AN INTRODUCTION TO COMMUNICATION NETWORK ANALYSIS

George Kesidis
Pennsylvania State University

BICENTENNIAL 1807 WILEY 2007 BICENTENNIAL

WILEY-INTERSCIENCE
A John Wiley & Sons, Inc., Publication

This Page Intentionally Left Blank

AN INTRODUCTION TO COMMUNICATION NETWORK ANALYSIS

The Wiley Bicentennial–Knowledge for Generations

Each generation has its unique needs and aspirations. When Charles Wiley first opened his small printing shop in lower Manhattan in 1807, it was a generation of boundless potential searching for an identity. And we were there, helping to define a new American literary tradition. Over half a century later, in the midst of the Second Industrial Revolution, it was a generation focused on building the future. Once again, we were there, supplying the critical scientific, technical, and engineering knowledge that helped frame the world. Throughout the 20th Century, and into the new millennium, nations began to reach out beyond their own borders and a new international community was born. Wiley was there, expanding its operations around the world to enable a global exchange of ideas, opinions, and know-how.

For 200 years, Wiley has been an integral part of each generation's journey, enabling the flow of information and understanding necessary to meet their needs and fulfill their aspirations. Today, bold new technologies are changing the way we live and learn. Wiley will be there, providing you the must-have knowledge you need to imagine new worlds, new possibilities, and new opportunities.

Generations come and go, but you can always count on Wiley to provide you the knowledge you need, when and where you need it!

William J. Pesce

William J. Pesce
President and Chief Executive Officer

Peter Booth Wiley

Peter Booth Wiley
Chairman of the Board

AN INTRODUCTION TO COMMUNICATION NETWORK ANALYSIS

George Kesidis
Pennsylvania State University

BICENTENNIAL 1807 WILEY 2007 BICENTENNIAL

WILEY-INTERSCIENCE
A John Wiley & Sons, Inc., Publication

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data:

Kesidis, George.
An introduction to communication network analysis / George Kesidis.
p. cm.
ISBN 978-0-471-37141-0 (cloth)
1. Telecommunication—Traffic—Textbooks. 2. Network performance (Telecommunication)—Textbooks. 3. Switching theory—Textbooks. I. Title.
TK5102.985.K47 2007
621.382'1—dc22 2007002535

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

For Selena, Emma and Cleo

CONTENTS

This Page Intentionally Left Blank

PREFACE

This book was the basis of a single graduate course on the general subject of "performance" of communication networks for students from a broad set of backgrounds in electrical engineering, computer science, or computer engineering. The student was assumed to have basic familiarity with networking concepts as discussed in introductory texts on the subject, e.g., [139, 172, 220]. Also the student was assumed to have undergraduate courses in probability theory and linear (matrix) algebra.

Background material on probability and statistics is reviewed in Chapter 1. Graduate courses on probability and stochastic processes in electrical and computer engineering tend to focus on wide-sense stationary processes, typically in order to study the effects of noise in communication and control systems. In two successive chapters this book covers Markov chains and introduces the topic of queueing. Though the continuous-time context is stressed (to facilitate the queueing material), the discrete-time context is covered at the end of each chapter.

The remaining chapters pertain more directly to networking. Chapter 4 is on the subject of traffic shaping and multiplexing using a localized bandwidth resource. The next chapter describes queueing networks with static routing in the rather classical contexts of loss networks and open Jackson networks. Chapter 6 is on dynamic routing and routing with incentives including a game-theoretic model. The final chapter is a discussion of peer-to-peer networking systems, specifically those for the purposes of file sharing.

In general, problems at the end of each chapter review the described concepts and cover more specialized related material that may be of interest to the networking researcher. Worked solutions or references for certain problems are given in an appendix.

The length of the book allows time for about two weeks of lectures on material of specific interest to the instructor. The amount of instructor discretionary time can be increased by, for example, omitting coverage of the concluding sections of Chapters 4, 6, and 7 that are largely drawn from the author's own publications.

I thank my wife Diane, the editing staff at Wiley, and several of my colleagues and students for their time spent reading draft manuscripts and for their sage though often conflicting advice. In particular, I thank my students Y. Jin and B. Mortazavi. All flaws in this book are the sole responsibility of the author. Please contact me to report errors found in the book as an errata sheet will be made available on the Web.

George Kesidis

State College, Pennsylvania
kesidis@gmail.com
April 2007

CHAPTER 1

REVIEW OF ELEMENTARY PROBABILITY THEORY

This book assumes some familiarity with elementary probability theory. Good introductory texts are [63, 192]. We will begin by briefly reviewing some of these concepts in order to introduce notation and for future reference. We will also introduce basic notions of statistics that are particularly useful. See [106] for additional discussion of related material in a computer systems context.

