



# PROBABILITY AND STOCHASTIC PROCESSES

IONUT FLORESCU

WILEY



# **PROBABILITY AND STOCHASTIC PROCESSES**



---

# PROBABILITY AND STOCHASTIC PROCESSES

---

**Ionuț Florescu**

Stevens Institute of Technology

WILEY

Copyright ©2015 by John Wiley & Sons, Inc. All rights reserved.

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

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 646-8600, or on the web at [www.copyright.com](http://www.copyright.com). Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008.

**Limit of Liability/Disclaimer of Warranty:** While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services please contact our Customer Care Department with the U.S. at 877-762-2974, outside the U.S. at 317-572-3993 or fax 317-572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print, however, may not be available in electronic format.

***Library of Congress Cataloging-in-Publication Data:***

Florescu, Ionuț, 1973–  
Probability and stochastic processes / Ionuț Florescu, Stevens Institute of Technology, Hoboken, NJ.  
pages cm  
Includes bibliographical references and index.  
ISBN 978-0-470-62455-5 (cloth)  
1. Probabilities. 2. Stochastic processes. I. Title.  
QA273.F655 2015  
519.2–dc23

2015018455

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

*To M.G. and C.*





# CONTENTS IN BRIEF

---

## PART I PROBABILITY

1	Elements of Probability Measure	9
2	Random Variables	45
3	Applied Chapter: Generating Random Variables	87
4	Integration Theory	123
5	Conditional Distribution and Conditional Expectation	157
6	Moment Generating Function. Characteristic Function.	181
7	Limit Theorems	213
8	Statistical Inference	259

## PART II STOCHASTIC PROCESSES

9	Introduction to Stochastic Processes	293
10	The Poisson process	307
11	Renewal Processes	331
12	Markov Chains	371
13	Semi-Markov and Continuous-time Markov Processes	411

<b>14</b>	<b>Martingales</b>	<b>437</b>
<b>15</b>	<b>Brownian Motion</b>	<b>465</b>
<b>16</b>	<b>Stochastic Differential Equations</b>	<b>485</b>

# CONTENTS

---

List of Figures	xvii
List of Tables	xx
Preface	xxi
Acknowledgments	xxiii
Introduction	1

## PART I PROBABILITY

<b>1</b>	<b>Elements of Probability Measure</b>	<b>9</b>
1.1	Probability Spaces	10
1.1.1	Null element of $\mathcal{F}$ . Almost sure (a.s.) statements.	
	Indicator of a set	21
1.2	Conditional Probability	22
1.3	Independence	29
1.4	Monotone Convergence Properties of Probability	31
1.5	Lebesgue Measure on the Unit Interval (0,1]	37
	Problems	40
		<b>ix</b>

<b>2</b>	<b>Random Variables</b>	<b>45</b>
2.1	Discrete and Continuous Random Variables	48
2.2	Examples of Commonly Encountered Random Variables	52
2.3	Existence of Random Variables with Prescribed Distribution	65
2.4	Independence	68
2.5	Functions of Random Variables. Calculating Distributions	72
	Problems	82
<b>3</b>	<b>Applied Chapter: Generating Random Variables</b>	<b>87</b>
3.1	Generating One-Dimensional Random Variables by Inverting the cdf	88
3.2	Generating One-Dimensional Normal Random Variables	91
3.3	Generating Random Variables. Rejection Sampling Method	94
3.4	Generating Random Variables. Importance Sampling	109
	Problems	119
<b>4</b>	<b>Integration Theory</b>	<b>123</b>
4.1	Integral of Measurable Functions	124
4.2	Expectations	130
4.3	Moments of a Random Variable. Variance and the Correlation Coefficient	143
4.4	Functions of Random Variables. The Transport Formula	145
4.5	Applications. Exercises in Probability Reasoning	148
4.6	A Basic Central Limit Theorem: The <i>DeMoivre–Laplace Theorem</i> :	150
	Problems	152
<b>5</b>	<b>Conditional Distribution and Conditional Expectation</b>	<b>157</b>
5.1	Product Spaces	158
5.2	Conditional Distribution and Expectation. Calculation in Simple Cases	162
5.3	Conditional Expectation. General Definition	165
5.4	Random Vectors. Moments and Distributions	168
	Problems	177
<b>6</b>	<b>Moment Generating Function. Characteristic Function</b>	<b>181</b>
6.1	Sums of Random Variables. Convolutions	181
6.2	Generating Functions and Applications	182
6.3	Moment Generating Function	188
6.4	Characteristic Function	192
6.5	Inversion and Continuity Theorems	199
6.6	Stable Distributions. Lévy Distribution	204
6.6.1	Truncated Lévy flight distribution	206
	Problems	208

