Probability and its Applications

A Series of the Applied Probability Trust

Editors: J. Gani, C.C. Heyde, T.G. Kurtz

Springer
New York
Berlin
Heidelberg
Hong Kong
London
Milan
Paris
Tokyo

D.J. Daley D. Vere-Jones

An Introduction to the Theory of Point Processes

Volume I: Elementary Theory and Methods

Second Edition

Springer

D.J. Daley
Centre for Mathematics and its Applications
Mathematical Sciences Institute
Australian National University
Canberra, ACT 0200, Australia
daryl@maths.anu.edu.au

D. Vere-Jones
School of Mathematical and Computing Sciences
Victoria University of Wellington
Wellington, New Zealand
David.Vere-Jones@mcs.vuw.ac.nz

Library of Congress Cataloging-in-Publication Data
Daley, Daryl J.
An introduction to the theory of point processes / D.J. Daley, D. Vere-Jones.
p. cm.
Includes bibliographical references and index.
Contents: v. 1. Elementary theory and methods
ISBN 0-387-95541-0 (alk. paper)
1. Point processes. I. Vere-Jones, D. (David) II. Title
QA274.42.D35 2002
519.2´3—dc21 2002026666

ISBN 0-387-95541-0 Printed on acid-free paper.

Printed in the United States of America.

9 8 7 6 5 4 3 2 1 SPIN 10885680

Typesetting: Photocomposed pages prepared by the authors using plain TeX files.

www.springer-ny.com

Springer-Verlag New York Berlin Heidelberg
A member of BertelsmannSpringer Science+Business Media GmbH

To Nola,
and in memory of Mary

Preface to the Second Edition

In preparing this second edition, we have taken the opportunity to reshape the book, partly in response to the further explosion of material on point processes that has occurred in the last decade but partly also in the hope of making some of the material in later chapters of the first edition more accessible to readers primarily interested in models and applications. Topics such as conditional intensities and spatial processes, which appeared relatively advanced and technically difficult at the time of the first edition, have now been so extensively used and developed that they warrant inclusion in the earlier introductory part of the text. Although the original aim of the book—to present an introduction to the theory in as broad a manner as we are able—has remained unchanged, it now seems to us best accomplished in two volumes, the first concentrating on introductory material and models and the second on structure and general theory. The major revisions in this volume, as well as the main new material, are to be found in Chapters 6–8. The rest of the book has been revised to take these changes into account, to correct errors in the first edition, and to bring in a range of new ideas and examples.

Even at the time of the first edition, we were struggling to do justice to the variety of directions, applications and links with other material that the theory of point processes had acquired. The situation now is a great deal more daunting. The mathematical ideas, particularly the links to statistical mechanics and with regard to inference for point processes, have extended considerably. Simulation and related computational methods have developed even more rapidly, transforming the range and nature of the problems under active investigation and development. Applications to spatial point patterns, especially in connection with image analysis but also in many other scientific disciplines, have also exploded, frequently acquiring special language and techniques in the different fields of application. Marked point processes, which were clamouring for greater attention even at the time of the first edition, have acquired a central position in many of these new applications, influencing both the direction of growth and the centre of gravity of the theory.

We are sadly conscious of our inability to do justice to this wealth of new material. Even less than at the time of the first edition can the book claim to provide a comprehensive, up-to-the-minute treatment of the subject. Nor are we able to provide more than a sketch of how the ideas of the subject have evolved. Nevertheless, we hope that the attempt to provide an introduction to the main lines of development, backed by a succinct yet rigorous treatment of the theory, will prove of value to readers in both theoretical and applied fields and a possible starting point for the development of lecture courses on different facets of the subject. As with the first edition, we have endeavoured to make the material as self-contained as possible, with references to background mathematical concepts summarized in the appendices, which appear in this edition at the end of Volume I.

We would like to express our gratitude to the readers who drew our attention to some of the major errors and omissions of the first edition and will be glad to receive similar notice of those that remain or have been newly introduced. Space precludes our listing these many helpers, but we would like to acknowledge our indebtedness to Rick Schoenberg, Robin Milne, Volker Schmidt, Günter Last, Peter Glynn, Olav Kallenberg, Martin Kalinke, Jim Pitman, Tim Brown and Steve Evans for particular comments and careful reading of the original or revised texts (or both). Finally, it is a pleasure to thank John Kimmel of Springer-Verlag for his patience and encouragement, and especially Eileen Dallwitz for undertaking the painful task of rekeying the text of the first edition.

