CRC

# MONTE CARLO SIMULATION *in the* RADIOLOGICAL SCIENCES

Richard L. Morin

CRC PRESS

# Monte Carlo Simulation in the Radiological Sciences

NOT FOR SALE

Editor

**Richard L. Morin, Ph.D.**
Senior Associate Consultant
and
Associate Professor
Department of Radiology
Mayo Clinic
Rochester, Minnesota
formerly
Director
Section of Physics
Department of Radiology
University of Minnesota
Minneapolis, Minnesota

馆藏专用章
图书室
重医附一院 00162781
2000年6月12

CRC Press, Inc.
Boca Raton, Florida

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

Monte Carlo simulation in the radiological sciences / editor, Richard
L. Morin.
p. cm.
Bibliography: p.
Includes index.
ISBN 0-8493-5559-1
1. Radiology, Medical--Statistical methods. 2. Monte Carlo
method. I. Morin, Richard L., 1949- .
R905.M66 1988
616.07'57--dc19 88-10478
CIP

This book represents information obtained from authentic and highly regarded sources. Reprinted material is quoted with permission, and sources are indicated. A wide variety of references are listed. Every reasonable effort has been made to give reliable data and information, but the author and the publisher cannot assume responsibility for the validity of all materials or for the consequences of their use.

All rights reserved. This book, or any parts thereof, may not be reproduced in any form without written consent from the publisher.

Direct all inquiries to CRC Press, Inc., 2000 Corporate Blvd., N.W., Boca Raton, Florida, 33431.

© 1988 by CRC Press, Inc.

International Standard Book Number 0-8493-5559-1

Library of Congress Card Number 88-10478
Printed in the United States

# PREFACE

This text is intended to aid Radiological Science investigators in understanding and using Monte Carlo methods in their research. There are certainly many fine discussions and reviews of the Monte Carlo method; however, the Monte Carlo novice often finds it disconcerting to assemble the many source texts and papers prior to utilizing this robust technique. We hope this volume reduces the new investigator's labors during the ''bootstrap'' process and aids in furthering the popularization of this technique within the Radiological Sciences.

The presentation is broadly organized into tutorial and application sections. Chapters 2 and 3 describe the fundamental inputs of the Monte Carlo technique, while Chapter 4 describes the fundamental mechanisms involved in obtaining a solution utilizing Monte Carlo simulation. Chapters 5 through 8 discuss in detail various applications in Diagnostic Radiology, Nuclear Medicine, and Radiation Therapy. The reader will note that occasional duplication occurs throughout the text. This is intentional. Explanation of the same concept from different points of view or through the thoughts of different authors is very often useful to the new investigator. In addition, since the nomenclature and approach to a particular simulation is often inherently linked to a particular application, explanation in the context of the simulation problem under discussion seemed to be the most beneficial and efficient means of presentation.

On a personal note, I thank the contributors for their diligent work and hope you find this material interesting and useful in your efforts.

**RLM**

# EDITOR

**Richard L. Morin, Ph.D.,** is a Senior Associate Consultant in the Radiology Department, Mayo Clinic and Associate Professor of Radiologic Physics, Mayo Medical School, Rochester, Minnesota.

Doctor Morin received his M.S. in Radiation Biophysics from the University of Florida in 1973 and his Ph.D. from the University of Oklahoma in Radiological Sciences in 1980. He was formerly Director of the Physics Section in the Radiology Department at the University of Minnesota. He is a diplomate of the American Board of Radiology in Diagnostic Radiological Physics and Nuclear Medical Physics.

Doctor Morin has presented numerous lectures at international and national scientific meetings and has published more than 35 research papers. His current research interests include computer applications in the radiological sciences with particular emphasis in quantitative computed tomography, magnetic resonance imaging, and medical imaging transmission, display, and analysis technologies.

