Applications of Mathematics

1 Fleming/Rishel, **Deterministic and Stochastic Optimal Control** (1975)
2 Marchuk, **Methods of Numerical Mathematics,** Second Edition (1982)
3 Balakrishnan, **Applied Functional Analysis,** Second Edition (1981)
4 Borovkov, **Stochastic Processes in Queueing Theory** (1976)
5 Liptser/Shiryayev, **Statistics of Random Processes I: General Theory** (1977)
6 Liptser/Shiryayev, **Statistics of Random Processes II: Applications** (1978)
7 Vorob'ev, **Game Theory: Lectures for Economists and Systems Scientists** (1977)
8 Shiryayev, **Optimal Stopping Rules** (1978)
9 Ibragimov/Rozanov, **Gaussian Random Processes** (1978)
10 Wonham, **Linear Multivariable Control: A Geometric Approach,** Third Edition (1985)
11 Hida, **Brownian Motion** (1980)
12 Hestenes, **Conjugate Direction Methods in Optimization** (1980)
13 Kallianpur, **Stochastic Filtering Theory** (1980)
14 Krylov, **Controlled Diffusion Processes** (1980)
15 Prabhu, **Stochastic Storage Processes: Queues, Insurance Risk, and Dams** (1980)
16 Ibragimov/Has'minskii, **Statistical Estimation: Asymptotic Theory** (1981)
17 Cesari, **Optimization: Theory and Applications** (1982)
18 Elliott, **Stochastic Calculus and Applications** (1982)
19 Marchuk/Shaidourov, **Difference Methods and Their Extrapolations** (1983)
20 Hijab, **Stabilization of Control Systems** (1986)
21 Protter, **Stochastic Integration and Differential Equations** (1990)
22 Benveniste/Métivier/Priouret, **Adaptive Algorithms and Stochastic Approximations** (1990)
23 Kloeden/Platen, **Numerical Solution of Stochastic Differential Equations** (1992)
24 Kushner/Dupuis, **Numerical Methods for Stochastic Control Problems in Continuous Time** (1992)
25 Flerning/Soner, **Controlled Markov Processes and Viscosity Solutions** (1993)
26 Baccelli/Brémaud, **Elements of Queueing Theory** (1994)
27 Winkler, **Image Analysis, Random Fields and Dynamic Monte Carlo Methods** (1995)
28 Kalpazidou, **Cycle Representations of Markov Processes** (1995)
29 Elliott/Aggoun/Moore, **Hidden Markov Models: Estimation and Control** (1995)
30 Hernández-Lerma/Lasserre, **Discrete-Time Markov Control Processes** (1995)

Peter E. Kloeden Eckhard Platen

Numerical Solution of Stochastic Differential Equations

With 85 Figures

Springer

Peter E. Kloeden

School of Computing and Mathematics, Deakin University
Geelong 3217, Victoria, Australia

Eckhard Platen

Institute of Advanced Studies, SMS, CFM, Australian National University
Canberra, ACT 0200, Australia

Second Corrected Printing 1995

Mathematics Subject Classification (1991): 60H10, 65C05

ISBN 3-540-54062-8 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-54062-8 Springer-Verlag New York Berlin Heidelberg

Library of Congress Cataloging-in-Publication Data.
Kloeden, Peter E. Numerical solution of stochastic differential equations/Peter E. Kloeden, Eckhard Platen. p. cm. – (Applications of mathematics; 23) "Second corrected printing" – T. p. verso.
Includes bibliographical references (p. –) and index.
ISBN 0-387-54062 (acid-free). – ISBN 3-540-54062-8 (acid-free)
1. Stochastic differential equations – Numerical solutions. I. Platen. Eckhard. II. Title. III. Series.
QA274.23.K56 1992b 519.2.– dc20 95-463 CIP

Printed in the United States of America

SPIN: 10502282 41/3143 - 5 4 3 2 1 0 – Printed on acid-free paper

Dedicated to Our Parents

Chapter 1

Probability and Statistics

The basic concepts and results of probability and stochastic processes needed later in the book are reviewed here. The emphasis is descriptive and PC-Exercises (PC= Personal Computer), based on pseudo-random number generators introduced in Section 3, are used extensively to help the reader to develop an intuitive understanding of the material. Statistical tests are discussed briefly in the final section.