The support of our two universities has been as unflagging for this endeavour as for the first edition; we would add thanks to host institutions of visits to the Technical University of Munich (supported by a Humboldt Foundation Award), University College London (supported by a grant from the Engineering and Physical Sciences Research Council) and the Institute of Mathematics and its Applications at the University of Minnesota.

Daryl Daley
Canberra, Australia

David Vere-Jones
Wellington, New Zealand

Preface to the First Edition

This book has developed over many years—too many, as our colleagues and families would doubtless aver. It was conceived as a sequel to the review paper that we wrote for the Point Process Conference organized by Peter Lewis in 1971. Since that time the subject has kept running away from us faster than we could organize our attempts to set it down on paper. The last two decades have seen the rise and rapid development of martingale methods, the surge of interest in stochastic geometry following Rollo Davidson's work, and the forging of close links between point processes and equilibrium problems in statistical mechanics.

Our intention at the beginning was to write a text that would provide a survey of point process *theory* accessible to beginning graduate students and workers in applied fields. With this in mind we adopted a partly historical approach, starting with an informal introduction followed by a more detailed discussion of the most familiar and important examples, and then moving gradually into topics of increased abstraction and generality. This is still the basic pattern of the book. Chapters 1–4 provide historical background and treat fundamental special cases (Poisson processes, stationary processes on the line, and renewal processes). Chapter 5, on finite point processes, has a bridging character, while Chapters 6–14 develop aspects of the general theory.

The main difficulty we had with this approach was to decide when and how far to introduce the abstract concepts of functional analysis. With some regret, we finally decided that it was idle to pretend that a general treatment of point processes could be developed without this background, mainly because the problems of existence and convergence lead inexorably to the theory of measures on metric spaces. This being so, one might as well take advantage of the metric space framework from the outset and let the point process itself be defined on a space of this character: at least this obviates the tedium of having continually to specify the dimensions of the Euclidean space, while in the context of completely separable metric spaces—and this is the greatest generality we contemplate—intuitive spatial notions still provide a reasonable guide to basic properties. For these reasons the general results from Chapter 6 onward are couched in the language of this setting, although the examples continue to be drawn mainly from the one- or two-dimensional Euclidean spaces $\mathbb{R}^1$ and $\mathbb{R}^2$. Two appendices collect together the main results we need from measure theory and the theory of measures on metric spaces. We hope that their inclusion will help to make the book more readily usable by applied workers who wish to understand the main ideas of the general theory without themselves becoming experts in these fields. Chapter 13, on the martingale approach, is a special case. Here the context is again the real line, but we added a third appendix that attempts to summarize the main ideas needed from martingale theory and the general theory of processes. Such special treatment seems to us warranted by the exceptional importance of these ideas in handling the problems of inference for point processes.

In style, our guiding star has been the texts of Feller, however many light-years we may be from achieving that goal. In particular, we have tried to follow his format of motivating and illustrating the general theory with a range of examples, sometimes didactical in character, but more often taken from real applications of importance. In this sense we have tried to strike a mean between the rigorous, abstract treatments of texts such as those by Matthes, Kerstan and Mecke (1974/1978/1982) and Kallenberg (1975, 1983), and practically motivated but informal treatments such as Cox and Lewis (1966) and Cox and Isham (1980).

Numbering Conventions. Each chapter is divided into sections, with consecutive labelling within each of equations, statements (encompassing Definitions, Conditions, Lemmas, Propositions, Theorems), examples, and the exercises collected at the end of each section. Thus, in Section 1.2, (1.2.3) is the third equation, **Statement 1.2.III** is the third statement, EXAMPLE 1.2(c) is the third example, and Exercise 1.2.3 is the third exercise. The exercises are varied in both content and intention and form a significant part of the text. Usually, they indicate extensions or applications (or both) of the theory and examples developed in the main text, elaborated by hints or references intended to help the reader seeking to make use of them. The symbol $\square$ denotes the end of a proof. Instead of a name index, the listed references carry page number(s) where they are cited. A general outline of the notation used has been included before the main text.