此为试读，需要完整PDF请访问： www.ertongbook.com

## CONTRIBUTORS


**Heang-Ping Chan, Ph.D.**
Associate Professor
Department of Radiology
University of Chicago
Chicago, Illinois

**Kunio Doi, Ph.D.**
Professor
Department of Radiology
University of Chicago
Chicago, Illinois

**James E. Goin, Ph.D.**
Biostatistican
Geometric Data
Wayne, Pennsylvania

**Ravinder Nath, Ph.D.**
Professor
Department of Therapeutic Radiology
Yale University Medical School
New Haven, Connecticut

**David E. Raeside, Ph.D.**
Professor
Department of Radiological Sciences
University of Oklahoma Health Sciences Center
Oklahoma City, Oklahoma

**John C. Widman, Ph.D.**
Radiological Safety Officer
KMS Fusion, Inc.
Ann Arbor, Michigan

**Jeffrey F. Williamson, Ph.D.**
Assistant Professor
Department of Radiation Oncology
University of Arizona Health Science Center
Tucson, Arizona



重庆医科大学第一医院
图书室

2000年6月12

*To*
*My Parents*
*Joyce & Charles Morin*

TABLE OF CONTENTS

Chapter 1

# INTRODUCTION

**Richard L. Morin**

The Monte Carlo simulation method has found widespread application in the radiological sciences. Its popularization as an investigational tool is demonstrated by the number of papers in *Medical Physics* and *Physics in Medicine and Biology* which utilize Monte Carlo simulation techniques (approximately a fivefold increase since 1983). This expansion is perhaps most likely due to the computing power which is easily accessible to the radiological physicist. At present, it is no longer necessary to consider performing Monte Carlo simulation solely on large mainframe computer systems. In addition, the presence and availability of software packages (e.g., ETRAN, OGRE-G, MORSE) has increased the utilization of the Monte Carlo simulation technique. In this chapter, we briefly review the chronology of the development of Monte Carlo simulation and describe the fundamentals of the technique.

The earliest use of the Monte Carlo simulation technique was probably around 1873, involving a method for the calculation of the constant pi.[1] This instance was reported by Hall[1,2] and concerned the work of a Captain Fox and others who performed the experiment while recovering from Civil War wounds. Surprisingly, modern Monte Carlo techniques were used around 1901 by Lord Kelvin[1,3] in a discussion of the Boltzman equation; however, he did little to propagate the method. The popularization of the technique and the term Monte Carlo is generally attributed to the work of J. von Neumann, S. Ulam, and E. Fermi during World War II.[1,4,5] The term ''Monte Carlo'' was given by von Neumann to a secret Los Alamos project concerning neutron diffusion[4] and reflected the idea that a conceptual roulette wheel could be employed to determine neutron absorption (the number of compartments following the probability distribution for neutron diffusion). The number of events to be analyzed was quite enormous, so a computer was substituted for an actual roulette wheel.

Since that time the Monte Carlo technique has found widespread application. A brief survey includes: the simulation of nerve muscle end plate potentials and genetic research (Biology), the analysis of phase shifts and gas kinetics (Chemistry), photon and particle diffusion (Physics), radiation dosimetry and characterization (Medical Physics), radiation protection (Health Physics), the evaluation of definite integrals (Mathematics), evaluation of beam shot noise (Electrical Engineering), forest growth and pollution studies (Environmental Engineering), the analysis of stock market performance and Gross National Product growth (Economics), and the behavior of customer arrival sequences (Business). Since the Monte Carlo technique is inherently linked to computers, this widespread use has given rise to specialized computed languages for it implementation such as GPSS, DYNAMO, and SIMSCRIPT.

The Monte Carlo simulation technique can be understood in terms of a black box diagram (Figure 1). The procedure can be considered as a two-input, one-output device. The inputs are a large source of high-quality random numbers and some probability law which we wish to examine. The output is the result of random sampling of the probability distribution under investigation. For example, if the underlying probability distribution for forest growth is known (perhaps from historical data), it is possible, using the Monte Carlo approach, to determine the way in which the forest would grow over some time period. In the same manner, the precise transmission of a Bremsstrahlung beam through tissue can be ascertained.