<b>7</b>	<b>Limit Theorems</b>	<b>213</b>
7.1	Types of Convergence	213
7.1.1	Traditional deterministic convergence types	214
7.1.2	Convergence in $L^p$	215
7.1.3	Almost sure (a.s.) convergence	216
7.1.4	Convergence in probability. Convergence in distribution	217
7.2	Relationships between Types of Convergence	221
7.2.1	A.S. and $L^p$	221
7.2.2	Probability, a.s., $L^p$ convergence	223
7.2.3	Uniform Integrability	226
7.2.4	Weak convergence and all the others	228
7.3	Continuous Mapping Theorem. Joint Convergence. Slutsky's Theorem	230
7.4	The Two Big Limit Theorems: LLN and CLT	232
7.4.1	A note on statistics	232
7.4.2	The order statistics	234
7.4.3	Limit theorems for the mean statistics	238
7.5	Extensions of CLT	245
7.6	Exchanging the Order of Limits and Expectations	251
	Problems	252
<b>8</b>	<b>Statistical Inference</b>	<b>259</b>
8.1	The Classical Problems in Statistics	259
8.2	Parameter Estimation Problem	260
8.2.1	The case of the normal distribution, estimating mean when variance is unknown	262
8.2.2	The case of the normal distribution, comparing variances	264
8.3	Maximum Likelihood Estimation Method	265
8.3.1	The bisection method	267
8.4	The Method of Moments	276
8.5	Testing, the Likelihood Ratio Test	277
8.5.1	The likelihood ratio test	280
8.6	Confidence Sets	284
	Problems	286

## PART II STOCHASTIC PROCESSES

<b>9</b>	<b>Introduction to Stochastic Processes</b>	<b>293</b>
9.1	General Characteristics of Stochastic Processes	294
9.1.1	The index set $\mathcal{I}$	294
9.1.2	The state space $\mathcal{S}$	294
9.1.3	Adaptiveness, filtration, standard filtration	294

9.1.4	Pathwise realizations	296
9.1.5	The finite distribution of stochastic processes	296
9.1.6	Independent components	297
9.1.7	Stationary process	298
9.1.8	Stationary and independent increments	299
9.1.9	Other properties that characterize specific classes of stochastic processes	300
9.2	A Simple Process – The Bernoulli Process	301
	Problems	304
<b>10</b>	<b>The Poisson Process</b>	<b>307</b>
10.1	Definitions	307
10.2	Inter-Arrival and Waiting Time for a Poisson Process	310
10.2.1	Proving that the inter-arrival times are independent	311
10.2.2	Memoryless property of the exponential distribution	315
10.2.3	Merging two independent Poisson processes	316
10.2.4	Splitting the events of the Poisson process into types	316
10.3	General Poisson Processes	317
10.3.1	Nonhomogenous Poisson process	318
10.3.2	The compound Poisson process	319
10.4	Simulation techniques. Constructing Poisson Processes	323
10.4.1	One-dimensional simple Poisson process	323
	Problems	326
<b>11</b>	<b>Renewal Processes</b>	<b>331</b>
11.0.2	The renewal function	333
11.1	Limit Theorems for the Renewal Process	334
11.1.1	Auxiliary but very important results. Wald's theorem. Discrete stopping time	336
11.1.2	An alternative proof of the elementary renewal theorem	340
11.2	Discrete Renewal Theory	344
11.3	The Key Renewal Theorem	349
11.4	Applications of the Renewal Theorems	350
11.5	Special cases of renewal processes	352
11.5.1	The alternating renewal process	353
11.5.2	Renewal reward process	358
11.6	The renewal Equation	359
11.7	Age-Dependent Branching processes	363
	Problems	366
<b>12</b>	<b>Markov Chains</b>	<b>371</b>
12.1	Basic Concepts for Markov Chains	371
12.1.1	Definition	371