1.1 Probabilities and Events

If we toss a die, then, excluding absurd situations, we always observe one of six basic outcomes; it lands with its upper face indicating one of the numbers 1, 2, 3, 4, 5 or 6. We shall denote these outcomes by ω_1, ω_2, ω_3, ω_4, ω_5 and ω_6, respectively, and call the set of outcomes $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$ the *sample space*. If we toss the die N times and count the number of times N_i that outcome ω_i occurs, we obtain a *relative frequency* $f_i(N) = N_i/N$. This number usually varies considerably with N, but experience tells us that as N becomes larger it approaches a limit $\lim_{N\to\infty} f_i(N) = p_i$, which we call the *probability* of outcome ω_i. Clearly $0 \leq p_i \leq 1$ for each $i = 1, 2, \ldots, 6$ and $\sum_{i=1}^{6} p_i = 1$; for a fair die each $p_i = 1/6$, giving a uniform distribution of probabilities over the outcomes.

Often we are interested in combinations of outcomes, that is subsets of the sample space Ω such as the subset $\{\omega_1, \omega_3, \omega_5\}$ of odd indexed outcomes. If we can distinguish such a combination by either its occurence or its nonoccurence we call it an *event*. Clearly if a subset A is an event, then its complement $A^c = \{\omega_i \in \Omega : \omega_i \notin A\}$ must also be an event. In particular, the whole sample space Ω is an event, which we call the *sure event* since one of its outcomes must always occur; its complement, the empty set $\emptyset$, is also an event but can never occur. We might think that every subset A of Ω should be an event, in which case we could determine its probability $P(A)$ from those of its constituent outcomes, that is as $P(A) = \sum_{\omega_i \in A} p_i$. However this corresponds to a situation of complete information about each of the outcomes, information which we may not always possess. For example, we may have only kept records of the occurences of odd or even indexed outcomes, but not of the actual outcomes themselves. Then we only distinguish and determine probabilities for the four subsets $\emptyset$, $O = \{\omega_1, \omega_3, \omega_5\}$, $E = \{\omega_2, \omega_4, \omega_6\}$ and Ω, which are thus the only events in this case. Actually, we could introduce new basic outcomes O (odd) and E (even) here;

shall return to it in Chapter 2 and present definitions there. Some terminology will however be useful before then. In particular, the collection of events $\mathcal{A}$ is known technically as a *σ-algebra* or *σ-field*. While $P(\emptyset) = 0$ always holds, there may also be nonempty events A with $P(A) = 0$; we call these *null events*. The sample space Ω is the *sure event*, and we say that any other event A with $P(A) = 1$ occurs *almost surely* (a.s.) or *with probability one* (w.p.1).

Regardless of how we actually evaluate it, the probability $P(A)$ of an event A is an indicator of the likelihood that A will occur. Our estimate of this likelihood may change if we possess some additional information, such as that another event has occured. For example, if we toss a fair die the probability of obtaining a 6 is $P(\{\omega_6\}) = p_6 = 1/6$ and the probability of obtaining an even number, that is the probability of the event $E = \{\omega_2, \omega_4, \omega_6\}$, is $P(E) = p_2 + p_4 + p_6 = 1/2$. If we know that an even number has been thrown, then, since this occurs in one of three equally likely ways, we might now expect that the probability of its being the outcome ω_6 is 1/3. We call this the conditional probability of the event $\{\omega_6\}$ given that the event E has occured and denote it by $P(\{\omega_6\}|E)$, noting that

$$P(\{\omega_6\}|E) = \frac{P(\{\omega_6\} \cap E)}{P(E)}$$

where $P(E) > 0$. In general, we define the *conditional probability* $P(A|B)$ of A given that an event B has occured by

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{1.8}$$

provided $P(B) > 0$ and define it to be equal to 0 (or we leave it undefined) in the vacuous case that $P(B) = 0$. This definition is readily suggested from the relative frequencies

$$\frac{N_{A\cap B}}{N_B} = \frac{N_{A\cap B}}{N} \Big/ \frac{N_B}{N},$$

where $N_{A\cap B}$ and N_B are the numbers of times that the events $A \cap B$ and B, respectively, occur out of N repetitions of what we usually call an *experiment*.

