

Christiane Fuchs

# Inference for Diffusion Processes

With Applications in Life Sciences

扩散过程推断理论  
及其在生命科学中的应用

Springer

世界图书出版公司  
[www.wpcbj.com.cn](http://www.wpcbj.com.cn)

Christiane Fuchs

# Inference for Diffusion Processes

With Applications in Life Sciences

 Springer

Christiane Fuchs  
Institute for Bioinformatics and Systems Biology  
Leibniz Center Hannover  
Hannover  
Germany

ISBN 978-3-642-25062-2  
DOI 10.1007/978-3-642-25062-2  
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012930004

© Springer-Verlag Berlin Heidelberg 2013

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recording, broadcasting, reproduction on microfilm or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through Rightslink in the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Reprint from English language edition

Inference for Diffusion Processes With Applications in Life Sciences

by Christiane Fuchs

Copyright © Springer Berlin Heidelberg 2013

Springer Berlin Heidelberg is a part of Springer Science+Business Media

All Rights Reserved

This reprint has been authorized by Springer Science & Business Media  
China Mainland only and not for export therefrom.

## 图书在版编目 ( CIP ) 数据

扩散过程推断理论及其在生命科学中的应用 = Inference for Diffusion

Processes : With Applications in Life Sciences : 英文 / ( 德 ) 福克斯 ( Fuchs, C. ) 著 .

—影印本 . —北京 : 世界图书出版公司北京公司 , 2016. 4

ISBN 978-7-5192-1168-4

I . ①扩… II . ①福… III . ①扩散过程—应用—生命科学—研究—英文  
IV . ① O211.6 ② Q1-0

中国版本图书馆 CIP 数据核字 ( 2016 ) 第 083175 号

著 者 : Christiane Fuchs

责任编辑 : 刘 慧 高 蓉

装帧设计 : 任志远

出版发行 : 世界图书出版公司北京公司

地 址 : 北京市东城区朝内大街 137 号

邮 编 : 100010

电 话 : 010-64038355 ( 发行 ) 64015580 ( 客服 ) 64033507 ( 总编室 )

网 址 : <http://www.wpcbj.com.cn>

邮 箱 : [wpcbjst@vip.163.com](mailto:wpcbjst@vip.163.com)

销 售 : 新华书店

印 刷 : 三河市国英印务有限公司

开 本 : 711mm × 1245 mm 1/24

印 张 : 19

字 数 : 364 千

版 次 : 2016 年 5 月第 1 版 2016 年 5 月第 1 次印刷

版权登记 : 01-2016-0566

定 价 : 99.00 元

---

版权所有 翻印必究

( 如发现印装质量问题 , 请与所购图书销售部门联系调换 )

Christiane Fuchs  
Institute for Bioinformatics and Systems Biology  
Helmholtz Zentrum München  
Neuherberg  
Germany

ISBN 978-3-642-25968-5 ISBN 978-3-642-25969-2 (eBook)

DOI 10.1007/978-3-642-25969-2

Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012950004

© Springer-Verlag Berlin Heidelberg 2013

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Reprint from English language edition:

Inference for Diffusion Processes: With Applications in Life Sciences

by Christiane Fuchs

Copyright © Springer Berlin Heidelberg 2013

Springer Berlin Heidelberg is a part of Springer Science+Business Media

All Rights Reserved

This reprint has been authorized by Springer Science & Business Media for distribution in China Mainland only and not for export therefrom.