12.1.2	Examples of Markov chains	372
12.1.3	The Chapman– Kolmogorov equation	378
12.1.4	Communicating classes and class properties	379
12.1.5	Periodicity	379
12.1.6	Recurrence property	380
12.1.7	Types of recurrence	382
12.2	Simple Random Walk on Integers in $d$ Dimensions	383
12.3	Limit Theorems	386
12.4	States in a MC. Stationary Distribution	387
12.4.1	Examples. Calculating stationary distribution	391
12.5	Other Issues: Graphs, First-Step Analysis	394
12.5.1	First-step analysis	394
12.5.2	Markov chains and graphs	395
12.6	A general Treatment of the Markov Chains	396
12.6.1	Time of absorption	399
12.6.2	An example	400
	Problems	406
<b>13</b>	<b>Semi-Markov and Continuous-time Markov Processes</b>	<b>411</b>
13.1	Characterization Theorems for the General semi- Markov Process	413
13.2	Continuous-Time Markov Processes	417
13.3	The Kolmogorov Differential Equations	420
13.4	Calculating Transition Probabilities for a Markov Process. General Approach	425
13.5	Limiting Probabilities for the Continuous-Time Markov Chain	426
13.6	Reversible Markov Process	429
	Problems	432
<b>14</b>	<b>Martingales</b>	<b>437</b>
14.1	Definition and Examples	438
14.1.1	Examples of martingales	439
14.2	Martingales and Markov Chains	440
14.2.1	Martingales induced by Markov chains	440
14.3	Previsible Process. The Martingale Transform	442
14.4	Stopping Time. Stopped Process	444
14.4.1	Properties of stopping time	446
14.5	Classical Examples of Martingale Reasoning	449
14.5.1	The expected number of tosses until a binary pattern occurs	449
14.5.2	Expected number of attempts until a general pattern occurs	451
14.5.3	Gambler's ruin probability – revisited	452

14.6	Convergence Theorems. $L^1$ Convergence. Bounded Martingales in $L^2$	456
	Problems	458
<b>15</b>	<b>Brownian Motion</b>	<b>465</b>
15.1	History	465
15.2	Definition	467
15.2.1	Brownian motion as a Gaussian process	469
15.3	Properties of Brownian Motion	471
15.3.1	Hitting times. Reflection principle. Maximum value	474
15.3.2	Quadratic variation	476
15.4	Simulating Brownian Motions	480
15.4.1	Generating a Brownian motion path	480
15.4.2	Estimating parameters for a Brownian motion with drift	481
	Problems	481
<b>16</b>	<b>Stochastic Differential Equations</b>	<b>485</b>
16.1	The Construction of the Stochastic Integral	487
16.1.1	Itô integral construction	490
16.1.2	An illustrative example	492
16.2	Properties of the Stochastic Integral	494
16.3	Itô lemma	495
16.4	Stochastic Differential Equations (SDEs)	499
16.4.1	A discussion of the types of solution for an SDE	501
16.5	Examples of SDEs	502
16.5.1	An analysis of Cox– Ingersoll– Ross (CIR) type models	507
16.5.2	Models similar to CIR	507
16.5.3	Moments calculation for the CIR model	509
16.5.4	Interpretation of the formulas for moments	511
16.5.5	Parameter estimation for the CIR model	511
16.6	Linear Systems of SDEs	513
16.7	A Simple Relationship between SDEs and Partial Differential Equations (PDEs)	515
16.8	Monte Carlo Simulations of SDEs	517
	Problems	522
<b>A</b>	<b>Appendix: Linear Algebra and Solving Difference Equations and Systems of Differential Equations</b>	<b>527</b>
A.1	Solving difference equations with constant coefficients	528
A.2	Generalized matrix inverse and pseudo-determinant	528
A.3	Connection between systems of differential equations and matrices	529