It is possible that the occurence or not of an event A is unaffected by whether or not another event B has occured. Then its conditional probability $P(A|B)$ should be the same as $P(A)$, which with (1.8) implies that

$$P(A \cap B) = P(A)P(B) \tag{1.9}$$

In this case we say that the events A and B are *independent*. For example, events A and B are independent if $P(A) = P(B) = 1/2$ and $P(A \cap B) = 1/4$. This particular situation occurs if we toss a fair coin twice, with A the event that we obtain a head on the first toss and B a head on the second toss, provided that the way we toss the coin the second time is not biased by the outcome of the first toss. This may seem tautological; in fact, we shall use the independence of outcomes of a repeated experiment to define *independent repetitions* of the experiment. Finally, we say that n events $A_1, A_2, \ldots, A_n$ are *independent* if

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_k}) \tag{1.10}$$

for all nonempty subsets $\{i_1, i_2, \ldots, i_k\}$ of the set of indices $\{1, 2, \ldots, n\}$.

1.2 Random Variables and Distributions

We are often interested in numerical quantities associated with the outcome of a probabilistic experiment, such as our winnings in a gambling game based on tossing a die or the revenue made by a telephone company based on the number of calls made. These numbers, $X(\omega)$ say, provide us with information about the experiment, which, of course, can never exceed that already summarized in its probability space $(\Omega, \mathcal{A}, P)$. They correspond to the values taken by some function $X : \Omega \to \Re$, which we call a random variable if its information content is appropriately restricted.

Consider the *indicator function* I_A of a subset A of Ω defined by

$$I_A(\omega) = \begin{cases} 0 & : \ \omega \notin A \\ 1 & : \ \omega \in A, \end{cases} \tag{2.1}$$

which is thus a function from Ω into $\Re$. For I_A to be a random variable we require A to be an event, or equivalently the subset

$$\{\omega \in \Omega : I_A(\omega) \le a\} = \begin{cases} \emptyset & : \ a < 0 \\ A^c & : \ 0 \le a < 1 \\ \Omega & : \ 1 \le a \end{cases}$$

to be an event for each $a \in \Re$.

In general, for a probability space $(\Omega, \mathcal{A}, P)$ we say that a function $X : \Omega \to \Re$ is a *random variable* if

$$\{\omega \in \Omega : X(\omega) \le a\} \in \mathcal{A} \quad \text{for each} \quad a \in \Re, \tag{2.2}$$

that is if $\{\omega \in \Omega : X(\omega) \le a\}$ is an event for each $a \in \Re$. This is not quite so restrictive as it may seem because it implies that $\{\omega \in \Omega : X(\omega) \in B\}$ is an event for any *Borel subset* B of $\Re$, that is any subset of $\Re$ in the σ-algebra $\mathcal{B}$ is generated from countable unions, intersections or complements of the semi-infinite intervals $\{x \in \Re; -\infty < x \le a\}$.

We call the function P_X defined for all $B \in \mathcal{B}$ by

$$P_X(B) = P(\{\omega \in \Omega : X(\omega) \in B\}) \tag{2.3}$$

the *distribution* of the random variable X, and note that the ordered triple $(\Re, \mathcal{B}, P_X)$ is a probability space which contains all of the essential information associated with the random variable X. Since point functions are simpler than set functions we often restrict attention to the function $F_X : \Re \to \Re$ defined for each $x \in \Re$ by

(2.4) $$F_X(x) = P_X((-\infty, x)) = P(\{\omega \in \Omega : X(\omega) < x\}),$$

which we call the ***distribution function*** of X. For example, if A is an event the distribution function for its indicator function I_A is