Christiane Eich  
**Inference for Diffusion Processes**

# Inference for Diffusion Processes

With Applications in Life Sciences

 Springer

## Foreword

Beginning with Brownian motion, modifications, diffusion processes and stochastic differential equations have a long tradition as mathematical models for stochastic phenomena evolving in continuous time, with applications in diverse substantive fields such as engineering, finance and life sciences. While many textbooks, in particular in financial mathematics, include concise and rigorous introductions to diffusion processes and stochastic calculus, they often provide only limited information on modern statistical inference and usually neglect the question of how to adequately approximate original phenomena through a diffusion process at all. These two issues are, however, of fundamental importance in a number of applied fields, in particular in modern life sciences: First, the original processes typically have a huge but discrete state space, and approximations through diffusion processes can be challenging. Second, data are usually observed with low frequency and often at non-equidistant time points, involve measurement errors etc. This requires modern tools such as simulation based Bayesian inference. Motivated through applications in epidemiology and molecular biology, this book closes the existing gap. For the first time, it provides a unified presentation of the approximation techniques, previously often developed for special cases only, and a thorough account of modern statistical inference, including a powerful Bayesian approach. Together with the two application chapters, this book will be of high value for theoretical and applied work. For me, it was a pleasure to watch it grow, to be able to give some advice, to read the final version, and, at last, to see it on my bookshelf!

Münster, Germany

Ludwig Palmrodt

# Foreword

Beginning with Brownian motion and its modifications, diffusion processes and stochastic differential equations have a long tradition as mathematical models for stochastic phenomena evolving in continuous time, with applications in diverse substantive fields such as engineering, finance and life sciences. While many textbooks, in particular in financial mathematics, include concise and rigorous introductions to diffusion processes and stochastic calculus, they often provide only limited information on modern statistical inference and usually neglect the question of how to adequately approximate original phenomena through a diffusion process at all. These two issues are, however, of fundamental importance in a number of applied fields, in particular in modern life sciences: First, the original processes typically have a large but discrete state space, and approximations through diffusion processes can be challenging. Second, data are usually observed with low frequency and often at non-equidistant time points, involve measurement error etc. This requires modern tools such as simulation-based Bayesian inference. Motivated through applications in epidemiology and molecular biology, this book closes the existing gap. For the first time, it provides a unified presentation of the approximation techniques, previously often developed for special cases only, and a thorough account of modern statistical inference, including a powerful Bayesian approach. Together with the two application chapters, this book will be of high value for theoretical and applied work. For me, it was a pleasure to watch it grow, to be able to give some advice, to read the final version, and, at last, to see it on my bookshelf!

Munich, Germany

Ludwig Fahrmeir

# Preface

This book originates from my doctoral thesis. One of the first problems of my work was to describe the spread of an infectious disease by a diffusion process and to statistically estimate the involved model parameters. At that time, I did not expect such a seemingly straightforward task to surface so many diverse open problems to fill an entire book. As a mathematician by training, I knew about stochastic calculus, but I did not anticipate the troubles of deriving diffusion approximations and inferring their parameters from real data.

When delving into the diffusion approximation literature, I noticed that there were several, at first sight, contradicting approaches, some of them formulated in generality, others being carried out for particular problems. Their appropriateness, differences and conformities, however, were unclear as well as their extension to more complex, e.g. multidimensional, processes. Furthermore, parameter estimation for diffusions is a challenging problem, in particular if the application of interest involves multi-dimensional processes, few observation times, latent variables and considerable measurement error. Under these circumstances, probably the only applicable technique is a popular Bayesian approach which is used in a number of scientific papers. I was astonished that I could not find any textbook which comprehensively explained it. Moreover, the method has a well-known but hard-to-grasp convergence problem, which has not been detailed in any book or review so far. Since I am convinced that these are subjects of wide-spread interest and importance, I dedicated to them the major chapter in each of the first two parts of this book. The third part finally addresses the initial project which triggered the theoretical questions: to estimate a diffusion model for the spread of diseases.

In contrast to existing literature, this book treats modelling and inference for diffusions under one umbrella. It thus covers both steps that necessarily arise in a real application. Importance is attached to presenting the methods both comprehensibly and mathematically well-founded. As such, the book addresses both theoreticians, like mathematicians and statisticians, as well as practitioners, like bioinformaticians and biologists. The reader is required to have basic knowledge about deterministic differential equations, probability theory and statistics. An introduction to stochastic calculus, in particular to diffusions, is provided in this book.