A.3.1	Writing a system of differential equations in matrix form	530
A.4	Linear Algebra results	533
A.4.1	Eigenvalues, eigenvectors of a square matrix	533
A.4.2	Matrix Exponential Function	534
A.4.3	Relationship between Exponential matrix and Eigenvectors	534
A.5	Finding fundamental solution of the homogeneous system	535
A.5.1	The case when all the eigenvalues are distinct and real	536
A.5.2	The case when some of the eigenvalues are complex	536
A.5.3	The case of repeated real eigenvalues	537
A.6	The nonhomogeneous system	538
A.6.1	The method of undetermined coefficients	538
A.6.2	The method of variation of parameters	539
A.7	Solving systems when $P$ is non-constant	540
	Bibliography	541
	Index	547



# LIST OF FIGURES

---

1.1	Example: Biker trajectories	24
1.2	Example: Monty Hall conditional probabilities	26
1.3	Example: Bertrand's box paradox	27
1.4	Example: Bayes rule	28
1.5	Example: A sequence with no limit	33
1.6	Example: A sequence with one limiting point	33
2.1	The pdf and cdf of the Uniform	55
2.2	<i>Exponential</i> distribution	56
2.3	<i>Normal</i> distribution	57
2.4	<i>Gamma</i> distribution	58
2.5	The $\chi_n^2$ distribution	59
2.6	Inverse Gamma distribution	60
2.7	Beta distributions for various parameter values	61

2.8	A $t$ -distribution versus a normal	62
2.9	$\text{Log} - \text{normal}$ distribution	63
2.10	A simple distribution function, yet with no density	64
2.11	Points where the two variables $X^\pm$ may have different outcomes	65
2.12	The area we need to calculate. The shaded parts need to be deleted.	68
3.1	Rejection sampling using a basic Uniform.	96
3.2	The ziggurat distribution for $n = 8$ .	97
3.3	The ziggurat distribution for $n = 256$ .	98
3.4	Generating from unknown distribution	103
3.5	Empirical histogram	105
3.6	Generating from Gamma mixture distribution	106
3.7	Empirical histogram of Gamma mixture	107
3.8	Importance sampling candidate densities	113
3.9	The evolution of different importance sampling estimators.	115
7.1	The continuous functions $f_1$ and $f_2$	220
7.2	Relations between different types of convergence	222
8.1	Log-likelihood plot for normal distribution	270
9.1	Paths for a stochastic process	297
9.2	White noise processes	298
9.3	Failure and waiting time	302
10.1	A general Poisson process	322
10.2	Illustration of the construction	323
10.3	Illustration of set $A$ and $B_i$	325
11.1	Relationship between $S_{N(t)}$ , $t$ , and $S_{N(t)+1}$	335
11.2	Representation of the new process	342
11.3	Deduction of the formula	351
11.4	Relationship between $X_{n+1}$ and $Y(t)$	356

12.1	A graphical description of the Ehrenfest chain	373
12.2	A graphical description of the Gambler's ruin problem	375
12.3	A graphical description of the birth and death chain	375
12.4	A realization of the simple queue system	377
12.5	A simple Markov Chain	379
12.6	Simple chains where $\sum_{n=0}^{\infty} P_{ij}^n = \infty \nleftrightarrow f_{ij} = 1$	382
12.7	Random walk in two dimensions	384
14.1	A $Z_t^T$ process which stops once it reaches 0.6	445
15.1	Reflected Brownian motion	474

# LIST OF TABLES

---

3.1	Running times for rejection sampling and mixture generating algorithms	109
8.1	Commonly encountered cases of pivotal quantities	284
9.1	Sample Outcome	301
11.1	Cost evolution with block value $K$	346
12.1	Transition probability for crossbreeding with genotype BB	402
12.2	Transition probability for crossbreeding with genotype Bb	402
12.3	Transition probability for crossbreeding with genotype bb	403
12.4	Transition probability for crossbreeding with red-eyed flies	404
12.5	Possible genotypes of parents	405
12.6	Transition probability for the interbreeding model	405

# PREFACE

---

This book originated as a series of lecture notes for a first year graduate class taught during two semesters at Stevens Institute of Technology. It covers probability, which is taught during the first semester, and stochastic processes, taught in the second semester. Thus the book is structured to cover both subject in a wide enough manner to allow applications to many domains.

