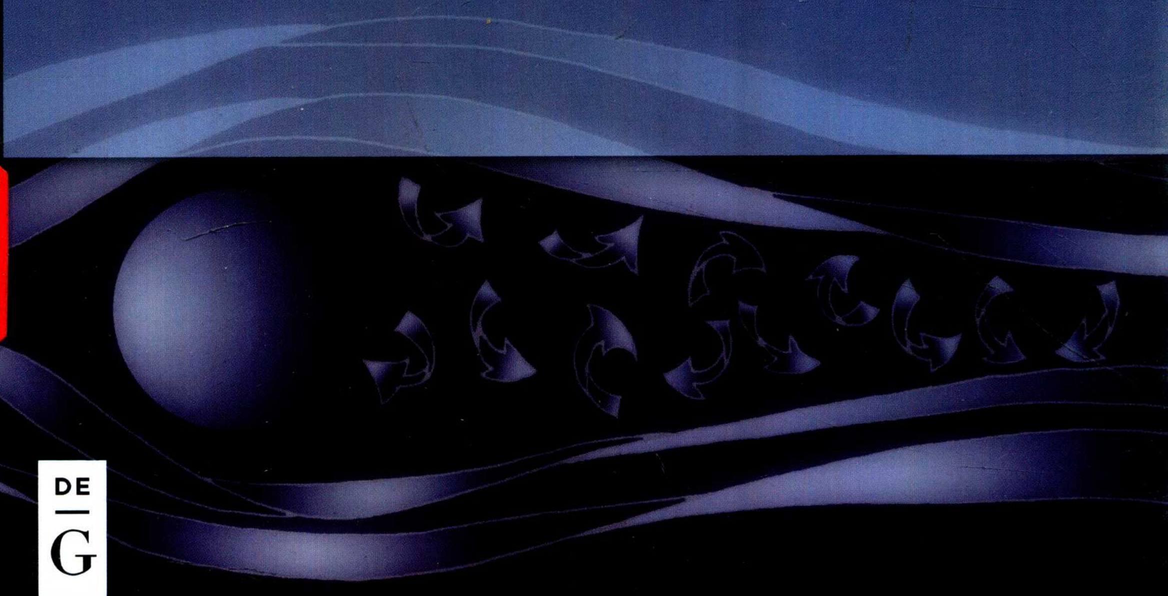
DE GRUYTER

Boling Guo, Hongjun Gao, Xueke Pu

STOCHASTIC PDES AND DYNAMICS



DE G

Boling Guo, Hongjun Gao, Xueke Pu

Stochastic PDEs and Dynamics

DE GRUYTER

Authors
Prof. Boling Guo
Laboratory of Computational Physics
Institute of Applied Physics & Computational Mathematics
Huayuan Road 6
100088 Beijing
China

Prof. Hongjun Gao
Nanjing Normal University
Institute of Mathematics, School of Mathematics Sciences
Wenyuan Road 1
210023 Nanjing
China

Dr. Xueke Pu
Chongqing University
College of Mathematics and Statistics
Shazheng Street 174
404100 Chongqing
China

ISBN 978-3-11-049510-2
e-ISBN (PDF) 978-3-11-049388-7
e-ISBN (EPUB) 978-3-11-049243-9
Set-ISBN 978-3-11-049389-4

Library of Congress Cataloging-in-Publication Data
A CIP catalog record for this book has been applied for at the Library of Congress.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.

© 2017 Walter de Gruyter GmbH, Berlin/Boston
Typesetting: Integra Software Services Pvt. Ltd.
Cover image: Chong Guo, guochong200805@163.com
Printing and binding: CPI books GmbH, Leck
♾ Printed on acid-free paper
Printed in Germany

MIX
Papier aus verantwortungsvollen Quellen
FSC www.fsc.org FSC® C083411

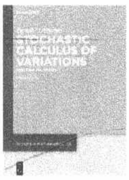
www.degruyter.com

Boling Guo, Hongjun Gao, Xueke Pu

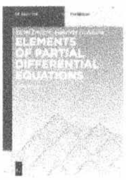
Stochastic PDEs and Dynamics

Also of interest

Stochastic Methods for Boundary Value Problems
Karl K. Sabelfeld, Nikolai A. Simonov, 2016
ISBN 978-3-11-047906-5, e-ISBN 978-3-11-047916-4