(2.5) $$F_{I_A}(x) = \begin{cases} 0 & : \ x < 0 \\ 1 - P(A) & : \ 0 \le x < 1 \\ 1 & : \ 1 \le x \end{cases}$$

Since for any $x < y$ we have $\{\omega \in \Omega : X(\omega) \le x\} \subseteq \{\omega \in \Omega : X(\omega) \le y\}$ and hence $P(\{\omega \in \Omega : X(\omega) \le x\}) \le P(\{\omega \in \Omega : X(\omega) \le y\})$, we can see that any distribution function F_X satisfies:

(2.6) $$\lim_{x \to -\infty} F_X(x) = 0 \quad \text{and} \quad \lim_{x \to +\infty} F_X(x) = 1$$

with

(2.7) $$F_X(x) \quad \text{nondecreasing in} \quad x.$$

The example (2.5) shows that F_X need not be a continuous function, but from properties (1.1)–(1.7) and (2.4) we can show that F_X is always *continuous from the right*, that is

(2.8) $$\lim_{h \to 0+} F_X(x + h) = F_X(x) \quad \text{for all} \quad x \in \Re.$$

Conversely, for any function $F : \Re \to \Re$ satisfying properties (2.6)–(2.8), we can define a random variable X which has F as its distribution function. We can use $(\Re, \mathcal{B}, P_X)$ as the underlying probability space for this random variable, with P_X defined for subintervals $(-\infty, x]$ in terms of F using (2.4) and then extended appropriately to the more general Borel subsets. In this setting the

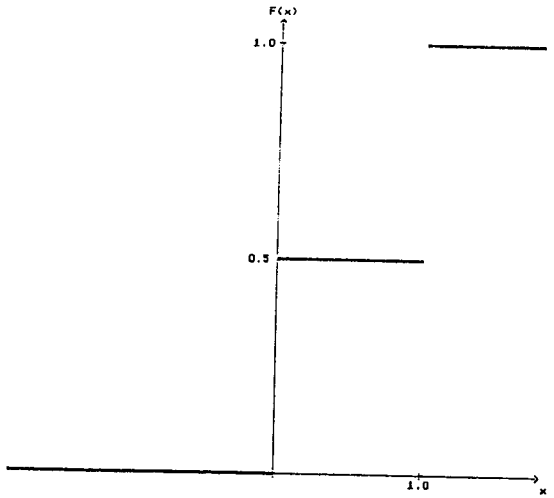

Figure 1.2.1 Distribution (2.5) for $P(A) = \frac{1}{2}$.

random variable is the identity function on $\Re$. Often we omit mention of this probability space and concentrate on the distribution function.

In applications the following examples are frequently encountered.

Example 1.2.1 *The simplest nontrivial random variable X takes just two distinct real values x_1 and x_2, where $x_1 < x_2$, with probabilities p_1 and $p_2 = 1 - p_1$, respectively. It is often called a* two-point random variable *and its distribution function is given by*

$$F_X(x) = \begin{cases} 0 & : \ x < x_1 \\ p_1 & : \ x_1 \le x < x_2 \\ 1 & : \ x_2 \le x \end{cases}$$

For instance, the indicator function I_A of an event A is such a random variable with $x_1 = 0$ and $x_2 = 1$ (see (2.5)). Another two-point random variable arises in the gambling game where we win a dollar when a tossed coin shows a head and lose a dollar when it shows a tail; here $x_1 = -1$ and $x_2 = +1$.

Example 1.2.2 *In radioactive decay the number of atoms decaying per unit time is a random variable X taking values* 0, 1, 2, ... *without any upper bound. The probabilities $p_n = P(X = n)$ often satisfy the* Poisson distribution *with*

$$p_n = \frac{\lambda^n}{n!} \exp(-\lambda)$$

for $n = 0, 1, 2, \ldots$, where $\lambda > 0$ is a given parameter.