1.1 SAMPLE SPACE, EVENTS, AND PROBABILITIES

Consider a random experiment resulting in an *outcome* (or "sample") represented by ω. For example, the experiment could be a pair of dice thrown onto a table and the outcome could be the exact orientation of the dice and their position on the table when they stop moving. The abstract space of all outcomes (called the *sample space*) is normally denoted by Ω, i.e., $\omega \in \Omega$.

An *event* is merely a subset of Ω. For example, in a dice throwing experiment, an event is "both dice land in a specific region of the table" or "the sum of the dots on the upward facing surfaces of the dice is 7." Clearly, many different individual outcomes ω belong to these events. We say that an event A has *occurred* if the outcome ω of the random experiment belongs to A, i.e., $w \in A$, where $A \subset \Omega$. Now consider two events A and B. We therefore say that A *and* B occur if the outcome $\omega \in A \cap B$. Also, we say that A *or* B occur if the outcome $\omega \in A \cup B$.

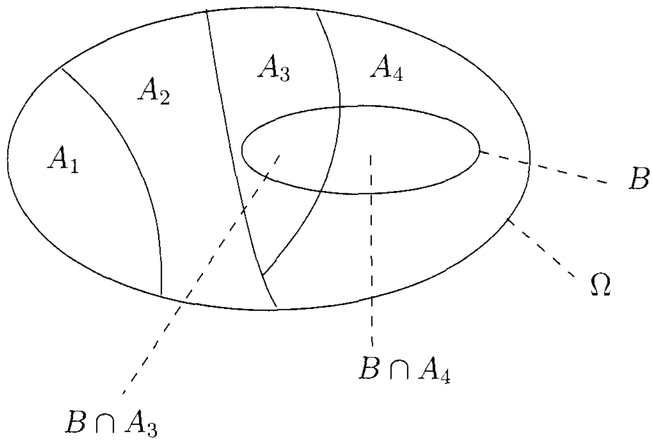

Figure 1.1 A partition of Ω.

A *probability measure* P maps each event $A \subset \Omega$ to a real number between zero and one inclusive, i.e., $\mathsf{P}(A) \in [0, 1]$. A probability measure has certain properties such as $\mathsf{P}(\Omega) = 1$ and

$$\mathsf{P}(A) \quad = \quad 1 - \mathsf{P}(\bar{A}),$$

where $\bar{A} = \{\omega \in \Omega \mid \omega \notin A\}$ is the complement of A. Also, if the events $\{A_i\}_{i=1}^n$ are disjoint (i.e., $A_i \cap A_j = \emptyset$ for all $i \neq j$), then

$$\mathsf{P}\left(\bigcup_{i=1}^n A_i\right) \quad = \quad \sum_{i=1}^n P(A_i),$$

i.e., P is finitely additive. Formally, a probability measure is defined to be countably additive. Also, $P(A)$ is defined only for events $A \subset \Omega$ that belong to a *σ-field* (sigma-field) or *σ-algebra* of events. These details are beyond the scope of this book.

The probability of an event A *conditioned on* (or "given that") another event B has occurred is

$$\mathsf{P}(A|B) \quad \equiv \quad \frac{\mathsf{P}(A \cap B)}{\mathsf{P}(B)},$$

where $\mathsf{P}(B) > 0$ is assumed. Now suppose the events $A_1, A_2, ..., A_n$ form a *partition* of Ω, i.e.,

$$\bigcup_{i=1}^n A_i = \Omega \quad \text{and} \quad A_i \cap A_j = \emptyset \text{ for all } i \neq j.$$

Assuming that $\mathsf{P}(A_i) > 0$ for all i, the *law of total probability* states that, for any event B,

$$\mathsf{P}(B) \quad = \quad \sum_{i=1}^n \mathsf{P}(B|A_i)\mathsf{P}(A_i). \tag{1.1}$$

Note that the events $A_i \cap B$ form a partition of $B \subset \Omega$. See Figure 1.1, where $B = (B \cap A_4) \cup (B \cap A_3)$ and $B \cap A_i = \emptyset$ for $i = 1, 2$.

A group of events $A_1, A_2, ..., A_n$ are said to be *mutually independent* (or just "independent") if

$$\mathsf{P}\left(\bigcap_{i \in \mathcal{I}} A_i\right) = \prod_{i \in \mathcal{I}} \mathsf{P}(A_i)$$

for all subsets $\mathcal{I} \subset \{1, 2, ..., n\}$. Note that if events A and B are independent and $\mathsf{P}(B) > 0$, then $\mathsf{P}(A|B) = \mathsf{P}(A)$; therefore, knowledge that the event B has occurred has no bearing on the probability that the event A has occurred as well.