It remains to acknowledge our indebtedness to many persons and institutions. Any reader familiar with the development of point process theory over the last two decades will have no difficulty in appreciating our dependence on the fundamental monographs already noted by Matthes, Kerstan and Mecke in its three editions (our use of the abbreviation MKM for the 1978 English edition is as much a mark of respect as convenience) and Kallenberg in its two editions. We have been very conscious of their generous interest in our efforts from the outset and are grateful to Olav Kallenberg in particular for saving us from some major blunders. A number of other colleagues, notably

David Brillinger, David Cox, Klaus Krickeberg, Robin Milne, Dietrich Stoyan, Mark Westcott, and Deng Yonglu, have also provided valuable comments and advice for which we are very grateful. Our two universities have responded generously with seemingly unending streams of requests to visit one another at various stages during more intensive periods of writing the manuscript. We also note visits to the University of California at Berkeley, to the Center for Stochastic Processes at the University of North Carolina at Chapel Hill, and to Zhongshan University at Guangzhou. For secretarial assistance we wish to thank particularly Beryl Cranston, Sue Watson, June Wilson, Ann Milligan, and Shelley Carlyle for their excellent and painstaking typing of difficult manuscript.

Finally, we must acknowledge the long-enduring support of our families, and especially our wives, throughout: they are not alone in welcoming the speed and efficiency of Springer-Verlag in completing this project.

Daryl Daley
Canberra, Australia

David Vere-Jones
Wellington, New Zealand

Contents

Chapter Titles for Volume II

Principal Notation

Very little of the general notation used in Appendices 1–3 is given below. Also, notation that is largely confined to one or two sections of the same chapter is mostly excluded, so that neither all the symbols used nor all the uses of the symbols shown are given. The repeated use of some symbols occurs as a result of point process theory embracing a variety of topics from the theory of stochastic processes. Where they are given, page numbers indicate the first or significant use of the notation. Generally, the particular interpretation of symbols with more than one use is clear from the context.

Throughout the lists below, N denotes a point process and ξ denotes a random measure.

Spaces

$\mathbb{C}$	complex numbers	
$\mathbb{R}^d$	d-dimensional Euclidean space	
$\mathbb{R} = \mathbb{R}^1$	real line	
$\mathbb{R}_+$	nonnegative numbers	
$\mathbb{S}$	circle group and its representation as $(0, 2\pi]$	
$\mathbb{U}^d_{2\alpha}$	d-dimensional cube of side length 2α and vertices $(\pm\,\alpha, \ldots, \pm\,\alpha)$	
$\mathbb{Z},\ \mathbb{Z}_+$	integers of $\mathbb{R}$, $\mathbb{R}_+$	
$\mathcal{X}$	state space of N or ξ; often $\mathcal{X} = \mathbb{R}^d$; always $\mathcal{X}$ is c.s.m.s. (complete separable metric space)	
Ω	space of probability elements ω	
$\emptyset,\ \emptyset(\cdot)$	null set, null measure	
$\mathcal{E}$	measurable sets in probability space	
$(\Omega, \mathcal{E}, \mathcal{P})$	basic probability space on which N and ξ are defined	158
$\mathcal{X}^{(n)}$	n-fold product space $\mathcal{X} \times \cdots \times \mathcal{X}$	123
$\mathcal{X}^{\cup}$	$= \mathcal{X}^{(0)} \cup \mathcal{X}^{(1)} \cup \cdots$	129

$\mathcal{B}(\mathcal{X})$	Borel σ-field generated by open spheres of c.s.m.s. $\mathcal{X}$	34
$\mathcal{B}_{\mathcal{X}}$	$= \mathcal{B}(\mathcal{X}),\ \mathcal{B} = \mathcal{B}_{\mathbb{R}} = \mathcal{B}(\mathbb{R})$	34, 374
$\mathcal{B}_{\mathcal{X}}^{(n)} = \mathcal{B}(\mathcal{X}^{(n)})$	product σ-field on product space $\mathcal{X}^{(n)}$	129
$\mathrm{BM}(\mathcal{X})$	measurable functions of bounded support	161
$\mathrm{BM}_{+}(\mathcal{X})$	measurable nonnegative functions of bounded support	161
$\mathcal{K}$	mark space for marked point process (MPP)	194
$\mathcal{M}_{\mathcal{X}}$ $(\mathcal{N}_{\mathcal{X}})$	totally finite (counting) measures on c.s.m.s. $\mathcal{X}$	158, 398
$\mathcal{M}_{\mathcal{X}}^{\#}$	boundedly finite measures on c.s.m.s. $\mathcal{X}$	158, 398
$\mathcal{N}_{\mathcal{X}}^{\#}$	boundedly finite counting measures on c.s.m.s. $\mathcal{X}$	131
$\mathcal{P}^{+}$	p.p.d. (positive positive-definite) measures	359
$\mathcal{S}$	infinitely differentiable functions of rapid decay	357
$\mathcal{U}$	complex-valued Borel measurable functions on $\mathcal{X}$ of modulus ≤ 1	144
$\mathcal{U} \otimes \mathcal{V}$	product topology on product space $\mathcal{X} \times \mathcal{Y}$ of topological spaces $(\mathcal{X},\mathcal{U})$, $(\mathcal{Y},\mathcal{V})$	378
$\mathcal{V} = \mathcal{V}(\mathcal{X})$	$[0,1]$-valued measurable functions $h(x)$ with $1 - h(x)$ of bounded support in $\mathcal{X}$	149, 152