Everybody who supported me during the writing of my thesis “Bayesian Inference for Diffusion Processes with Applications in Life Sciences”, submitted in 2010 at Ludwig-Maximilians-Universität Munich under my maiden name Christiane Dargatz, also supported the making of this book. My sincere gratitude is due to my supervisors Ludwig Fahrmeir and Gareth Roberts. They enriched my work through their advice, ideas and encouragement. I deeply appreciate the careful proof-reading and helpful comments by Michael Höhle. Katrin Schneider and Lothar Schermelleh deserve my thanks for having initiated the collaboration on the FRAP project and having collected all the data. I thank my former and present colleagues at Ludwig-Maximilians-Universität Munich and Helmholtz Centre Munich for their interest and helpful discussions. Furthermore, I thank Niels Thomas and Alice Blanck from Springer for the friendly and constructive cooperation. My family has been a constant source of support, which I greatly acknowledge. My heartfelt gratitude is due to my husband Florian, who caringly accompanied me and my work all chapters long.

Munich, Germany

Christiane Fuchs

# Notation

## Symbols

|                                                                                                                                             |                                                                                                                                                                                                  |
|---------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| $\mathbb{N} = \{1, 2, 3, \dots\}$                                                                                                           | the natural numbers                                                                                                                                                                              |
| $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$                                                                                                      | the non-negative whole numbers                                                                                                                                                                   |
| $\mathbb{Z}$                                                                                                                                | the whole numbers                                                                                                                                                                                |
| $\mathbb{Q}$                                                                                                                                | the rational numbers                                                                                                                                                                             |
| $\mathbb{R}$                                                                                                                                | the real numbers                                                                                                                                                                                 |
| $\mathbb{R}_+$                                                                                                                              | the strictly positive real numbers                                                                                                                                                               |
| $\mathbb{R}_0 = \mathbb{R}_+ \cup \{0\}$                                                                                                    | the non-negative real numbers                                                                                                                                                                    |
| $x$                                                                                                                                         | a column vector                                                                                                                                                                                  |
| $x'$                                                                                                                                        | the transpose of $x$                                                                                                                                                                             |
| $\mathbf{0} = (0, \dots, 0)'$                                                                                                               | the null vector                                                                                                                                                                                  |
| $e_i = (0, \dots, 1, \dots, 0)'$                                                                                                            | the $i$ th unit vector                                                                                                                                                                           |
| $I$                                                                                                                                         | the identity matrix                                                                                                                                                                              |
| $\text{diag}(\mathbf{A})$                                                                                                                   | the main diagonal of the quadratic matrix $\mathbf{A}$                                                                                                                                           |
| $\text{diag}(\mathbf{a})$                                                                                                                   | quadratic matrix with main diagonal $\mathbf{a}$ and zero entries otherwise                                                                                                                      |
| $ a $                                                                                                                                       | the absolute value of $a \in \mathbb{R}$                                                                                                                                                         |
| $\ \mathbf{A}\ $                                                                                                                            | Euclidean distance, i. e. $\ \mathbf{A}\ ^2 = \text{tr}(\mathbf{A}'\mathbf{A})$ for a vector or matrix $\mathbf{A}$                                                                              |
| $ \mathbf{k}  = \sum_{i=1}^n k_i$ , where $\mathbf{k} = (k_1, \dots, k_n)'$ (in the context of Sect. B.1)                                   |                                                                                                                                                                                                  |
| $\mathbf{u} \diamond \mathbf{v} = (u_1v_1, \dots, u_nv_n)'$ , where $\mathbf{u} = (u_1, \dots, u_n)'$ and $\mathbf{v} = (v_1, \dots, v_n)'$ |                                                                                                                                                                                                  |
| $(\Omega, \mathcal{F}^*, \mathcal{F}, \mathbb{P})$                                                                                          | a filtered probability space with sample space $\Omega$ , $\sigma$ -algebra $\mathcal{F}^*$ , natural filtration $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$ and probability measure $\mathbb{P}$ |
| $\mathcal{L}$                                                                                                                               | the $\sigma$ -algebra of Lebesgue subsets of $\mathbb{R}$                                                                                                                                        |