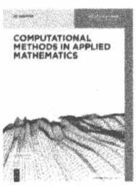


Stochastic Calculus of Variations. For Jump Processes
Yasushi Ishikawa, 2016
ISBN 978-3-11-037776-7, e-ISBN 978-3-11-039232-6

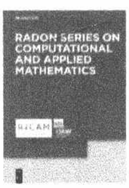


Elements of Partial Differential Equations
Pavel Drábek, Gabriela Holubová, 2014
ISBN 978-3-11-031665-0, e-ISBN 978-3-11-037404-9

Journal of Numerical Mathematics
Ronald H. W. Hoppe, Yuri Kuznetsov (Editor-in-Chief)
ISSN 1570–2820, e-ISSN 1569–3953



Computational Methods in Applied Mathematics
Carsten Carstensen (Editor-in-Chief)
ISSN 1609–4840, e-ISSN 1609–9389



Radon Series on Computational and Applied Mathematics
Ulrich Langer, Hansjörg Albrecher, Heinz W. Engl, Ronald H. W. Hoppe, Karl Kunisch, Harald Niederreiter, Christian Schmeiser (Eds.)
ISSN 1865–3707

此为试读，需要完整PDF请访问：www.ertongbook.com

Preface

The past two decades has witnessed the emergence of stochastic nonlinear partial differential equations (SPDEs) and their dynamics in different fields such as physics, mechanics, finance and biology etc. For example, there are corresponding SPDEs descriptions for the atmosphere-oceanic circulation, for the nonlinear wave propagation in random media, for the pricing of risky assets, and the law of fluctuation of stock market prices. In early 1970s, mathematicians such as Bensoussan, Temam, Pardoux, just to name only a few, initiated the mathematical studies of SPDEs and the studies of corresponding random dynamical systems began slightly later afterwards. In the middle 90's, Crauel, Da Prato, Debussche, Flandoli, Schmalfuss, Zabczyk *et al* established the framework of random attractors, Hausdorff dimension estimates and Invariant measure theory of random dynamic systems with applications to stochastic nonlinear evolutionary equations. Recently, there are theoretical and numerical aspects of nonlinear SPDEs has been developed, resulting in many fruitful achievement and subsequently, many monographs were published.

The authors of this book had been working in the fields of nonlinear SPDEs and random dynamics as well as stochastic processes such as Lévy process and fractional Wiener process for more than a decade. Seminars were held and discussions had been going with scholars all over the world since then. Interesting and preliminary results were made on some mathematical problems in climate, ocean circulation and propagation of nonlinear waves in random media.

The aim of book is twofold. First, to give some preliminaries that are of importance to SPDEs. Second, to introduce latest recent results concerning several important SPDEs such as Ginzburg-Landau equation, Ostrovsky equation, geostrophic equations and primitive equations in climate. Materials are presented in a concise way, hoping to bring readers into such an interesting field of modern applied mathematics.

Chapter one introduces preliminaries in probability and stochastic processes, and Chapter 2 briefly presents the stochastic integral and Ito formula, which plays a vital role in stochastic partial differential equations. Chapter 3 discusses the Ornstein-Ulenbeck process and some linear SDEs. Chapter 4 establishes the basic framework of stochastic dynamic systems. In Chapter 5, latest results on several SPDEs emerging from various physics backgrounds are given.

Last but not the least, I would like to take the opportunity to express sincere gratitudes to Dr. Mufa Chen, Member of Chinese Academy of Sciences, and Dr. Jian Wang at Fuzhou University, from whom we benefited constantly in preparing this book.

Boling Guo

August 20, 2016

Contents

1 Preliminaries

This chapter contains some preliminaries in probability and stochastic processes, especially some basic properties of the Wiener process, Poisson process and Lévy process. Because many contents in this chapter can be found in the other literature, we here only give the conclusions and omit the proofs.

1.1 Preliminaries in probability

1.1.1 Probability space

There are many uncertainties and randomness in our natural world and social environments. A lot of observations and tests are asking for the research in random phenomena. A random experiment must contain certain properties, usually requiring that (i) the experiment can be repeated arbitrarily many times under the same conditions and (ii) the outcome for the experiment may be more than one and all the possible outcomes are known, but one can't accurately predict which outcome would appear in one trial.

The random experiment is usually called test for short and is expressed as E. Each possible result in E is called a basic event or a sample point, expressed as ω. The set of all sample points in E, denoted by Ω, is called the space for basic event, and the set of sample points is called event and is expressed in capital letters $A, B, C, \cdots$. The event A occurs if and only if one of the sample points in A occurs.