Probability is an old subject. Stochastic processes is a new subject that is quickly becoming old. So why write a book on an old subject? In my opinion, this book is necessary at the current day and age. The fundamental textbooks are becoming too complex for the new students and, in an effort to make the material more accessible to students, new applied probability books discard the rigor of the old books and the painstaking details that is put forward in these old textbooks. At times, reading these new books feels like the authors are inventing new notions to be able to skip the old reasoning. I believe that this is not needed. I believe that it is possible to have a mathematically rigorous textbook which is at the same time accessible to students. The result is this work. This book does not try to reinvent the concepts only to put them into an accessible format. Throughout, I have tried to explain complex notions with as many details as possible. For this reason to a versed reader in the subject, many of the derivations will seem to contain unnecessary details. Let me assure you that for a student seeing the concepts for the first time these derivations are vital.

This textbook is not a replacement for the fundamental textbooks. Many results are not proven and, for a deeper understanding of each of the subjects, the reader is

advised to delve deeper into these fundamental textbooks. However, in my opinion this textbook contains all the material needed to start research in probability, complete a qualifying exam in probability and stochastic processes, or make sound probability reasoning for applied problems.

I. FLORESCU

*Hoboken, New Jersey*  
*April, 2014*



# ACKNOWLEDGMENTS

---

I want to acknowledge all my teachers of Mathematics and Probability. They are the ones who made this book possible. These great people that taught me probability are listed in the order in which they taught me: *Nicolae Gheorghilă, Valentin Nicula, Gabriela Licea, Ioan Cuculescu, Constantin Tudor, Vasile Preda, Thomas Selke, Ioannis Kontoyiannis, Rebecca Doerge, Burgess Davis, Tony Cai, Herman Rubin, Phillip Protter*, and of course my advisor *Frederi Viens*.

I would also like to acknowledge all my students who helped improve this book by providing feedback and corrections, among them Lazer Teran, Pojen Huan, Chris Flynn, and Brad Warren.

I. F.



# INTRODUCTION

---

What is Probability? In essence:

Mathematical modeling of random events and phenomena. It is fundamentally different from modeling deterministic events and functions, which constitutes the traditional study of Mathematics.

However, the study of probability uses concepts and notions straight from Mathematics; in fact Measure Theory and Potential Theory are expressions of abstract mathematics generalizing the Theory of Probability.

Like so many other branches of mathematics, the development of probability theory has been stimulated by the variety of its applications. In turn, each advance in the theory has enlarged the scope of its influence. Mathematical statistics is one important branch of applied probability; other applications occur in such widely different fields as genetics, biology, psychology, economics, finance, engineering, mechanics, optics, thermodynamics, quantum mechanics, computer vision, geophysics, etc. In fact I compel the reader to find one area in today's science where no applications of probability theory can be found.

## Early history

In the XVII-th century the first notions of Probability Theory appeared. More precisely, in 1654 Antoine Gombaud Chevalier de Méré, a French nobleman with an interest in gaming and gambling questions, was puzzled by an apparent contradiction

concerning a popular dice game. The game consisted of throwing a pair of dice 24 times; the problem was to decide whether or not to bet even money on the occurrence of at least one "double six" during the 24 throws. A seemingly well-established gambling rule led de Méré to believe that betting on a double six in 24 throws would be profitable (based on the payoff of the game). However, his own calculations based on many repetitions of the 24 throws indicated just the opposite. Using modern probability language de Méré was trying to establish if such an event has probability greater than 0.5 (we are looking at this question in example 1.7). Puzzled by this and other similar gambling problems he called on the famous mathematician Blaise Pascal. This, in turn led to an exchange of letters between Pascal and another famous French mathematician Pierre de Fermat. This is the first known documentation of the fundamental principles of the theory of probability. Before this famous exchange of letters, a few other simple problems on games of chance had been solved in the XV-th and XVI-th centuries by Italian mathematicians; however, no general principles had been formulated before this famous correspondence.