|                                                             |                                                                                                        |
|-------------------------------------------------------------|--------------------------------------------------------------------------------------------------------|
| $\mathbb{L}$                                                | Lebesgue measure                                                                                       |
| $\mathbb{W}$                                                | Wiener measure                                                                                         |
| $\mathbb{P}_1 \ll \mathbb{P}_2$                             | the measure $\mathbb{P}_1$ is absolutely continuous with respect to $\mathbb{P}_2$                     |
| $\mathbb{P}_1 \perp \mathbb{P}_2$                           | the measures $\mathbb{P}_1$ and $\mathbb{P}_2$ are mutually singular, i. e. they have disjoint support |
| $\mathbb{P}_1 \otimes \mathbb{P}_2$                         | factorisation of measures                                                                              |
| $\mathfrak{L}(X)$                                           | the distribution of the random variable $X$                                                            |
| $\mathbb{1}(A)$                                             | indicator function; equal to one if $A$ is true and zero otherwise                                     |
| $\delta(x - y)$                                             | Dirac delta function; equal to $\infty$ if $x = y$ and zero otherwise                                  |
| $a = o(h) \Leftrightarrow \lim_{h \rightarrow 0} (a/h) = 0$ |                                                                                                        |
| $\biguplus_{i=1}^n A_i$                                     | disjoint union of sets $A_1, \dots, A_n$                                                               |
| $\Gamma$                                                    | Gamma function, defined as $\Gamma(x) = \int_0^\infty t^{x-1} \exp(-t) dt$                             |
| $d/dt$                                                      | total derivative, i. e.                                                                                |

$$\frac{d}{dt} f(t, \xi_1, \dots, \xi_m) = \frac{\partial f(t, \xi_1, \dots, \xi_m)}{\partial t} + \sum_{j=1}^n \frac{\partial f(t, \xi_1, \dots, \xi_m)}{\partial \xi_j} \frac{d\xi_j}{dt}$$

$\partial/\partial t$  partial derivative, where all arguments but  $t$  remain constant

Abbreviations

|        |                                         |
|--------|-----------------------------------------|
| SDE    | stochastic differential equation        |
| ODE    | ordinary differential equation          |
| i.i.d. | independent and identically distributed |
| a.s.   | almost surely                           |

Distributions

|                               |                                                                                                                                                                                                                                                                                                                                                   |
|-------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Normal distribution           | $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $\boldsymbol{\mu} \in \mathbb{R}^n$ , $\boldsymbol{\Sigma} \in \mathbb{R}^{n \times n}$ symmetric and positive definite                                                                                                                                                 |
| Log-normal distribution       | $\mathbf{X} \sim \mathcal{LN}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \Leftrightarrow \log(\mathbf{X}) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$                                                                                                                                                                                    |
| Truncated normal distribution | $\mathbf{X} \sim \mathcal{N}_{\text{trunc}}(\boldsymbol{\mu}, \sigma^2)$ ; generates random numbers from $\mathcal{N}(\boldsymbol{\mu}, \sigma^2)$ restricted to the positive real line, i. e. the density $f(x)$ of this distribution is proportional to the density of $\mathcal{N}(\boldsymbol{\mu}, \sigma^2)$ for $x > 0$ and zero otherwise |
| Multivariate t distribution   | $\mathbf{X} \sim t_\nu(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $\nu \in \mathbb{R}_+$ degrees of freedom, $\boldsymbol{\mu} \in \mathbb{R}^n$ , $\boldsymbol{\Sigma} \in \mathbb{R}^{n \times n}$ symmetric and positive definite and density                                                                                                |

$$f(\mathbf{x}) = \frac{\Gamma\left(\frac{\nu+n}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)(\nu\pi)^{n/2}} |\Sigma|^{-\frac{1}{2}} \left(1 + \frac{1}{\nu}(\mathbf{x} - \boldsymbol{\mu})' \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu})\right)^{-\frac{\nu+n}{2}} \text{ for } \mathbf{x} \in \mathbb{R}^n$$

|                          |                                                                   |
|--------------------------|-------------------------------------------------------------------|
| Uniform distribution     | $X \sim U(A)$ for a discrete or continuous set $A$                |
| Poisson distribution     | $X \sim \text{Po}(\lambda)$ with $\lambda \in \mathbb{R}_+$       |
| Exponential distribution | $X \sim \text{Exp}(\lambda)$ with $\lambda \in \mathbb{R}_+$      |
| Gamma distribution       | $X \sim \text{Ga}(a, b)$ with $a, b \in \mathbb{R}_+$ and density |