Take the "roll the dice" game as a simple example. The outcome can't be predicted in the experiment when the dice was rolled once, but certainly it is one of the outcomes "point one," $\cdots$, "point six." Hence, $\Omega = \{1, 2, 3, 4, 5, 6\}$ consists of six elements, representing the six possible outcomes in the "roll the dice" experiment. In this experiment, "roll prime number point" is an event and consists of three basic events 2, 3, 5, which we denote as $A = \{2, 3, 5\}$.

In practice, various manipulations such as intersection, union or complement to subsets are needed. It is nature to ask that whether the result is still an event after such manipulations. This leads to the concept of σ-algebra.

Definition 1.1.1. *Let Ω be a sample space, then the set $\mathcal{F} = \{A : A \subset \Omega\}$ is called a* σ-algebra *if it satisfies*

(1) $\Omega \in \mathcal{F}$;

(2) *if* $A \in \mathcal{F}$, *then* $A^c := \Omega \backslash A \in \mathcal{F}$;

(3) *if* $A_i \in \mathcal{F}$, *then* $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Then $(\Omega, \mathcal{F})$ is called a measurable space and each element in $\mathcal{F}$ is measurable. Two trivial examples are $\mathcal{F} = \{\emptyset, \Omega\}$ and $\mathcal{F}$ contains all the subsets of Ω. These two examples are too special to be studied in mathematics. The first one is so small that we cannot get enough information that we are interested in, while the latter one is so big that it is difficult to define a probability measure on it. Therefore, we need to consider other intermediate σ-algebra we are interested in. For a collection $\mathcal{C}$ of subsets of Ω, we denote $\sigma(\mathcal{C})$ the σ-algebra generated by $\mathcal{C}$, that is the smallest σ-algebra containing $\mathcal{C}$.

Definition 1.1.2. *Let $(\Omega, \mathscr{F})$ be a measurable space and P be a real valued function defined on the event field $\mathscr{F}$. If P satisfies*

(1) *for each $A_i \in \mathscr{F}$, it has $P(A_i) \geq 0$;*

(2) *$P(\Omega) = 1$;*

(3) *for $A_i \in \mathscr{F}(i = 1, 2, \cdots, \infty)$ with $i \neq j$, $A_iA_j \doteq A_i \cap A_j = \varnothing$, it has*

$$P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i),$$

then P is called probability measure, and probability for short.

The above three properties are called Kolmogorov's axioms, named after Andrey Kolmogorov, one of the greatest Russian mathematicians. Such triple $(\Omega, \mathscr{F}, P)$ is called a measure space or a probability space in probability theory. In Kolmogrov's probability theory, $\mathscr{F}$ doesn't have to include all the possible subsets of Ω, but only includes the subsets we are interested in. In such a measure space, $\mathscr{F}$ is usually called an event field and the element in $\mathscr{F}$ is called an event or a measurable set. The event $A = \Omega$ is called certain event since the possibility for A to occur is $P(\Omega) = 1$ and the event $A = \emptyset$ is called impossible event accordingly since $P(\emptyset) = 0$ thanks to properties (2) and (3).

In the following, we regard Ω in $(\Omega, \mathscr{F}, P)$ as sample space, $\mathscr{F}$ as event field in Ω, and P as a determinate probability corresponding to $(\Omega, \mathscr{F})$. The properties in the definition are called non negativity, normalization, and complete additivity of probability, respectively.

We also note that for a fixed sample space Ω, many σ-algebra can be constructed (hence not unique), but not every σ-algebra is an event field. For example, let $\mathscr{F}_1 = \{\emptyset, \Omega\}$ and $\mathscr{F}_2 = \{\emptyset, A, A^c, \Omega\}$, where $A \subset \Omega$. By definition, $\mathscr{F}_1$ is certainly an event field, but $\mathscr{F}_2$ is not necessarily an event field since A is possibly not measurable under P.

After introducing the probability space, the relations and operations among events and the conditional probability can be considered. Two events A and B are called mutually exclusive, if both A, B can't occur in the same experiment (but it is possible that neither of them occurs). If any two events are exclusive, then these events are called mutually exclusive pairwise.