In the following, a comma between events will represent an intersection symbol, for example, the probability that A *and* B occur is

$$\mathsf{P}(A, B) \equiv \mathsf{P}(A \cap B).$$

1.2 RANDOM VARIABLES

A *random variable* X is a real-valued function with domain Ω. That is, for each outcome ω, $X(\omega)$ is a real number representing some feature of the outcome. For example, in a dice-throwing experiment, $X(\omega)$ could be defined as the sum of the dots on the upward-facing surfaces of outcome ω. Formally, random variables are defined to be *measurable* in the sense that the event $X^{-1}(B) = \{\omega \in \Omega \mid X(\omega) \in B\}$ is an event (a member of the σ-field on Ω) for all events $B \subset \mathbb{R}$ (belonging to the *Borel* σ-field of subsets of $\mathbb{R}$). In this way, the quantity $\mathsf{P}(X \in B)$ is well defined for any set B that is of interest. Again, the details of this measurability condition are beyond the scope of this book. In the following, all functions are implicitly assumed to be measurable.

The strict range of X is defined to be the *smallest* subset R_X of $\mathbb{R}$ such that

$$\mathsf{P}(X \in R_X) = 1,$$

where $\mathsf{P}(X \in R_X)$ is short for $\mathsf{P}(\{\omega \in \Omega \mid X(\omega) \in R_X\})$.

Note that a (Borel-measurable) function g of a random variable X, $g(X)$, is also a random variable.

A group of random variables $X_1, X_2, ..., X_n$ are said to be mutually independent (or just "independent") if, for any collection $\{B_i\}_{i=1}^n$ of subsets of $\mathbb{R}$, the events $\{X_i \in B_i\}_{i=1}^n$ are independent; see Section 1.9.

1.3 CUMULATIVE DISTRIBUTION FUNCTIONS, EXPECTATION, AND MOMENT GENERATING FUNCTIONS

The probability *distribution* of a random variable X connotes the information $\mathsf{P}(X \in B)$ for all events $B \in \mathbb{R}$. We need only stipulate

$$\mathsf{P}(X \leq x) \equiv \mathsf{P}(X \in (-\infty, x])$$

for all $x \in \mathbb{R}$ to completely specify the distribution of X; see Equation (1.4). This leads us to define the *cumulative distribution function* (CDF) F_X of a random variable X as

$$F_X(x) = \mathsf{P}(X \leq x) \tag{1.2}$$

for $x \in \mathbb{R}$, where $\mathsf{P}(X \leq x)$ is, again, short for $\mathsf{P}(\{\omega \in \Omega \mid X(\omega) \leq x\})$. Clearly, a CDF F_X takes values in $[0, 1]$, is nondecreasing on $\mathbb{R}$, $F_X(x) \to 1$ as $x \to \infty$, and $F_X(x) \to 0$ as $x \to -\infty$.

The *expectation* of a random variable is simply its average (or "mean") value. We can define the expectation of a function g of a random variable X as

$$\mathsf{E}(g(X)) = \int_{-\infty}^{\infty} g(x)\, \mathrm{d}F_X(x), \tag{1.3}$$

where we have used a Stieltjes integral [133] that will be explained via explicit examples in the following. Note here that the expectation (when it exists) is simply a real number or $\pm\infty$. Also note that expectation is a linear operation over random variables. That is, for any two random variables X and Y and any two real constants a and b,

$$\mathsf{E}(aX + bY) = a\mathsf{E}X + b\mathsf{E}Y.$$

■ EXAMPLE 1.1

Suppose g is an *indicator function*, i.e., for some event $B \subset \mathbb{R}$

$$\begin{aligned} g(X(\omega)) &\equiv \mathbf{1}\{X(\omega) \in B\} \\ &\equiv \begin{cases} 1 & \text{if } X(\omega) \in B, \\ 0 & \text{else.} \end{cases} \end{aligned}$$

In this case,

$$\mathsf{E}g(X) = \mathsf{P}(X \in B) = \int_B \mathrm{d}F_X(x), \tag{1.4}$$

where the notation refers to integration over the set B.

