. Bhat, *An Introduction to Queueing Theory*, DOI: 10.1007/978-0-8176-4725-4_1,