Theorem 1.1.1. *The probability for the sum of some of mutual exclusion events is equal to the sum of every event, i.e.,*

$$P(A_1 + A_2 + \cdots) = P(A_1) + P(A_2) + \cdots,$$

if $A_i \cap A_j = \emptyset$ *whenever* $i \neq j$.

The conditional probability measures the probability of an event under the assumption that another event has occurred. For example, we are interested in the probability of "prime number occurs" in the "roll the dice" game, under the assumption that we have known that the outcome in one trial is an odd number. We give the definition below.

Definition 1.1.3. *Let* $(\Omega, \mathscr{F}, P)$ *be a probability space,* $A, B \in \mathscr{F}$ *and* $P(B) \neq 0$. *Then the conditional probability of A given B or the probability of A under the condition B is defined by*

$$P(A|B) = P(AB)/P(B), \tag{1.1.1}$$

where $P(AB) = P(A \cap B) = P(A \text{ and } B \text{ both occur})$.

We also take the "roll the dice" as an example. Consider the event A="prime point occurs," B="odd point occurs," and C="even point occurs," that is

$$A = \{2, 3, 5\}, \quad B = \{1, 3, 5\}, \quad C = \{2, 4, 6\}.$$

It can be calculated that the (unconditional) probability of A is 1/2. Now, if the event B is assumed to have occurred, then we ask for the probability of A. By the definition, the conditional probability of A under B is $P(A|B) = P(AB)/P(B) = P(\{3, 5\})/P(B) = 2/3$. Similarly, if the event C is assumed to have occurred, or we know that C has occurred, then the conditional probability of A is $P(A|C) = 1/3$.

Another important concept in probability is independence. Consider two events A and B. Generally speaking, the conditional probability $P(A)$ of A is different from $P(A|B)$. If $P(A|B) > P(A)$, then the occurrence of B enlarges than the probability of A. Otherwise, if $P(A) = P(A|B)$, then the occurrence of B has no influence on A. In the latter case, the events A, B are said to be independent and

$$P(AB) = P(A)P(B). \tag{1.1.2}$$

Definition 1.1.4. *If* A, B *satisfy eq.* (1.1.2), *then* A, B *are said to be independent.*

Definition 1.1.5. *Let* $A_1, A_2, \cdots$ *be at most countably many events. If for any finite events* $A_{i_1}, A_{i_2}, \cdots, A_{i_m}$, *there holds*

$$P(A_{i_1}A_{i_2}\cdots A_{i_m}) = P(A_{i_1})P(A_{i_2})\cdots P(A_{i_m}), \tag{1.1.3}$$

then the events $A_1, A_2, \cdots$, *are said to be independent.*

It is noted that the events in a subset of independent events are also independent.

此为试读，需要完整PDF请访问：www.ertongbook.com

Theorem 1.1.2. *Let $A_1, A_2, \cdots, A_n$ be independent, then*

$$P(A_1A_2\cdots A_n) = P(A_1)P(A_2)\cdots P(A_n).$$

Next, we introduce the law of total probability and Bayes formula. Let $B_1, B_2, \cdots$ be at most countably many events, mutually exclusive, and at least one of them happens in an experiment. That is, $B_iB_j = \emptyset$ (impossible event), when $i \neq j$ and $B_1 + B_2 + \cdots = \Omega$ (certain event). Given any event A, noting Ω is a certain event, one gets $A = A\Omega = AB_1 + AB_2 + \cdots$, where $AB_1, AB_2, \cdots$ are mutually exclusive as $B_1, B_2, \cdots$ are mutually exclusive. Hence, by Theorem 1.1.1

$$P(A) = P(AB_1) + P(AB_2) + \cdots, \tag{1.1.4}$$

and by the definition of conditional probability, we have $P(AB_i) = P(B_i)P(A|B_i)$, which follows that

$$P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + \cdots. \tag{1.1.5}$$

This formula is called the law of total probability. By eqs (1.1.4) and (1.1.5), the probability of $P(A)$ is decomposed into the sum of many parts. It can be understood that the events B_i is a possible cause leading to A. The probability of A, under the possible cause B_i, is the conditional probability $P(A|B_i)$. Intuitively, the probability of A, $P(A)$, must be between the smallest and largest $P(A|B_i)$ under this mechanism, and also because the probabilities of $P(B_i)$ are different in all kinds of causes, the probability $P(A)$ should be a weighted average of $P(A|B_i)$, with the weight being $P(B_i)$.