It should be noted that the Monte Carlo technique is sometimes discussed in the literature under terms such as simulation, gaming, or modeling; however, the converse is not always true. The hallmark of Monte Carlo is the use of random sampling. A simulation which is

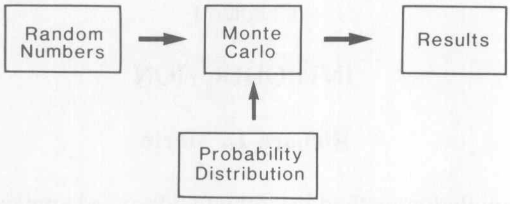


FIGURE 1. Black Box illustration of the Monte Carlo technique.

strictly analytical in nature is not a Monte Carlo simulation. Throughout this text, unless otherwise stated, simulation is understood to mean Monte Carlo simulation. Further information concerning the Monte Carlo technique can be found in the tutorial references 1 and 4 through 7. In order to understand the use of this fundamental technique, we shall now examine each input of the Monte Carlo black box and present a variety of simulation techniques and results.

## REFERENCES

1. **Hammersley, J. M. and Handscomb, D. C.,** *Monte Carlo Methods,* Methuen, London, 1964.
2. **Hall, A.,** On an experimental determination of $\pi$, *Messeng Math,* 2, 113, 1873.
3. **Kelvin, L.,** Nineteenth century clouds over the dynamical theory of heat and light, *Philos. Mag.,* 2, 1, 1901.
4. **McGraken, D.,** The Monte Carlo method, *Sci. Am.,* 192(5), 90, 1955.
5. **Halton, J. H.,** A retrospective and prospective survey of the Monte Carlo method, *SIAM Rev.,* 12, 1, 1970.
6. **Fluendy, M.,** Monte Carlo methods, in *Markov Chains and Monte Carlo Calculations in Polymer Science,* Lowry, G. G., Ed., Marcel Dekker, New York, 1970.
7. **Raeside, D. E.,** Monte Carlo principles and applications, *Phys. Med. Biol.,* 21, 181, 1978.

Chapter 2

# PROBABILITY AND STATISTICS

**David E. Raeside**

## TABLE OF CONTENTS



此为试读，需要完整PDF请访问： www.ertongbook.com

## I. INTRODUCTION

The proper design of a Monte Carlo simulation requires an understanding of probability and statistics. In addition, such an understanding is required for the intelligent interpretation of Monte Carlo simulation results. This chapter is an introduction to probability and statistics.

## II. ELEMENTS OF PROBABILITY THEORY

This section is a brief introduction to the basic ideas of probability theory. Readers requiring a more complete presentation should consult one of the numerous books on the subject. Papoulis' book[1] is recommended because it contains many worked examples. The two-volume work by Feller[2,3] has become a classic; it is recommended most highly. The two slim volumes by Lindley[4,5] are well-written and they also are recommended.

Any quantity that cannot be specified without the use of probability laws is called a random variable. A random variable will be denoted by a capital letter with the lower case of the letter being reserved for a realization or fixed value of that random variable. A random variable X is said to have a discrete distribution if X can take only a countable number of distinct values $x_1$, $x_2$, . . . (this set of values can be finite or infinite). A random variable X is said to have a continuous distribution if X can take any value between the limits $x_1$ and $x_2$ (there may be more than one such interval for which this is true). For a continuous random variable, X, the probability that X lies within the interval $(x_1,x_2)$, $\mathrm{Prob}(x_1 < X < x_2)$, is given by an integral:

$$\mathrm{Prob}(x_1 < X < x_2) = \int_{x_1}^{x_2} p(x)dx \tag{1}$$

This relation defines a function p(x) which is called the probability density of X. The following convention distinguishes between the probability laws governing discrete and continuous random variables: the upper case letter P denotes a probability (also called a probability distribution function or a probability mass function) while the lower case of that letter (p) denotes a probability density. It is required that $\int_{-\infty}^{\infty} p(x)dx = 1$, where p(x) is integrated over all possible values of the continuous random variable X. Similarly, any probability distribution function P(x) must satisfy the requirement $\sum_x P(x) = 1$, where the sum is over all possible values of x.