The above two examples are typical of a discrete random variable X taking a finite or countably infinite number of distinct values $x_0 < x_1 < \cdots < x_n < \cdots$ with probabilities $p_n = P(X = x_n)$ for $n = 0, 1, 2, \ldots$. The distribution function F_X here satisfies

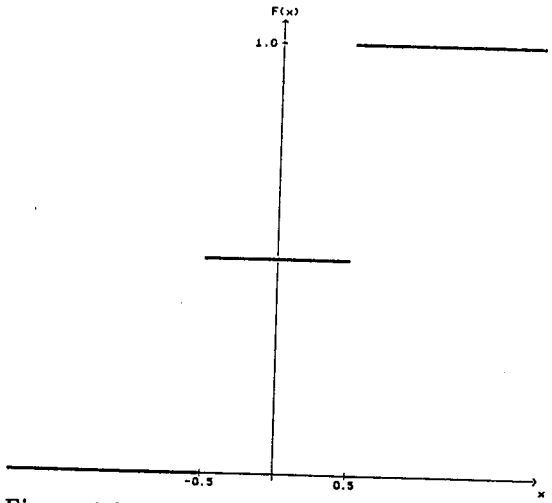

Figure 1.2.2 Distribution for Example 1.2.1 with $-x_1 = x_2$ and $p_1 = 0.5$.

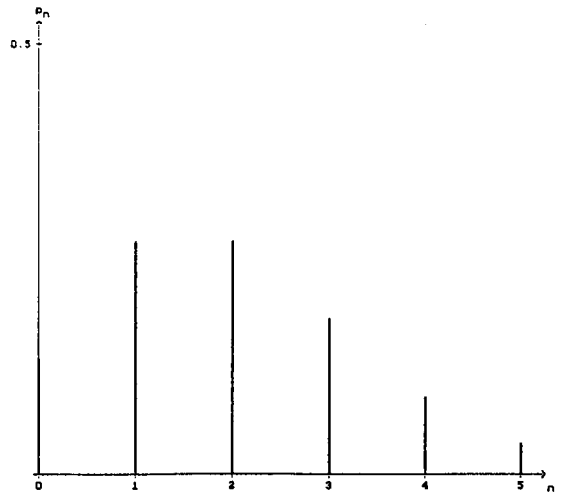

Figure 1.2.3 Poisson probabilities for $\lambda = 2$.

$$F_X(x) = \begin{cases} 0 & : \quad x < x_0 \\ \sum_{i=0}^{n} p_i & : \quad x_n \le x < x_{n+1},\ n = 0, 1, \ldots \end{cases} \tag{2.9}$$

F_X is a step-function with steps of height p_n at $x = x_n$. For such a random variable the set $\{x_0, x_1, x_2, \ldots\}$ could be used as the sample space Ω, with all of its subsets being events.

In sharp contrast are the random variables taking all possible values in $\Re$. We call such a random variable X a *continuous random variable* if the probability $P(\{\omega \in \Omega : X(\omega) = x\})$ is zero for all $x \in \Re$. In this case the distribution function F_X is often differentiable, that is there exists a nonnegative function p, called the *density function*, such that $F_X'(x) = p(x)$ for each $x \in \Re$; when F_X is only piecewise differentiable this holds everywhere except at certain isolated points. Then

$$F_X(x) = \int_{-\infty}^{x} p(s)\, ds \tag{2.10}$$

for all $x \in \Re$, including the above mentioned exceptional points in the piecewise differentiable case. Such a distribution function is usually said to be *absolutely continuous*. The following are commonly occuring examples.

Example 1.2.3 *Consider a random variable X which only takes values in a finite interval $a \le x \le b$, such that the probability of its being in a given subinterval is proportional to the length of the subinterval. Then the distribution function is given by*

$$F_X(x) = \begin{cases} 0 & : \quad x < a \\ \frac{x-a}{b-a} & : \quad a \le x \le b \\ 1 & : \quad b < x \end{cases}$$

which is differentiable everywhere except at $x = a$ and $x = b$. The corresponding density function is given by

$$p(x) = \begin{cases} 0 & : \quad x \notin [a, b] \\ \frac{1}{b-a} & : \quad x \in [a, b] \end{cases}$$

We say that the random variable X in Example 1.2.3 is *uniformly distributed* on $[a, b]$ and denote this by $X \sim U(a, b)$. Alternatively, we say that it has a *rectangular* density function.