Under the assumption of the law of total probability, one has

$$\begin{aligned} P(B|A) &= P(AB_i)/P(A) \\ &= P(B_i)P(A|B_i)/\sum_j P(B_j)P(A|B_j). \end{aligned} \tag{1.1.6}$$

This formula is called Bayes formula. Formally, it is just a simple deduction of the conditional probability and the law of total probability. It is famous for its explanation in reality and philosophical significance. For $P(B_i)$, it is the probability of B_i under no further information. Now, if it has new information (we know that A has occurred), then the probability of B_i has a new estimate. If the event of A is viewed as a result, and $B_1, B_2, \cdots$ are the possible causes of A, then we can formally view the law of total probability as "from the reason to result," while Bayes formula can be viewed as "from the result to reasons." In fact, a comprehensive set of statistical inference methods has been developed by the idea, which is called "Bayes statistics."

1.1.2 Random variable and probability distribution

Random variable, as the name indicates, is a variable whose value is determined randomly. Strictly speaking, given a probability space $(\Omega, \mathscr{F}, P)$, random variable X is defined as a measurable mapping from Ω to R^d. When $d \geq 2$, X is usually called a random vector, and d is the dimension. Random vectors can be divided into discrete and continuous types according to the value of random variables. The research of random variable is the content in probability, because in a random experiment, what is concerned are variables, which are usually random and are often associated with certain problems of interests.

Next, we consider the distribution of random variable.

Definition 1.1.6. *Let X be a random variable. Then the function*

$$F(x) = P(X \leq x), \quad -\infty < x < \infty, \tag{1.1.7}$$

is called the probability distribution function of X, where $P(X \leq x)$ denotes the probability of the event $\{\omega : X(\omega) \leq x\}$.

Here, it doesn't request the random variable to be discrete or continuous. It's obvious that the distribution function has the following properties: (1) $F(x)$ is a monotonically nondecreasing function, (2) $F(x) \to 0$ as $x \to -\infty$, and (3) $F(x) \to 1$ as $x \to \infty$.

First, let us consider a discrete random variable X taking possible values $a_1, a_2, \cdots$. Then $p_i = P(X = a_i)$, $i = 1, 2, \cdots$ is called the probability function of X. An important example of the discrete distribution is the Poisson distribution. If X is a non-negative integer-valued random variable with its probability function $p_i = P(X = i) = e^{-\lambda}\lambda^i/i!$, then X is said to subject to Poisson distribution, denoted by $X \sim P(\lambda)$, where $\lambda > 0$ is a constant.

For the distribution of continuous random variable, it can't be described as the discrete ones. One method to describe continuous random variable is to use distribution function and probability density function.

Definition 1.1.7. *Let $F(x)$ be the distribution function of a continuous random variable X, then the derivative $f(x) = F'(x)$ of $F(x)$, if exists, is called the probability density function of X.*

The density function $f(x)$ has the following properties:

(1) $f(x) \geq 0$.

(2) $\int_{-\infty}^{\infty} f(x)dx = 1$.

(3) For any $a < b$, there holds

$$P(a \le X \le b) = F(b) - F(a) = \int_a^b f(x)dx.$$

An important example of continuous distribution is the normal distribution, whose probability density function is

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}, \quad -\infty < x < \infty.$$

The associated random variable is usually denoted by $X \sim N(\mu, \sigma^2)$.

The above conclusions can be generalized to random vectors. Consider a d-dimensional random vector $X = (X_1, \cdots, X_n)$, whose components $X_1, \cdots, X_n$ are one-dimensional random variable. For $A \subset R^n$, $X \in A$ denotes $\{\omega : X(\omega) \in A\}$.

Definition 1.1.8. *A nonnegative function $f(x_1, \cdots, x_n)$ on R^n is said to be the probability density function of X if*

$$P(X \in A) = \int_A f(x_1, \cdots, x_n)dx_1 \cdots dx_n, \tag{1.1.8}$$

for any $A \subset R^n$.