General

Unless otherwise specified, $A \in \mathcal{B}_{\mathcal{X}}$, k and $n \in \mathbb{Z}_{+}$, t and $x \in \mathbb{R}$, $h \in \mathcal{V}(\mathcal{X})$, and $z \in \mathbb{C}$.

$\tilde{\ }$	$\tilde{\nu}, \widetilde{F}$ = Fourier–Stieltjes transforms of measure ν or d.f. F	411–412
	$\tilde{\phi}$ = Fourier transform of Lebesgue integrable function ϕ for counting measures	357
$\breve{\ }$	reduced (ordinary or factorial) (moment or cumulant) measure	160
#	extension of concept from totally finite to boundedly finite measure space	158
$\Vert\mu\Vert$	variation norm of measure μ	374
a.e. μ, μ-a.e.	almost everywhere with respect to measure μ	376
a.s., $\mathcal{P}$-a.s.	almost sure, $\mathcal{P}$-almost surely	376
$A^{(n)}$	n-fold product set $A \times \cdots \times A$	130
$\mathcal{A}$	family of sets generating $\mathcal{B}$; semiring of bounded Borel sets generating $\mathcal{B}_{\mathcal{X}}$	31, 368
B_u (T_u)	backward (forward) recurrence time at u	58, 76
$c_k,\ c_{[k]}$	kth cumulant, kth factorial cumulant, of distribution $\{p_n\}$	116
$c(x) = c(y, y + x)$	covariance density of stationary mean square continuous process on $\mathbb{R}^d$	160, 358

$C_{[k]}(\cdot),\ c_{[k]}(\cdot)$	factorial cumulant measure and density	147
$\breve{C}_2(\cdot), \breve{c}(\cdot)$	reduced covariance measure of stationary N or ξ	292
$\breve{c}(\cdot)$	reduced covariance density of stationary N or ξ	160, 292
$\delta(\cdot)$	Dirac delta function	
$\delta_x(A)$	Dirac measure, $= \int_A \delta(u-x)\,\mathrm{d}u = I_A(x)$	382
$\Delta F(x) = F(x) - F(x-)$	jump at x in right-continuous function F	107
$\mathrm{e}_\lambda(x) = (\frac{1}{2}\lambda)^d \exp\big(-\lambda \sum_{i=1}^d \lvert x_i \rvert\big)$	two-sided exponential density in $\mathbb{R}^d$	359
F	renewal process lifetime d.f.	67
F^{n*}	n-fold convolution power of measure or d.f. F	55
$F(\cdot\,;\cdot)$	finite-dimensional (fidi) distribution	158–161
$\boldsymbol{\mathcal{F}}$	history	236, 240
$\Phi(\cdot)$	characteristic functional	15
$G[h]$	probability generating functional (p.g.fl.) of N,	15, 144
$G[h \mid x]$	member of measurable family of p.g.fl.s	166
$G_c[\cdot],\ G_m[\cdot \mid x]$	p.g.fl.s of cluster centre and cluster member processes N_c and $N_m(\cdot \mid x)$	178
$G,\ G_I$	expected information gain (per interval) of stationary N on $\mathbb{R}$	280, 285
$\Gamma(\cdot),\ \gamma(\cdot)$	Bartlett spectrum, its density when it exists	304
$H(\mathcal{P};\mu)$	generalized entropy	277, 283
$\boldsymbol{\mathcal{H}},\ \boldsymbol{\mathcal{H}}^*$	internal history of ξ on $\mathbb{R}_+$, $\mathbb{R}$	236
$I_A(x) = \delta_x(A)$	indicator function of element x in set A	
$I_n(x)$	modified Bessel function of order n	72
$J_n(A_1 \times \cdots \times A_n)$	Janossy measure	124
$j_n(x_1,\ldots,x_n)$	Janossy density	125
$J_n(\cdot \mid A)$	local Janossy measure	137
K	compact set	371
$K_n(\cdot),\ k_n(\cdot)$	Khinchin measure and density	146
$\ell(\cdot)$	Lebesgue measure in $\mathcal{B}(\mathbb{R}^d)$,	31
	Haar measure on σ-group	408–409
$L_u = B_u + T_u$	current lifetime of point process on $\mathbb{R}$	58, 76
$L[f]\ (f \in BM_+(\mathcal{X}))$	Laplace functional of ξ	161
$L_\xi[1-h]$	p.g.fl. of Cox process directed by ξ	170
$L_2(\xi^0),\ L_2(\Gamma)$	Hilbert spaces of square integrable r.v.s ξ^0, and of functions square integrable w.r.t. measure Γ	332
$L_A(x_1,\ldots,x_n), = j_N(x_1,\ldots,x_N \mid A)$	likelihood, local Janossy density, $N \equiv N(A)$	22, 212
λ	rate of N, especially intensity of stationary N	46
$\lambda^*(t)$	conditional intensity function	231
$m_k\ (m_{[k]})$	kth (factorial) moment of distribution $\{p_n\}$	115