The $U(0, 1)$ random variables are a special case of the beta- distributed random variable with parameters $\alpha = \beta = 1$. In general, the *beta-distribution* with positive parameters α and β has the density function

$$p(x) = \begin{cases} 0 & : \quad x \notin [0, 1] \\ \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)} & : \quad x \in [0, 1] \end{cases} \tag{2.11}$$

where

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\, dx.$$

Example 1.2.4 *The life-span of a light bulb is a random variable X which is often modelled by the* exponential distribution

$$F_X(x) = \begin{cases} 0 & : \quad x < 0 \\ 1 - \exp(-\lambda x) & : \quad x \ge 0 \end{cases}$$

for some intensity parameter $\lambda > 0$. F_X is differentiable everywhere except for $x = 0$ and has the density function

$$p(x) = \begin{cases} 0 & : \quad x < 0 \\ \lambda \exp(-\lambda x) & : \quad x \ge 0. \end{cases}$$

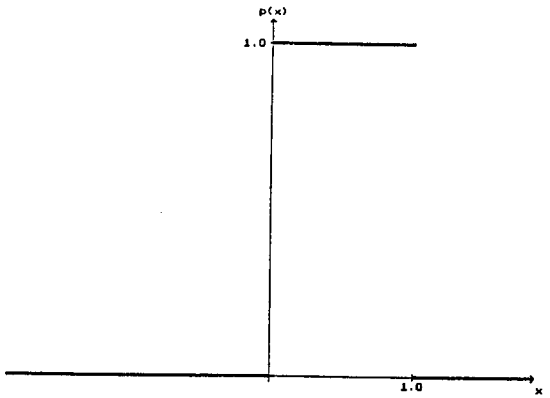

Figure 1.2.4 The rectangular density with $a = 0$ and $b = 1$.

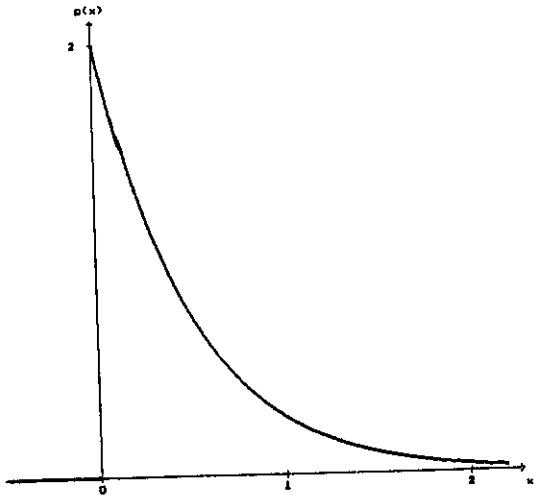

Figure 1.2.5 The exponential density with $\lambda = 2$.

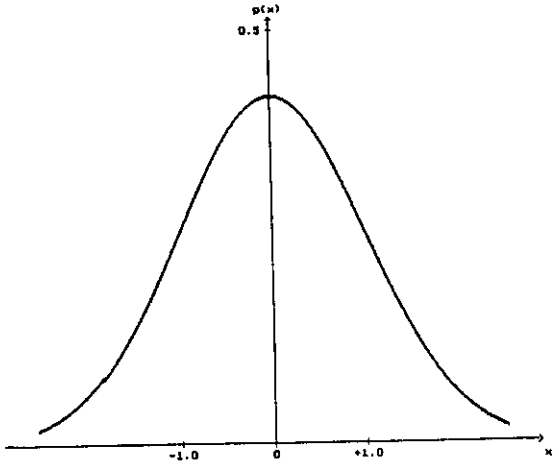

Figure 1.2.6 The standard Gaussian density.

Example 1.2.5 *The density function*

$$p(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) \tag{2.12}$$

has a bell-shaped graph which is symmetric about $x = 0$. The corresponding distribution function $F_X(x)$ is differentiable everywhere and has a sigmoidal-shaped graph, but must be evaluated numerically or taken from tables since no anti-derivative is known in analytical form for (2.12). A random variable with this density function is called a standard Gaussian random variable.