In 1655 during his first visit to Paris, the Dutch scientist Christian Huygens learned of the work on probability carried out in this correspondence. On his return to Holland in 1657, Huygens wrote a small work *De Ratiociniis in Ludo Aleae*, the first printed work on the calculus of probabilities. It was a treatise on problems associated with gambling. Because of the inherent appeal of games of chance, probability theory soon became popular, and the subject developed rapidly during the XVIII-th century.

## The XVIII-th century

The major contributors during this period were Jacob Bernoulli (1654–1705) and Abraham de Moivre (1667–1754). Jacob (Jacques) Bernoulli was a Swiss mathematician who was the first to use the term integral. He was the first mathematician in the Bernoulli family, a family of famous scientists of the XVIII-th century. Jacob Bernoulli's most original work was *Ars Conjectandi* published in Basel in 1713, eight years after his death. The work was incomplete at the time of his death but it still was a work of the greatest significance in the development of the Theory of Probability. De Moivre was a French mathematician who lived most of his life in England<sup>1</sup>. De Moivre pioneered the modern approach to the Theory of Probability, in his work *The Doctrine of Chance: A Method of Calculating the Probabilities of Events in Play* in the year 1718. A Latin version of the book had been presented to the Royal Society and published in the *Philosophical Transactions* in 1711. The definition of statistical independence appears in this book for the first time. *The Doctrine of Chance* appeared in new expanded editions in 1718, 1738 and 1756. The birthday problem (example 1.12) appeared in the 1738 edition, the gambler's ruin problem (example 1.11) in the 1756 edition. The 1756 edition of *The Doctrine of Chance* contained what is probably de Moivre's most significant contribution to probability, namely the approximation of the binomial distribution with the normal distribution in the case of a large number of trials - which is now known by most probability textbooks as "The First Central Limit Theorem" (we will discuss this theorem in Chapter 4). He understood the notion of

<sup>1</sup> A protestant, he was pushed to leave France after Louis XIV revoked the Edict of Nantes in 1685, leading to the expulsion of the Huguenots

standard deviation and is the first to write the normal integral (and the distribution density). In *Miscellanea Analytica* (1730) he derives Stirling's formula (wrongly attributed to Stirling) which he uses in his proof of the central limit theorem. In the second edition of the book in 1738 de Moivre gives credit to Stirling for an improvement to the formula. De Moivre wrote:

"I desisted in proceeding farther till my worthy and learned friend Mr James Stirling, who had applied after me to that inquiry, [discovered that  $c = \sqrt{2}$ ]."

De Moivre also investigated mortality statistics and the foundation of the theory of annuities. In 1724 he published one of the first statistical applications to finance *Annuities on Lives*, based on population data for the city of Breslau. In fact, in *A History of the Mathematical Theory of Probability* (London, 1865), Isaac Todhunter says that probability:

... owes more to [de Moivre] than any other mathematician, with the single exception of Laplace.

De Moivre died in poverty. He did not hold a university position despite his influential friends Leibnitz, Newton, and Halley, and his main income came from tutoring.

De Moivre, like Cardan (Girolamo Cardano), predicted the day of his own death. He discovered that he was sleeping 15 minutes longer each night and summing the arithmetic progression, calculated that he would die on the day when he slept for 24 hours. He was right!