$$f(x) = \frac{b^a}{\Gamma(a)} x^{a-1} \exp(-bx) \text{ for } x \in \mathbb{R}_+$$

|                            |                                                                   |
|----------------------------|-------------------------------------------------------------------|
| Inverse gamma distribution | $X \sim \text{IG}(a, b) \Leftrightarrow 1/X \sim \text{Ga}(a, b)$ |
|----------------------------|-------------------------------------------------------------------|

anyone from molecular processes at molecular level up to the worldwide spread of infectious diseases in humans. Mathematical models are an indispensable tool for the understanding of such complex natural phenomena.

In order to describe the time continuous evolution of a given system, deterministic models are often favoured as they allow comparatively simple simulation and analytical techniques. Such models, however, do not capture the randomness of the underlying dynamics and therefore turn out to be inadequate in many applications. Determination of exact individual based stochastic models, on the other hand, appears to be infeasible in practice when the considered organism involves large numbers of objects. A natural and powerful compromise is the application of stochastic differential equations (SDEs), whose solutions are given by diffusion processes. Hence, diffusions have become an increasingly important tool for the statistical analysis of real world phenomena.

However, approximation of a given dynamic system in effect done heuristically in the literature, leading to diffusions that do not correctly mirror the true dynamics of the original process. Furthermore, the statistical inference for diffusions typically turns out to be demanding in real data situations as described below. Hence, the statistical estimation of complex diffusion models as applied to real datasets is not widely spread. These issues are addressed in the present book.

## 1.1 Aims of This Book

The main objectives of this book are threefold. First of all, given a dynamical system of interest with the aim to describe its temporal evolution by means of a diffusion process, one needs to construct this process such that it adequately captures the characteristics of the considered real phenomenon. In applications in life sciences, the original process typically concerns either numbers of objects such as the numbers of infectious individuals in a population or the numbers of patients in a

# Contents

|                                    |                                                                              |           |
|------------------------------------|------------------------------------------------------------------------------|-----------|
| <b>1</b>                           | <b>Introduction</b> .....                                                    | <b>1</b>  |
| 1.1                                | Aims of This Book .....                                                      | 1         |
| 1.2                                | Outline of This Book .....                                                   | 3         |
| <b>Part I Stochastic Modelling</b> |                                                                              |           |
| <b>2</b>                           | <b>Stochastic Modelling in Life Sciences</b> .....                           | <b>9</b>  |
| 2.1                                | Compartment Models .....                                                     | 10        |
| 2.2                                | Modelling the Spread of Infectious Diseases .....                            | 11        |
| 2.2.1                              | History of Epidemic Modelling .....                                          | 12        |
| 2.2.2                              | SIR Model .....                                                              | 13        |
| 2.2.3                              | Model Extensions .....                                                       | 18        |
| 2.3                                | Modelling Processes in Molecular Biology,<br>Biochemistry and Genetics ..... | 19        |
| 2.3.1                              | History of Chemical Reaction Modelling .....                                 | 20        |
| 2.3.2                              | Chemical Reaction Kinetics .....                                             | 20        |
| 2.3.3                              | Reaction Kinetics in the Biological Sciences .....                           | 23        |
| 2.4                                | Algorithms for Simulation .....                                              | 25        |
| 2.4.1                              | Simulation of Continuous-Time Markov Jump Processes .....                    | 25        |
| 2.4.2                              | Simulation of Solutions of ODEs and SDEs .....                               | 26        |
| 2.5                                | Conclusion .....                                                             | 27        |
|                                    | References .....                                                             | 28        |
| <b>3</b>                           | <b>Stochastic Differential Equations and Diffusions in a Nutshell</b> .....  | <b>31</b> |
| 3.1                                | Brownian Motion and Gaussian White Noise .....                               | 32        |
| 3.1.1                              | Brownian Motion .....                                                        | 32        |
| 3.1.2                              | Brownian Bridge .....                                                        | 34        |
| 3.1.3                              | Gaussian White Noise .....                                                   | 34        |
| 3.1.4                              | Excursus: Lévy Processes .....                                               | 35        |