The nth *moment* of X is $\mathsf{E}(X^n)$ and the *variance* of X is

$$\sigma_X^2 \equiv \text{var}(X) \equiv \mathsf{E}(X - \mathsf{E}X)^2,$$

i.e., the variance is the second *centered* moment. The *standard deviation* of X is the square root of the variance, $\sigma_X \geq 0$. The *moment generating function* (MGF) of X is

$$m_X(\theta) = \mathsf{E}e^{\theta X},$$

where θ is a real number. The moment generating function can also be used to completely describe the distribution of a random variable.

1.4 DISCRETELY DISTRIBUTED RANDOM VARIABLES

For a *discretely distributed* (or just "discrete") random variable X, there is a set of countably many real numbers $\{a_i\}_{i=1}^{\infty}$ such that

$$\sum_{i=1}^{\infty} \mathsf{P}(X = a_i) = 1.$$

Assuming $\mathsf{P}(X = a_i) > 0$ for all i, the countable set $\{a_i\}_{i=1}^{\infty}$ is the strict range of X, i.e., $R_X = \{a_i\}_{i=1}^{\infty}$. So, a discrete random variable has a piecewise constant CDF with a countable number of jump discontinuities occurring at the a_i's. That is, if the a_i are defined so as to be an increasing sequence, F is constant on each open interval (a_i, a_{i+1}), $F(x) = 0$ for $x < a_1$, and

$$F(a_i) = \sum_{j=1}^{i} \mathsf{P}(X = a_j) = \sum_{j=1}^{i} p(a_j),$$

where p is the probability mass function (PMF) of the discrete random variable X, i.e.,

$$p(a_i) \quad \equiv \quad \mathsf{P}(X = a_i).$$

Note that we have dropped the subscript "X" on the PMF and CDF for notational convenience. Moreover, for any $B \subset \mathbb{R}$ and any real-valued function g over $\mathbb{R}$,

$$\mathsf{P}(X \in B) \quad = \quad \sum_{a_j \in B} p(a_j)$$

and

$$\mathsf{E}g(X) \quad = \quad \sum_{j=1}^{\infty} g(a_j)p(a_j) = \sum_{a \in R_X} g(a)p(a).$$

To see the connection between this expression and (1.3), note that

$$\begin{aligned} \mathrm{d}F(x) &= F'(x)\,\mathrm{d}x \\ &= \sum_{i=1}^{\infty} p(a_i)\delta(x - a_i)\,\mathrm{d}x, \end{aligned}$$

where δ is the Dirac delta function [164]. That is, δ is the unit impulse satisfying $\delta(t) = 0$ for all $t \neq 0$ and

$$\int_{-\infty}^{\infty} \delta(t)\,\mathrm{d}t \quad = \quad 1.$$

1.4.1 The Bernoulli distribution

A random variable X that is *Bernoulli* distributed has strict range consisting of two elements, typically $R_X = \{0, 1\}$. So, there is a real parameter $q \in (0, 1)$ such that $q = \mathsf{P}(X = 1) = 1 - \mathsf{P}(X = 0)$. Also,

$$\mathsf{E}g(X) \;=\; (1 - q) \cdot g(0) + q \cdot g(1)$$

with $\mathsf{E}X = q$ in particular.

1.4.2 The geometric distribution

A random variable X that is *geometrically* distributed has a single parameter $\lambda > 0$ and its strict range is the nonnegative integers, i.e.,

$$R_X = \mathbb{Z}^+ \equiv \{0, 1, 2, ...\}.$$

The parameter λ satisfies $0 < \lambda < 1$. The CDF of X is piecewise constant with

$$F(i) = 1 - \lambda^{i+1}$$

for all $i \in \mathbb{Z}^+$. The PMF of X is $p(i) = (1-\lambda)\lambda^i$ for $i \in \mathbb{Z}^+$. To compute $\mathsf{E}X$, we rely on a little trick involving a derivative:

$$\begin{aligned}
\mathsf{E}X &= \sum_{i=0}^{\infty} ip(i) \\
&= (1-\lambda)\lambda \sum_{i=1}^{\infty} i\,\lambda^{i-1} \\
&= (1-\lambda)\lambda \frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\sum_{i=1}^{\infty} \lambda^i\right) \\
&= (1-\lambda)\lambda \frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{1}{1-\lambda} - 1\right) \\
&= (1-\lambda)\lambda \frac{1}{(1-\lambda)^2} \\
&= \frac{\lambda}{1-\lambda}.
\end{aligned}$$

Similarly, the moment generating function is

$$\begin{aligned}
m(\theta) &= (1-\lambda)\sum_{i=0}^{\infty}(\mathrm{e}^{\theta}\lambda)^i \\
&= \frac{1-\lambda}{1-\lambda\mathrm{e}^{\theta}}
\end{aligned}$$

for $\mathrm{e}^{\theta}\lambda < 1$, i.e., $\theta < -\log\lambda$.