$\breve{m}_2$, $\breve{M}_2$	reduced second-order moment density, measure, of stationary N	289
m_{g}	mean density of ground process N_{g} of MPP N	198, 323
$N(A)$	number of points in A	42
$N(a,b]$	number of points in half-open interval $(a,b]$,	19
	$= N((a,b])$	42
$N(t)$	$= N(0,t] = N((0,t])$	42
N_c	cluster centre process	176
$N(\cdot \mid x)$	cluster member or component process	176
$\{(p_n, \Pi_n)\}$	elements of probability measure for finite point process	123
$P(z)$	probability generating function (p.g.f.) of distribution $\{p_n\}$	10, 115
$P(x,A)$	Markov transition kernel	92
$P_0(A)$	avoidance function	31, 135
$\mathcal{P}_{jk}$	set of j-partitions of $\{1,\ldots,k\}$	121
$\mathcal{P}$	probability measure of stationary N on $\mathbb{R}$,	53
	probability measure of N or ξ on c.s.m.s. $\mathcal{X}$	158
$\{\pi_k\}$	batch-size distribution	28, 51
$q(x) = f(x)/[1-F(x)]$	hazard function for lifetime d.f. F	2, 106
$Q(z)$	$= -\log P(z)$	27
$Q(\cdot)$, $Q(t)$	hazard measure, integrated hazard function (IHF)	109
$\rho(x,y)$	metric for x,y in metric space	370
$\{S_n\}$	random walk, sequence of partial sums	66
$S(x) = 1 - F(x)$	survivor function of d.f. F	2, 109
$S_r(x)$	sphere of radius r, centre x, in metric space $\mathcal{X}$	35, 371
$t(x) = \prod_{i=1}^{d}(1-\lvert x_i\rvert)_+$	triangular density in $\mathbb{R}^d$	359
T_u	forward recurrence time at u	58, 75
$\mathcal{T} = \{S_1(\mathcal{T}),\ldots,S_j(\mathcal{T})\}$	a j-partition of k	121
$\mathcal{T} = \{\mathcal{T}_n\} = \{\{A_{ni}\}\}$	dissecting system of nested partitions	382
$U(A) = \mathrm{E}[N(A)]$	renewal measure	67
$U(x)$	$= U([0,x])$, expectation function,	61
	renewal function $(U(x) = 1 + U_0(x))$	67
$V(A)$	$= \operatorname{var} N(A)$, variance function	295
$V(x) = V((0,x])$	variance function for stationary N or ξ on $\mathbb{R}$	80, 301
$\{X_n\}$	components of random walk $\{S_n\}$,	66
	intervals of Wold process	92

Concordance of Statements from the First Edition

The table below lists the identifying number of formal statements of the first edition (1988) of this book and their identification in this volume.