We remark that similar to the one-dimensional case, we can introduce the probability distribution function

$$F(x_1, x_2, \cdots, x_n) = P(X_1 \le x_1, X_2 \le x_2, \cdots, X_n \le x_n),$$

for any random vector $X = (X_1, \cdots, X_n)$. For the random vector X, each competent X_i is one-dimensional random variable and has its own one-dimensional distribution function F_i, for $i = 1, \cdots, n$, which are called the "marginal distribution" of distribution F or of random vector X. It is easy to see that the marginal distribution is completely determined by the distribution F. For example, let $X = (X_1, X_2)$ with probability density function $f(x_1, x_2)$. Since $(X_1 \le x_1) = (X_1 \le x_1, X_2 < \infty)$, we have

$$F_1(x_1) = P(X_1 \le x_1) = \int_{-\infty}^{x_1} dt_1 \int_{-\infty}^{\infty} f(t_1, t_2)dt_2,$$

and the probability density function of X_1 is given by

$$f_1(x_1) := \frac{dF_1(x_1)}{dx_1} = \int_{-\infty}^{\infty} f(x_1, x_2)dx_2.$$

Similarly, in the multi dimensional case, we have for $X = (X_1, \cdots, X_n)$ that

$$f(x_1) := \frac{dF_1(x_1)}{dx_1} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, x_2, \cdots, x_n)dx_2 \cdots dx_n.$$

We recall the conditional probability, based on which independence is introduced. Now, we discuss the conditional probability distribution and the independence of random variables. We also take the continuous random variables for example. Given two random variables X_1 and X_2, we let $f(x_1, x_2)$ be the probability density function of the two-dimensional random vector $X = (X_1, X_2)$. Consider the conditional distribution of X_1 under the condition $a \le X_2 \le b$. Since

$$P(X_1 \le x_1 | a \le X_2 \le b) = P(X_1 \le x_1, a \le X_2 \le b)/P(a \le X_2 \le b),$$

by the marginal distribution function f_2 of X_2, it follows

$$P(X_1 \le x_1 | a \le X_2 \le b) = \int_{-\infty}^{x_1} dt_1 \int_a^b f(t_1, t_2)dt_2 / \int_a^b f_2(t_2)dt_2,$$

which is the conditional distribution function of X_1. The conditional density function can be obtained by derivative on x_1, i.e.,

$$f_1(X_1 | a \le X_2 \le b) = \int_a^b f(x_1, t_2)dt_2 / \int_a^b f_2(t_2)dt_2.$$

It is interesting to consider the limited case $a = b$. In this limit, we obtain

$$\begin{aligned} f_1(x_1|x_2) &= f_1(x_1|X_2 = x_2) \\ &= \lim_{h\to 0} f_1(x_1|x_2 \le X_2 \le x_2 + h) \\ &= \lim_{h\to 0} \int_{x_1}^{x_2+h} f(x_1, t_2)dt_2 / \lim_{h\to 0} \int_{x_1}^{x_2+h} f_2(t_2)dt_2 \\ &= f(x_1, x_2)/f_2(x_2). \end{aligned}$$

This is the conditional density function of X_1 under the condition $X_2 = x_2$, and we need $f_2(x_2) > 0$ such that the above equality makes sense. It can be rewritten as

$$f(x_1, x_2) = f_2(x_2) f_1(x_1|x_2), \tag{1.1.9}$$

corresponding to the conditional probability formula $P(AB) = P(A)P(B)$. In higher dimensional case, $X = (X_1, \cdots, X_n)$ with probability density function $f(x_1, \cdots, x_n)$, one has

$$f(x_1, \cdots, x_n) = g(x_1, \cdots, x_k) h(x_{k+1}, \cdots, x_n | x_1, \cdots, x_k),$$

where g is the probability density of $(X_1, \cdots, X_k)$, and h is the conditional probability density of $(X_{k+1}, \cdots, X_n)$ with the condition $X_1 = x_1, \cdots, X_k = x_k$. The formula can also be regarded as definition of the conditional probability density h. Integrating eq. (1.1.9) w.r.t. x_2, we have

$$f_1(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2)dx_2 = \int_{-\infty}^{\infty} f_1(x_1|x_2)f_2(x_2)dx_2. \tag{1.1.10}$$

Next, we discuss the independence of random variables. By the above notations, if $f_1(x_1|x_2)$ depends only on x_1 and is independent of x_2, then the distribution of X_1 is completely unrelated with the value of X_2. That is the stochastic variables X_1 and X_2 are independent in probability.

Definition 1.1.9. *Let $f(x_1, \cdots, x_n)$ be the joint probability density function of the n-dimensional random variable $X = (X_1, \cdots, X_n)$, and the marginal density functions of X_i are $f_i(x_i)$, $i = 1, \cdots, n$. If*

$$f(x_1, \cdots, x_n) = f_1(x_1) \cdots f_n(x_n),$$

then the stochastic variables $X_1, \cdots, X_n$ are mutually independent or independent for short.