Gaussian random variables occur so commonly in applications, for example as measurement errors in laboratory experiments, that they are often said

to be *normally distributed.* Their ubiquity is explained by the fundamental theorem of probability and statistics, the Central Limit Theorem, which will be discussed in Section 5.

In anticipation of the next section, we say that two random variables X and Y are *independent* if the events $\{\omega : X(\omega) \in A\}$ and $\{\omega : Y(\omega) \in B\}$ are independent for all Borel sets A and B. Essentially, the values taken by either of the random variables are uninfluenced by those taken by the other. More will be said about this concept in Section 4.

1.3 Random Number Generators

The numerical simulation of a mathematical model of a complicated probabilistic system often provides information about the behaviour of the model, and hopefully of the original system itself, which cannot be obtained directly or easily by other means. Numerical values of each of the random variables must be provided for a test run of the model, and then the outputs of many test runs are analysed statistically. This procedure requires the generation of large quantities of random numbers with the specified statistical properties. Originally such numbers were taken directly from actual random variables, generated, for example, mechanically by tossing a die or electronically by the noisy output of a valve, and often listed in random number tables. This proved impractical for large scale simulations and the numbers were not always statistically reliable. In addition, a particular sequence of random numbers could not always be reproduced, an important feature for comparative studies, and so had to be stored. The advent of electronic computers lead to the development of simple deterministic algorithms to generate sequences of random variables, quickly and reproducably. Such numbers are consequently not truly random, but with sufficient care they can be made to resemble random numbers in most properties, in which case they are called *pseudo-random numbers.*

These days most digital computers include a *linear congruential pseudo-random number generator.* These have the recursive form

$$X_{n+1} = aX_n + b \pmod{c} \tag{3.1}$$

where a and c are positive integers and b a nonnegative integer. For an integer initial value or *seed* X_0, the algorithm (3.1) generates a sequence taking integer values from 0 to $c-1$, the remainders when the aX_n+b are divided by c. When the coefficients a, b and c are chosen appropriately the numbers

$$U_n = X_n/c \tag{3.2}$$

seem to be uniformly distributed on the unit interval $[0, 1]$. Since only finitely many different numbers occur, the modulus c should be chosen as large as possible, and perhaps also as a power of 2 to take advantage of the binary arithmetic used in computers. To prevent cycling with a period less than c the multiplier a should also be taken relatively prime to c. Typically b is

chosen equal to zero, the resulting generator then being called a *multiplicative generator*. A much used example was the *RANDU* generator of the older *IBM* Scientific Subroutine Package with multiplier $a = 65,539 = 2^{16} + 3$ and modulus $c = 2^{31}$; the *IBM* System 360 Uniform Random Number Generator uses the multiplier $a = 16,807 = 7^5$ and modulus $c = 2^{31} - 1$, which is a prime number.

The reader is referred to specialist textbooks for an extensive discussion on pseudo-random number generators and their properties. In Section 9 we shall mention some basic tests for checking their statistical properties and reliability. For the remainder of this section we shall assume that we have a subroutine *RDN* which provides us with $U(0,1)$ uniformly distributed pseudo-random numbers by means of (3.2) and a generator (3.1). We shall show how we can then use this subroutine to generate pseudo-random numbers with other commonly encountered distributions, in particular those described in Section 2.

A two-point random variable X (see Example 1.2.1) taking values $x_1 < x_2$ with probabilities p_1 and $p_2 = 1 - p_1$ can be generated easily from a $U(0,1)$ random variable U, namely with

$$X = \begin{cases} x_1 & : \quad 0 \le U \le p_1 \\ x_2 & : \quad p_1 < U \le 1. \end{cases} \tag{3.3}$$

This idea extends readily to an N-state random variable X taking values $x_1 < x_2 < \cdots < x_N$ with nonzero probabilities $p_1, p_2, \ldots, p_N$ where $\sum_{i=1}^{N} p_i = 1$. With $s_0 = 0$ and $s_j = \sum_{i=1}^{j} p_i$ for $j = 1, 2, \ldots, N$ we set $X = x_{j+1}$ if $s_j < U \le s_{j+1}$ for $j = 0, 1, 2, \ldots, N-1$. We could approximate an infinite-state Poisson distributed random variable (see Example 1.2.2) by an N-state random variable for some large N by coalescing all of the remaining, less probable, states into the final state. However, computationally more efficient algorithms for Poisson random variables have been developed, but will not be given here.