1988 edition	this volume	1988 edition	this volume
2.2.I–III	2.2.I–III	8.1.II	6.1.II, IV
2.3.III	2.3.I	8.2.I	6.3.I
2.4.I–II	2.4.I–II	8.2.II	6.3.II, (6.3.6)
2.4.V–VIII	2.4.III–VI	8.3.I–III	6.3.III–V
3.2.I–II	3.2.I–II	8.5.I–III	6.2.II
3.3.I–IX	3.3.I–IX	11.1.I–V	8.6.I–V
3.4.I–II	3.4.I–II	11.2.I–II	8.2.I–II
3.5.I–III	3.5.I–III	11.3.I–VIII	8.4.I–VIII
3.6.I–V	3.6.I–V	11.4.I–IV	8.5.I–IV
4.2.I–II	4.2.I–II	11.4.V–VI	8.5.VI–VII
4.3.I–III	4.3.I–III	13.1.I–III	7.1.I–III
4.4.I–VI	4.4.I–VI	13.1.IV–VI	7.2.I–III
4.5.I–VI	4.5.I–VI	13.1.VII	7.1.IV
4.6.I–V	4.6.I–V	13.4.III	7.6.I
5.2.I–VII	5.2.I–VII	A1.1.I–5.IV	A1.1.I–5.IV
5.3.I–III	5.3.I–III	A2.1.I–III	A2.1.I–III
5.4.I–III	5.4.I–III	A2.1.IV	A1.6.I
5.4.IV–VI	5.4.V–VII	A2.1.V–VI	A2.1.IV–V
5.5.I	5.5.I	A2.2.I–7.III	A2.2.I–7.III
7.1.XII–XIII	6.4.I(a)–(b)	A3.1.I–4.IX	A3.1.I–4.IX

CHAPTER 1

Early History

The ancient origins of the modern theory of point processes are not easy to trace, nor is it our aim to give here an account with claims to being definitive. But any retrospective survey of a subject must inevitably give some focus on those past activities that can be seen to embody concepts in common with the modern theory. Accordingly, this first chapter is a historical indulgence but with the added benefit of describing certain fundamental concepts informally and in a heuristic fashion prior to possibly obscuring them with a plethora of mathematical jargon and techniques. These essentially simple ideas appear to have emerged from four distinguishable strands of enquiry—although our division of material may sometimes be a little arbitrary. These are

(i) life tables and the theory of self-renewing aggregates;
(ii) counting problems;
(iii) particle physics and population processes; and
(iv) communication engineering.

The first two of these strands could have been discerned in centuries past and are discussed in the first two sections. The remaining two essentially belong to the twentieth century, and our comments are briefer in the remaining section.

1.1. Life Tables and Renewal Theory

Of all the threads that are woven into the modern theory of point processes, the one with the longest history is that associated with intervals between events. This includes, in particular, renewal theory, which could be defined in a narrow sense as the study of the sequence of intervals between successive replacements of a component that is liable to failure and is replaced by a new component every time a failure occurs. As such, it is a subject that developed during the 1930s and reached a definitive stage with the work of Feller, Smith, and others in the period following World War II. But its roots extend back much further than this, through the study of 'self-renewing aggregates' to problems of statistical demography, insurance, and mortality tables—in short, to one of the founding impulses of probability theory itself. It is not easy to point with confidence to any intermediate stage in this chronicle that recommends itself as the natural starting point either of renewal theory or of point process theory more generally. Accordingly, we start from the beginning, with a brief discussion of life tables themselves. The connection with point processes may seem distant at first sight, but in fact the theory of life tables provides not only the source of much current terminology but also the setting for a range of problems concerning the evolution of populations in time and space, which, in their full complexity, are only now coming within the scope of current mathematical techniques.