The two most famous probability laws encountered in the radiological sciences are the normal and Poisson probability laws. The normal (or Gauss) probability law associated with the continuous random variable X is $p(x) = (2\pi\sigma^2)^{-1/2} \exp\{-(x - \mu)^2/2\sigma^2\}$, where $-\infty < x < \infty$ ($\mu$ and $\sigma^2$ are parameters of the density function). If one wished to know the probability that a normally distributed random variable fell within the interval $(x_1,x_2)$, it would be necessary to integrate this density function between the limits $x_1$ and $x_2$. A discrete random variable Y is said to have a Poisson distribution if the probability that Y = y is given by $P(y) = \lambda^y e^{-\lambda}/y!$, where $\lambda$ is a parameter of the distribution and $y = 0,1,2, \ldots$ . The probability that the random variable Y takes a particular value is simply the function P evaluated at that value.

The Poisson distribution has an interesting history. Textbooks that bother to credit original sources invariably claim that the first statement of the Poisson distribution is contained in Poisson's 1837 manuscript "Recherches sur la Probabilité des Judgments en Matière Criminelle et en Matière Civile, Précédées des Regles Générales du Calcul des Probabilités".[6] Although it is correct to cite this work as a source of an independent statement of the Poisson distribution, it is necessary to go back more than a century before Poisson's work to find the origin of the "Poisson distribution" in de Moivre's 1718 treatise "The Doctrine of Chances"[7] (see Reference 8, page 113). According to Haight,[8] not only was the knowledge of the Poisson distribution "lost" in the dozen decades between de Moivre and Poisson, but even after Poisson's work it lay dormant until the close of the nineteenth century. To quote Haight[8] (page 114): "The splendor of normal theory dazzled scientific writers of the period and led many of them to believe that it was a 'universal' law, governing every possible kind of variation." Ladislaus von Bortkewitsch altered that attitude forever with his 1898 book "Das Gesetz der Kleinen Zahlen",[9] a compendium of knowledge about the Poisson distribution with illustrative applications of the theory. In 1910, Bateman wrote an appendix for a paper by the physicists Rutherford and Geiger.[10] In this historic work, Bateman showed that the count rate of a radioactive sample satisfied a set of differential equations which have the Poisson distribution as their solution. This was not only an independent discovery of the Poisson distribution, but it was the first time that the Poisson distribution had been connected so intimately with physics (Reference 8, page 120). The identification of radioactive decay as a Poisson process causes one to wonder why von Bortkewitsch's book[9] has the title "The Law of Small Numbers". It certainly is not true that count rate measurements of radioactive samples always yield small numbers (in fact, the rates are often sufficiently large to "paralyze" the counting apparatus). What then did von Bortkewitsch have in mind? There is no doubt that it was a "law of rare events" that was being enunciated by von Bortkewitsch rather than a "law of small numbers". The decay of a particular nucleus in a radioactive sample is a rare event, but, because the number of radioactive nuclei present in the sample is usually extremely large, the product of these two quantities need not be small.

It is common to require the probability of several random variables simultaneously taking certain values (or falling within certain ranges). For the discrete random variables X, Y, and Z, for example, this involves an extension of the concept of a probability distribution function called the joint probability distribution function, denoted by

$$P(X = x, \quad Y = y, \quad Z = z) = P(x, y, z) \tag{2}$$

where P(x, y, z) is read "the probability that the random variable X is x, the random variable Y is y, and the random variable Z is z". For the continuous random variables H, I, J, the analogous quantity is the joint probability density p(h, i, j) with the properties

$$\text{Prob}(h_1 < H < h_2, \quad i_1 < I < i_2, \quad j_1 < J < j_2)$$

$$= \int_{h_1}^{h_2}\int_{i_1}^{i_2}\int_{j_1}^{j_2} p(h, i, j)dhdidj \tag{3}$$

and

$$\text{Prob}(h_1 < H < h_2) = \int_{h_1}^{h_2} p(h, i, j)dh \tag{4}$$

with similar definitions for Prob $(i_1 < I < i_2)$ and Prob$(j_1 < J < j_2)$. These definitions extend, of course, to cases involving more than three random variables.