The corresponding method for a continuous random variable X requires the probability distribution function F_X to be inverted when this is possible. For a number $0 < U < 1$ we define $x(U)$ by $U = F_X(x(U))$, so $x(U) = F_X^{-1}(U)$ if F_X^{-1} exists, or in general

$$x(U) = \inf\{x : U \le F_X(x)\}. \tag{3.4}$$

This is called the *inverse transform method* and is best used when (3.4) is easy to evaluate. For example, the exponential random variable with parameter $\lambda > 0$ (see Example 1.2.4) has an invertible distribution function with

$$x(U) = F_X^{-1}(U) = -\ln(1-U)/\lambda \quad \text{for} \quad 0 < U < 1.$$

When U is $U(0,1)$ distributed, then so too is $1 - U$ and a little computational effort can be spared by using *RDN* directly to generate $1 - U$.

In principle the inverse transform method could be used for any continuous random variable, but may require too much computational effort to evaluate

(3.4). This is the situation with the standard Gaussian random variable, since the integrals for its distribution function must be evaluated numerically. The *Box-Muller method* for generating standard Gaussian random variables avoids this problem. It is based on the observation that if U_1 and U_2 are two independent $U(0,1)$ uniformly distributed random variables, then N_1 and N_2 defined by

$$\begin{aligned} N_1 &= \sqrt{-2\ln(U_1)}\cos(2\pi U_2) \\ N_2 &= \sqrt{-2\ln(U_1)}\sin(2\pi U_2) \end{aligned} \tag{3.5}$$

are two independent standard Gaussian random variables. This can be verified with a change of coordinates from cartesian coordinates (N_1, N_2) to polar coordinates (r, θ) and then to $U_1 = \exp(-\frac{1}{2}r^2)$ and $U_2 = \theta/2\pi$ (see Exercise1.4.11).

A variation of the Box-Muller method which avoids the time consuming calculation of trigonometric functions is the *Polar Marsaglia method*. It is based on the facts that $V = 2U - 1$ is $U(-1,1)$ uniformly distributed if U is $U(0,1)$ distributed and that for two such random variables V_1 and V_2 with $W = V_1^2 + V_2^2 \le 1$, W is $U(0,1)$ distributed and $\theta = \arctan(V_1/V_2)$ is $U(0, 2\pi)$ distributed. Since the inscribed unit circle has $\pi/4$ of the area of the square $[-1,1]^2$, the point (V_1, V_2) will take values inside this circle with probability $\pi/4 \approx 0.7864816\cdots$. We only consider these points, discarding the others. Using

$$\cos\theta = \frac{V_1}{\sqrt{W}}, \quad \sin\theta = \frac{V_2}{\sqrt{W}}$$

when $W = V_1^2 + V_2^2 \le 1$ we can rewrite (3.5) (see Exercise 1.4.11) as

$$\begin{aligned} N_1 &= V_1\sqrt{-2\ln(W)/W} \\ N_2 &= V_2\sqrt{-2\ln(W)/W} \end{aligned} \tag{3.6}$$

Although a proportion of the generated uniformly distributed numbers are discarded, this method is often computationally more efficient than the Box-Muller method when a large quantity of numbers is to be generated.

We shall use the above methods in exercises in the following sections, assuming that they do indeed generate random numbers with the asserted properties. In Section 9 we shall examine the validity of this assumption.

PC-Exercise 1.3.1 *Write a program for two-point, exponential and Gaussian random variables based on the above methods to generate a list of pseudo-random numbers.*