|          |                                                                   |            |
|----------|-------------------------------------------------------------------|------------|
| 3.2      | Itô Calculus .....                                                | 35         |
| 3.2.1    | Stochastic Integral and Stochastic Differential Equations .....   | 36         |
| 3.2.2    | Different Stochastic Integrals .....                              | 37         |
| 3.2.3    | Existence and Uniqueness of Solutions .....                       | 38         |
| 3.2.4    | Transition Density and Likelihood Function .....                  | 39         |
| 3.2.5    | Itô Diffusion Processes .....                                     | 40         |
| 3.2.6    | Sample Path Properties .....                                      | 41         |
| 3.2.7    | Ergodicity .....                                                  | 41         |
| 3.2.8    | Kolmogorov Forward and Backward Equations .....                   | 42         |
| 3.2.9    | Infinitesimal Generator .....                                     | 43         |
| 3.2.10   | Itô Formula.....                                                  | 44         |
| 3.2.11   | Lamperti Transformation.....                                      | 45         |
| 3.2.12   | Girsanov Formula.....                                             | 45         |
| 3.3      | Approximation and Simulation .....                                | 46         |
| 3.3.1    | Convergence and Consistency .....                                 | 47         |
| 3.3.2    | Numerical Approximation .....                                     | 47         |
| 3.3.3    | Simulation of Brownian Bridge.....                                | 50         |
| 3.4      | Concluding Remarks .....                                          | 50         |
|          | References .....                                                  | 51         |
| <b>4</b> | <b>Approximation of Markov Jump Processes by Diffusions .....</b> | <b>55</b>  |
| 4.1      | Characterisation of Processes .....                               | 56         |
| 4.2      | Motivation and Purpose .....                                      | 61         |
| 4.3      | Approximation Methods .....                                       | 63         |
| 4.3.1    | Convergence of the Master Equation .....                          | 65         |
| 4.3.2    | Convergence of the Infinitesimal Generator .....                  | 69         |
| 4.3.3    | Langevin Approach.....                                            | 71         |
| 4.3.4    | Kramers-Moyal Expansion.....                                      | 75         |
| 4.3.5    | Van Kampen Expansion .....                                        | 77         |
| 4.3.6    | Other Approaches .....                                            | 83         |
| 4.4      | Extensions to Systems with Multiple Size Parameters .....         | 83         |
| 4.4.1    | Convergence of the Master Equation .....                          | 84         |
| 4.4.2    | Convergence of the Infinitesimal Generator .....                  | 85         |
| 4.4.3    | Langevin Approach.....                                            | 86         |
| 4.4.4    | Kramers-Moyal Expansion.....                                      | 87         |
| 4.4.5    | Van Kampen Expansion .....                                        | 88         |
| 4.5      | Choice of Stochastic Integral .....                               | 91         |
| 4.6      | Discussion and Conclusion .....                                   | 92         |
|          | References.....                                                   | 97         |
| <b>5</b> | <b>Diffusion Models in Life Sciences .....</b>                    | <b>101</b> |
| 5.1      | Standard SIR Model.....                                           | 102        |
| 5.1.1    | Model .....                                                       | 102        |
| 5.1.2    | Jump Process.....                                                 | 103        |

|       |                                             |     |
|-------|---------------------------------------------|-----|
| 5.1.3 | Diffusion Approximation .....               | 104 |
| 5.1.4 | Summary .....                               | 109 |
| 5.1.5 | Illustration .....                          | 110 |
| 5.2   | Multitype SIR Model .....                   | 111 |
| 5.2.1 | Model .....                                 | 111 |
| 5.2.2 | Jump Process .....                          | 113 |
| 5.2.3 | Diffusion Approximation .....               | 115 |
| 5.2.4 | Summary .....                               | 121 |
| 5.2.5 | Illustration and Further Remarks .....      | 123 |
| 5.3   | Existence and Uniqueness of Solutions ..... | 125 |
| 5.4   | Conclusion .....                            | 127 |
|       | References .....                            | 127 |