Conditional probabilities are used to deal with problems involving two random variables X and Y, such that the distribution of X given that Y is equal to some particular value y is of interest. The defining equation for a conditional probability law is

$$\text{Prob}(X = x \mid Y = y) \equiv \frac{\text{Prob}(X = x,\quad Y = x)}{\text{Prob}(Y = y)} \tag{5}$$

where Prob(X = x | Y = y) is read "the probability that X = x given that Y = y". A mixture of "P's" and "p's" may appear in this expression. The concept of conditional probability can be extended to more than two random variables. For two discrete random variables X and Y, P(x,y) = P(y,x), or since P(x,y) = P(x|y)P(y) and P(y,x) = P(y|x)P(x),

$$\begin{aligned} P(x|y) &= \frac{P(x,\ y)}{P(y)} \\ &= \frac{P(x,\ y)}{\sum_x P(x,\ y)} \\ &= \frac{P(y|x)P(x)}{\sum_x P(y|x)P(x)} \end{aligned} \tag{6}$$

This is a form of one of the most important theorems in probability: Bayes' theorem. It is possible to state Bayes's theorem so as to include probability densities as well as probability distributions. An example of Bayes' theorem in one of these alternate forms can be given for the case of the Poisson distribution:

$$p(\lambda|x) = \frac{P(x|\lambda)p(\lambda)}{\int_0^\infty P(x|\lambda)p(\lambda)d\lambda} \tag{7}$$

where $P(x|\lambda) = \lambda^x e^{-\lambda}/x!$ is a form of writing the Poisson distribution that will be dealt with in greater detail below.

Two discrete random variables X and Y are said to be independent if P(x,y) = P(x)P(y). Analogously, two continuous random variables U and V are said to be independent if p(u,v) = p(u)p(v). The reader easily can show that if the discrete random variables X and Y are independent, then P(x|y) = P(x) and P(y|x) = P(y). Similarly, if the continuous random variables U and V are independent, then p(u|v) = p(u) and p(v|u) = p(v).

If two events are mutually exclusive, the probabilities of these events merely add together to give the probability of either event occurring. This definition can be extended to the case where several random variables are involved.

To tie together the above ideas in one problem, consider two independent random variables X and Y which can take only nonnegative integer values. How can we determine the distribution of their sum Z = X + Y? The following procedure shows how to write P(z) in terms of P(x) and P(y):

$$\begin{aligned} P(z) &= P(Z = z) \\ &= P(X + Y = z) \\ &= P(X = z,\ Y = 0;\ \text{or},\ X = z - 1,\ Y = 1;\ \text{or},\ \ldots;\ \text{or},\ X = 0,\ Y = z) \\ &= P(X = z,\ Y = 0) + P(X = z - 1,\ Y = 1) + \ldots + P(X = 0,\ Y = z) \\ &= P(X = z)P(Y = 0) \\ &\quad + P(X = z - 1)P(Y = 1) + \ldots + P(X = 0)P(Y = z) \\ &= \sum_{i=0}^{z} P(X = z - i)P(Y = i) \end{aligned} \tag{8}$$

Note that both the additive property associated with the probabilities of mutually exclusive events and the multiplicative property associated with the probabilities of independent events have been used. To carry the calculation any further, it is necessary to specify the X and Y distributions. If X and Y have Poisson distributions

$$P(x) = \lambda^x e^{-\lambda}/x!$$
$$P(y) = \mu^y e^{-\mu}/y! \tag{9}$$

where $x = 0,1,2, \ldots, \lambda \geqslant 0$, $y = 0,1,2, \ldots$, and $\mu \geqslant 0$, then

$$\begin{aligned} P(z) &= \sum_{i=0}^{z} \frac{\lambda^{z-i}e^{-\lambda}}{(z-i)!}\frac{\mu^i e^{-\mu}}{i!} \\ &= \frac{e^{-(\lambda+\mu)}}{z!}\sum_{i=0}^{z}\frac{z!}{i!(z-i)!}\mu^i\lambda^{z-i} \\ &= \frac{(\lambda+\mu)^z e^{-(\lambda+\mu)}}{z!} \end{aligned} \tag{10}$$

because

$$\sum_{i=0}^{\infty}\frac{z!}{i!(z-i)!}\mu^i\lambda^{z-i} = (\lambda+\mu)^z \tag{11}$$