The concept of independence of variables can also be considered in the following view. If $X_1, \cdots, X_n$ are independent, then the probabilities of the variables are not affected by other variables, hence the events

$$A_1 = (a_1 \leq X_1 \leq b_1), \cdots, A_n = (a_n \leq X_n \leq b_n)$$

are independent.

1.1.3 Mathematical expectation and momentum

The probability distribution of random variable we introduced above is the most complete characterization of the probability properties of random variables. Next, we consider the mathematical expectation, momentum, and related topics. Let us first consider the mathematical expectation.

Definition 1.1.10. *If X is a discrete random variable, taking countable values $a_1, a_2, \cdots$ with probability distribution $P(X = a_i) = p_i, i = 1, 2, \cdots$, and $\sum_{i=1}^{\infty} |a_i|p_i < \infty$, then $EX = \sum_{i=1}^{\infty} a_i p_i$ is called the mathematical expectation of X.*

If X is a continuous random variable with probability density function $f(x)$ and $\int_{-\infty}^{\infty} |x|f(x)dx < \infty$, then $E(X) = \int_{-\infty}^{\infty} xf(x)dx$ is defined as the mathematical expectation of X.

Next, we consider the conditional mathematical expectation of random variables. Let X, Y be two random variables, we need to compute the expectation $E(Y|X = x)$ or simply $E(Y|x)$ of Y under the given condition $X = x$. Suppose that the joint density of

(X, Y) is given, then the conditional probability density function of Y is $f(y|x)$ under the given condition $X = x$. We then have by definition that

$$E(Y|x) = \int_{-\infty}^{\infty} yf(y|x)dy.$$

The conditional mathematical expectation reflects the mean change of Y with respect to x. Hence, $E(Y|X)$ is a random variable and changes with X. In statistics, the conditional expectation $E(Y|x)$ is regarded as a function of x and is usually termed as the "regression function" of Y to X.

From conditional mathematical expectation, we can get an important formula to the unconditional mathematical expectation. Recall the law of total probability $P(A) = \sum_i P(B_i)P(A|B_i)$. This can be understood as to find the unconditional conditional probability $P(A)$ from the conditional probability $P(A|B_i)$ of A. In this regard, $P(A)$ is the weighted average of conditional probability $P(A|B_i)$ with weight being the probability $P(B_i)$. By analogy, the unconditional expectation of Y should be equal to the weighted average of the conditional expectation $E(Y|x)$ of x with weight proportional to the probability density $f_1(x)$ of X, i.e.,

$$E(Y) = \int_{-\infty}^{\infty} E(Y|x)f_1(x)dx.$$

The proof is not difficult and omitted here. Recalling that right-hand side (RHS) member of this formula is just the mathematical expectation of the random variable $E(Y|X)$ with respect to X, hence we have

$$E(Y) = E(E(Y|X)).$$

Next, we consider the conditional expectation under σ-subalgebra of $\mathscr{F}$.

Definition 1.1.11. *Let $(\Omega, \mathscr{F}, P)$ be a probability space and $\mathscr{G} \subset \mathscr{F}$. If $X : \Omega \to R^n$ is an integrable random variable, then $E(X|\mathscr{G})$ is defined as a random variable satisfying*
(i) *$E(X|\mathscr{G})$ is $\mathscr{G}$ measurable;*
(ii) *$\int_A XdP = \int_A E(X|\mathscr{G})dP \ \forall A \in \mathscr{G}$.*

The conditional mathematical expectation has the following properties:

Proposition 1.1.1.
(i) *Let X be $\mathscr{G}$ measurable, then $E(X|\mathscr{G}) = X$ almost surely (a.s.)*
(ii) *Let a, b be constants, then $E(aX + bY|\mathscr{G}) = aE(X|\mathscr{G}) + bE(Y|\mathscr{G})$ a.s.*
(iii) *Let X is $\mathscr{G}$ measurable and XY be integrable, then $E(XY|\mathscr{G}) = XE(Y|\mathscr{G})$ a.s.*
(iv) *Let X be independent of $\mathscr{G}$, then $E(X|\mathscr{G}) = E(X)$ a.s.*
(v) *Let $\mathscr{E} \subset \mathscr{G}$, then*

此为试读,需要完整PDF请访问: www.ertongbook.com