## Part II Statistical Inference

|       |                                                                                |     |
|-------|--------------------------------------------------------------------------------|-----|
| 6     | <b>Parametric Inference for Discretely-Observed Diffusions</b> .....           | 133 |
| 6.1   | Preliminaries .....                                                            | 134 |
| 6.1.1 | Time-Continuous Observation .....                                              | 135 |
| 6.1.2 | Time-Discrete Observation .....                                                | 136 |
| 6.1.3 | Time Scheme .....                                                              | 137 |
| 6.2   | Naive Maximum Likelihood Approach .....                                        | 138 |
| 6.3   | Approximation of the Likelihood Function .....                                 | 139 |
| 6.3.1 | Analytical Approximation of the Likelihood Function ...                        | 139 |
| 6.3.2 | Numerical Solutions of the Kolmogorov<br>Forward Equation .....                | 142 |
| 6.3.3 | Simulated Maximum Likelihood Estimation .....                                  | 145 |
| 6.3.4 | Local Linearisation .....                                                      | 149 |
| 6.4   | Alternatives to Maximum Likelihood Estimation .....                            | 150 |
| 6.4.1 | Estimating Functions .....                                                     | 150 |
| 6.4.2 | Generalised Method of Moments .....                                            | 156 |
| 6.4.3 | Simulated Moments Estimation .....                                             | 157 |
| 6.4.4 | Indirect Inference .....                                                       | 158 |
| 6.4.5 | Efficient Method of Moments .....                                              | 160 |
| 6.5   | Exact Algorithm .....                                                          | 160 |
| 6.6   | Discussion and Conclusion .....                                                | 165 |
|       | References .....                                                               | 167 |
| 7     | <b>Bayesian Inference for Diffusions with Low-Frequency Observations</b> ..... | 171 |
| 7.1   | Concepts of Bayesian Data Augmentation for Diffusions .....                    | 172 |
| 7.1.1 | Preliminaries and Notation .....                                               | 173 |
| 7.1.2 | Path Update .....                                                              | 175 |
| 7.1.3 | Parameter Update .....                                                         | 185 |
| 7.1.4 | Generalisation to Several Observation Times .....                              | 188 |
| 7.1.5 | Generalisation to Several Observed Diffusion Paths .....                       | 189 |
| 7.1.6 | Practical Concerns .....                                                       | 190 |

|       |                                                                 |     |
|-------|-----------------------------------------------------------------|-----|
| 7.1.7 | Example: Ornstein-Uhlenbeck Process .....                       | 192 |
| 7.1.8 | Discussion .....                                                | 197 |
| 7.2   | Extension to Latent Data and Observation with Error .....       | 217 |
| 7.2.1 | Latent Data .....                                               | 217 |
| 7.2.2 | Observation with Error .....                                    | 225 |
| 7.3   | Convergence Problems .....                                      | 229 |
| 7.4   | Improvements of Convergence .....                               | 233 |
| 7.4.1 | Changing the Factorisation of the Dominating Measure ..         | 233 |
| 7.4.2 | Time Change Transformations .....                               | 235 |
| 7.4.3 | Particle Filters .....                                          | 238 |
| 7.4.4 | Innovation Scheme on Infinite-Dimensional<br>State Spaces ..... | 239 |
| 7.5   | Discussion and Conclusion .....                                 | 272 |
|       | References .....                                                | 277 |