## The XIX-th century

This century saw the development and generalization of the early Probability Theory. Pierre-Simon de Laplace (1749–1827) published *Théorie Analytique des Probabilités* in 1812. This is the first fundamental book in probability ever published (the second being Kolmogorov's 1933 monograph). Before Laplace, probability theory was solely concerned with developing a mathematical analysis of games of chance. The first edition was dedicated to Napoleon-le-Grand, but the dedication was removed in later editions!<sup>2</sup>

The work consisted of two books and a second edition two years later saw an increase in the material by about 30 per cent. The work studies generating functions, Laplace's definition of probability, Bayes rule (so named by Poincaré many years later), the notion of mathematical expectation, probability approximations, a discussion of the method of least squares, Buffon's needle problem, and inverse Laplace transform. Later editions of the *Théorie Analytique des Probabilités* also contains supplements which consider applications of probability to determine errors in observations arising in astronomy, the other passion of Laplace.

On the morning of Monday 5 March 1827, Laplace died. Few events would cause the Academy to cancel a meeting but they did so on that day as a mark of respect for one of the greatest scientists of all time.

<sup>2</sup>The close relationship between Laplace and Napoleon is well documented and he became Count of the Empire in 1806. However, when it was clear that royalists were coming back he offered his services to the Bourbons and in 1817 he was rewarded with the title of marquis.

## Century XX and modern times

Many scientists have contributed to the theory since Laplace's time; among the most important are Chebyshev, Markov, von Mises, and Kolmogorov.

One of the difficulties in developing a mathematical theory of probability has been to arrive at a definition of probability that is precise enough for use in mathematics, yet comprehensive enough to be applicable to a wide range of phenomena. The search for a widely acceptable definition took nearly three centuries and was marked by much controversy. The matter was finally resolved in the 20th century by treating probability theory on an axiomatic basis. In 1933, a monograph by the Russian *giant mathematician* Andrey Nikolaevich Kolmogorov (1903–1987) outlined an axiomatic approach that forms the basis for the modern theory. In 1925, the year he started his doctoral studies, Kolmogorov published his first paper with Khinchin on the probability theory. The paper contains, among other inequalities about partial series of random variables, the three series theorem which provides important tools for stochastic calculus. In 1929, when he finished his doctorate, he already had published 18 papers. Among them were versions of the strong law of large numbers and the law of iterated logarithm.

In 1933, two years after his appointment as a professor at Moscow University, Kolmogorov published *Grundbegriffe der Wahrscheinlichkeitsrechnung* his most fundamental book. In it he builds up probability theory in a rigorous way from fundamental axioms in a way comparable with Euclid's treatment of geometry. He gives a rigorous definition of the conditional expectation which later became fundamental for the definition of Brownian motion, stochastic integration, and Mathematics of Finance. (Kolmogorov's monograph is available in English translation as *Foundations of Probability Theory*, Chelsea, New York, 1950). In 1938 he publishes the paper *Analytic methods in probability theory* which lay the foundation for the Markov processes, leading toward a more rigorous approach to the Markov chains.

Kolmogorov later extended his work to study the motion of the planets and the turbulent flow of air from a jet engine. In 1941 he published two papers on turbulence which are of fundamental importance in the field of fluid mechanics. In 1953–54 two papers by Kolmogorov, each of four pages in length, appeared. These are on the theory of dynamical systems with applications to Hamiltonian dynamics. These papers mark the beginning of KAM-theory, which is named after Kolmogorov, Arnold and Moser. Kolmogorov addressed the International Congress of Mathematicians in Amsterdam in 1954 on this topic with his important talk *General Theory of Dynamical Systems and Classical Mechanics*. He thus demonstrated the vital role of probability theory in physics. His contribution in the topology theory is also of outmost importance<sup>3</sup>.

Closer to the modern era, I have to mention Joseph Leo Doob (1910–2004), who was one of the pioneers in the modern treatment of stochastic processes. His book *Stochastic Processes* (Doob, 1953) is one of the most influential in the treatment of modern stochastic processes (specifically martingales). Paul-André Meyer

<sup>3</sup>Kolmogorov had many interests outside mathematics, for example he was interested in the form and structure of the poetry of the greatest Russian poet Alexander Sergeyevich Pushkin (1799–1837).