In its basic form, a life table consists of a list of the number of individuals, usually from an initial group of 1000 individuals so that the numbers are effectively proportions, who survive to a given age in a given population. The most important parameters are the number ℓ_x surviving to age x, the number d_x dying between the ages x and $x+1$ ($d_x = \ell_x - \ell_{x+1}$), and the number q_x of those surviving to age x who die before reaching age $x+1$ ($q_x = d_x/\ell_x$). In practice, the tables are given for discrete ages, with the unit of time usually taken as 1 year. For our purposes, it is more appropriate to replace the discrete time parameter by a continuous one and to replace numbers by probabilities for a single individual. Corresponding to ℓ_x we have then the survivor function

$$S(x) = \Pr\{\text{lifetime} > x\}.$$

To d_x corresponds $f(x)$, the density of the lifetime distribution function, where

$$f(x)\,\mathrm{d}x = \Pr\{\text{lifetime terminates between } x \text{ and } x + \mathrm{d}x\},$$

while to q_x corresponds $q(x)$, the *hazard function*, where

$$q(x)\,\mathrm{d}x = \Pr\{\text{lifetime terminates between } x \text{ and } x + \mathrm{d}x \mid \text{it does not terminate before } x.\}$$

Denoting the lifetime distribution function itself by $F(x)$, we have the following important relations between the functions above:

$$S(x) = 1 - F(x) = \int_x^\infty f(y)\,\mathrm{d}y = \exp\Big(-\int_0^x q(y)\,\mathrm{d}y\Big), \tag{1.1.1}$$

$$f(x) = \frac{\mathrm{d}F}{\mathrm{d}x} = \frac{\mathrm{d}S}{\mathrm{d}x}, \tag{1.1.2}$$

$$q(x) = \frac{f(x)}{S(x)} = \frac{\mathrm{d}}{\mathrm{d}x}[\log S(x)] = -\frac{\mathrm{d}}{\mathrm{d}x}\{\log[1 - F(x)]\}. \tag{1.1.3}$$

The first life table appeared, in a rather crude form, in John Graunt's (1662) *Observations on the London Bills of Mortality*. This work is a landmark in the early history of statistics, much as the famous correspondence between Pascal and Fermat, which took place in 1654 but was not published until 1679, is a landmark in the early history of formal probability. The coincidence in dates lends weight to the thesis (see e.g. Maistrov, 1967) that mathematical scholars studied games of chance not only for their own interest but for the opportunity they gave for clarifying the basic notions of chance, frequency, and expectation, already actively in use in mortality, insurance, and population movement contexts.

An improved life table was constructed in 1693 by the astronomer Halley, using data from the smaller city of Breslau, which was not subject to the same problems of disease, immigration, and incomplete records with which Graunt struggled in the London data. Graunt's table was also discussed by Huyghens (1629–1695), to whom the notion of expected length of life is due. A. de Moivre (1667–1754) suggested that for human populations the function $S(x)$ could be taken to decrease with equal yearly decrements between the ages 22 and 86. This corresponds to a uniform density over this period and a hazard function that increases to infinity as x approaches 86. The analysis leading to (1.1.1) and (1.1.2), with further elaborations to take into account different sources of mortality, would appear to be due to Laplace (1747–1829). It is interesting that in *A Philosophical Essay on Probabilities* (1814), where the classical definition of probability based on equiprobable events is laid down, Laplace gave a discussion of mortality tables in terms of probabilities of a totally different kind. Euler (1707–1783) also studied a variety of problems of statistical demography.

From the mathematical point of view, the paradigm distribution function for lifetimes is the exponential function, which has a constant hazard independent of age: for $x > 0$, we have

$$f(x) = \lambda \mathrm{e}^{-\lambda x}, \quad q(x) = \lambda, \quad S(x) = \mathrm{e}^{-\lambda x}, \quad F(x) = 1 - \mathrm{e}^{-\lambda x}. \tag{1.1.4}$$

The usefulness of this distribution, particularly as an approximation for purposes of interpolation, was stressed by Gompertz (1779–1865), who also suggested, as a closer approximation, the distribution function corresponding to a power-law hazard of the form

$$q(x) = A\mathrm{e}^{\alpha x} \qquad (A > 0,\ \alpha > 0,\ x > 0). \tag{1.1.5}$$

With the addition of a further constant [i.e. $q(x) = B + A\mathrm{e}^{\alpha x}$], this is known in demography as the *Gompertz–Makeham* law and is possibly still the most widely used function for interpolating or graduating a life table.

Other forms commonly used for modelling the lifetime distribution in different contexts are the *Weibull*, *gamma*, and *log normal* distributions, corresponding, respectively, to the formulae

$$q(x) = \beta\lambda x^{\beta-1} \qquad \text{with } \ S(x) = \exp(-\lambda x^{\beta}) \qquad (\lambda > 0,\ \beta > 0), \tag{1.1.6}$$