1.4.3 The binomial distribution

A random variable Y is *binomially distributed* with parameters n and q if $R_Y = \{0, 1, ..., n\}$ and, for $k \in R_Y$,

$$\mathsf{P}(Y = k) = \binom{n}{k} q^k (1-q)^{n-k},$$

where $n \in \mathbb{Z}^+$, $0 < q < 1$, and

$$\binom{n}{k} \equiv \frac{n!}{k!(n-k)!}. \tag{1.5}$$

That is, $\binom{n}{k}$ is "n choose k," see Example 1.4. It is easy to see that, by the binomial theorem, $\sum_{k=0}^{n} \mathsf{P}(Y = k) = 1$, i.e.,

$$\sum_{k=0}^{n} \binom{n}{k} q^k (1-q)^{n-k} = (q + (1-q))^n = 1.$$

Also,

$$\begin{aligned} m(\theta) &= \sum_{k=0}^{n} \binom{n}{k} q^k (1-q)^{n-k} e^{\theta k} \\ &= (qe^\theta + (1-q))^n. \end{aligned}$$

■ **EXAMPLE 1.2**

If we are given n independent Bernoulli distributed random variables, X_i, each having the same parameter q, then $Y = \sum_{i=1}^{n} X_i$ is binomially distributed with parameters n and q. That is, for $k \in \{0, 1, 2, ..., n\}$, the event $\{Y = k\}$ can be written as a union on disjoint component events, where k of the X_i equal 1 and $n - k$ of the X_i equal 0. Each such component event occurs with probability $p^k(1-p)^{n-k}$. The number of such events, i.e., the number of ways the random vector $(X_1, X_2, ..., X_n)$ has exactly k ones, is

$$\frac{n!}{k!(n-k)!} = \binom{n}{k},$$

where $n!$ is the number of *permutations* (ordered arrangements) of n different objects and the factors $k!$ and $(n-k)!$ in the denominator account for the k ones being indistinguishable and the $n - k$ zeros being indistinguishable.

1.4.4 The Poisson distribution

A random variable X is *Poisson* distributed with parameter $\lambda > 0$ if $R_X = \mathbb{Z}^+$ and the PMF is

$$p(i) = \frac{\lambda^i}{i!} e^{-\lambda}$$

for $i \in \mathbb{Z}^+$. We can check that $\mathsf{E}X = \lambda$ as in the geometric case. The MGF is

$$\begin{aligned} m(\theta) &= \sum_{i=0}^{\infty} e^{\theta n} \frac{\lambda^n}{n!} e^{-\lambda} \\ &= e^{-\lambda} \sum_{i=0}^{\infty} \frac{(e^\theta \lambda)^n}{n!} \\ &= \exp((e^\theta - 1)\lambda). \end{aligned}$$

1.4.5 The discrete uniform distribution

A discrete random variable X is *uniformly* distributed on a *finite* range

$$R_X \quad \subset \quad \mathbb{R}$$

if

$$\mathsf{P}(X = x) \quad = \quad \frac{1}{|R_X|}$$

for all $x \in R_X$, where $|R_X|$ is the size of (the number of elements in) R_X. Clearly, therefore, for any $A \subset R_X$,

$$\mathsf{P}(X \in A) \quad = \quad \frac{|A|}{|R_X|},$$

i.e., to compute this probability, one needs to *count* the number of elements in A and R_X.

■ **EXAMPLE 1.3**

Suppose that a random experiment consists of tossing two different six-sided dice on the floor. Consider the events consisting of all outcomes having the same numbers (d_1, d_2) on the upturned faces of the dice. Note that there are $6 \times 6 = 36$ such events. Assume that the probability of each such event is $\frac{1}{36}$, i.e., the dice are "fair." This implies that the random variables d_i are independent and uniformly distributed on their state space $\{1, 2, 3, 4, 5, 6\}$.

Suppose that we are interested in $\mathsf{P}(X \in \{7, 11\})$, where the random variable

$$X \quad \equiv \quad d_1 + d_2.$$

That is, we are interested in the event

$$(d_1, d_2) \in \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1), (5,6), (6,5)\}$$

with eight members. So, $\mathsf{P}(X \in \{7, 11\}) = \frac{8}{36}$.

■ **EXAMPLE 1.4**

Suppose that five cards (a poker hand) are drawn, without replacement, from a standard deck of 52 different playing cards. The random variable X enumerates each *combination* (not considering the order in which the individual cards were drawn) of poker hands beginning with 1 and ending with the total number of different poker