### Part III Applications

|          |                                                                         |     |
|----------|-------------------------------------------------------------------------|-----|
| <b>8</b> | <b>Application I: Spread of Influenza</b> .....                         | 281 |
| 8.1      | Simulation Study .....                                                  | 281 |
| 8.1.1    | Data .....                                                              | 282 |
| 8.1.2    | Parameter Estimation .....                                              | 284 |
| 8.2      | Example: Influenza in a Boarding School .....                           | 287 |
| 8.2.1    | Data .....                                                              | 288 |
| 8.2.2    | Parameter Estimation .....                                              | 288 |
| 8.3      | Example: Influenza in Germany .....                                     | 294 |
| 8.3.1    | Data .....                                                              | 294 |
| 8.3.2    | Parameter Estimation .....                                              | 299 |
| 8.4      | Conclusion and Outlook .....                                            | 301 |
|          | References .....                                                        | 302 |
| <b>9</b> | <b>Application II: Analysis of Molecular Binding</b> .....              | 305 |
| 9.1      | Problem Statement .....                                                 | 306 |
| 9.1.1    | Data Acquisition by Fluorescence Recovery<br>After Photobleaching ..... | 306 |
| 9.1.2    | Research Questions .....                                                | 307 |
| 9.2      | Preliminary Analysis .....                                              | 309 |
| 9.2.1    | Impact of Binding .....                                                 | 309 |
| 9.2.2    | Impact of Diffusion .....                                               | 310 |
| 9.3      | General Model .....                                                     | 311 |
| 9.3.1    | Compartmental Description .....                                         | 311 |
| 9.3.2    | Diffusion Approximation .....                                           | 314 |
| 9.3.3    | Deterministic Approximation .....                                       | 318 |
| 9.3.4    | Simulation Study .....                                                  | 319 |
| 9.4      | Refinement of the General Model .....                                   | 325 |
| 9.4.1    | Compartmental Description .....                                         | 326 |

|       |                                                                  |     |
|-------|------------------------------------------------------------------|-----|
| 9.4.2 | Diffusion Approximation .....                                    | 326 |
| 9.4.3 | Deterministic Approximation .....                                | 328 |
| 9.4.4 | Simulation Study .....                                           | 328 |
| 9.5   | Extension of the General Model to Multiple Mobility Classes .... | 330 |
| 9.5.1 | Compartmental Description .....                                  | 331 |
| 9.5.2 | Diffusion Approximation .....                                    | 333 |
| 9.5.3 | Deterministic Approximation .....                                | 335 |
| 9.5.4 | Simulation Study .....                                           | 336 |
| 9.6   | Data Preparation .....                                           | 347 |
| 9.6.1 | Triple Normalisation .....                                       | 348 |
| 9.6.2 | Double Normalisation .....                                       | 352 |
| 9.6.3 | Single Normalisation .....                                       | 353 |
| 9.7   | Application .....                                                | 355 |
| 9.7.1 | Data .....                                                       | 356 |
| 9.7.2 | Bayesian Estimation .....                                        | 356 |
| 9.7.3 | Least Squares Estimation .....                                   | 359 |
| 9.7.4 | Conclusion .....                                                 | 364 |
| 9.8   | Diffusion-Coupled FRAP .....                                     | 364 |
| 9.9   | Conclusion and Outlook .....                                     | 367 |
|       | References .....                                                 | 368 |

## Conclusion and Appendix

|       |                                                               |     |
|-------|---------------------------------------------------------------|-----|
| 10    | Summary and Future Work .....                                 | 371 |
| 10.1  | Summary .....                                                 | 371 |
| 10.2  | Future Work .....                                             | 372 |
| A     | Benchmark Models .....                                        | 375 |
| A.1   | Geometric Brownian Motion .....                               | 375 |
| A.2   | Ornstein-Uhlenbeck Process .....                              | 376 |
| A.3   | Cox-Ingersoll-Ross Process .....                              | 376 |
|       | References .....                                              | 377 |
| B     | Miscellaneous .....                                           | 379 |
| B.1   | Difference Operators .....                                    | 379 |
| B.2   | Lipschitz Continuity for SIR Models .....                     | 385 |
| B.2.1 | Standard SIR Model .....                                      | 385 |
| B.2.2 | Multitype SIR Model .....                                     | 386 |
| B.3   | On the Choice of the Update Interval .....                    | 387 |
| B.4   | Posteriori Densities for the Ornstein-Uhlenbeck Process ..... | 388 |
| B.5   | Inefficiency Factors .....                                    | 397 |
| B.6   | Path Proposals in the Latent Data Framework .....             | 398 |
| B.7   | Derivation of Radon-Nikodym Derivatives .....                 | 402 |
| B.8   | Derivation of Acceptance Probability .....                    | 409 |
|       | References .....                                              | 409 |