(that is, because of the binomial theorem of algebra). Thus, the sum of two independent Poisson random variables is also a Poisson random variable. If the sum of two independent Poisson random variables is a Poisson random variable, then is it also true that their difference is a Poisson random variable? The answer is no, as the following argument shows:

$$\begin{aligned} P(z') &= \sum_{i=0}^{\infty} P(X = i + z')P(Y = i) \\ &= \sum_{i=0}^{\infty}\frac{\lambda^{i+z'}e^{-\lambda}}{(i+z')!}\frac{\mu^i e^{-\mu}}{i!} \\ &= \lambda^{z'}e^{-(\lambda+\mu)}\sum_{i=0}^{\infty}\frac{(\lambda\mu)^i}{i!(i+z')!} \end{aligned} \tag{12}$$

This expression for the difference $Z' = X - Y$ cannot be written in terms of any familiar probability distribution (although Skellam[11] has shown that it can be written in terms of Bessel functions).

Once the probability law governing a random variable is known, P(x) in the case of a discrete random variable and p(x) in the case of a continuous random variable, it is possible to calculate expectation values such as the mean and variance. The expectation value, E{f (X) }, of a function f(X) of a random variable X is defined by

$$E\{f(X)\} = \sum_x f(x)P(x) \tag{13}$$

for the case of discrete random variable and

$$E\{f(X)\} = \int f(x)p(x)dx \tag{14}$$

for the case of a continuous random variable. The sum in the definition for the discrete case includes all values of x such that P(x) is not zero, for the continuous case the integration includes all intervals for which the integral of f(x)p(x) does not vanish. Usually, it will be obvious what probability law is to be used in the calculation of a particular expectation value, but in cases of ambiguity the following notation can be used:

$$\underset{W|x}{E}\{f(W)\} \equiv \int f(w)p(w|x)dw \tag{15}$$

The mean of a random variable X is defined to be E{X}; this quantity also is called the average. The variance of a random variable is defined to be $E\{[X - E\{X\}]^2\}$; this quantity also is called the dispersion. The square root of the variance is called the standard deviation and the relative standard deviation (also called the coefficient of variation) is defined to be the ratio of the standard deviation to the mean. The following formula often is used to calculate the variance:

$$\begin{aligned} Var\{X\} &\equiv E\{[X - E\{X\}]^2\} \\ &= E\{X^2\} - 2[E\{X\}]^2 + [E\{X\}]^2 \\ &= E\{X^2\} - [E\{X\}]^2 \end{aligned} \tag{16}$$

(in deriving this expression for the variance, use must be made of the fact that E{X} is a deterministic quantity, not a random variable). The calculation of the standard deviation using a digital computer and the above formula can give inaccurate results (for reasons which concern computer arithmetic, not statistics). The recursive method given by Forsythe[12] is recommended for computer calculations of the variance (methods described by Chan and Lewis[13] and West[14] also can be used).

## III. THE BINOMIAL, NORMAL, AND POISSON DISTRIBUTIONS

Because the normal and Poisson probability laws are two of the most important probability laws in science, it is important to understand as much as possible about them. Both laws are limiting cases of the binomial probability law, the probability law which governs a sum of random variables each of which is governed by the same Bernoulli probability law.

The Bernoulli, binomial, and Poisson probability laws govern discrete random variables (that is, random variables that take only integer values). This means that questions such as the following are not senseful: ''What is the probability that the random variable of the Poisson distribution takes the value 13.35?'' It makes sense, of course, to ask questions such as '' What is the probability that the random variable of the Poisson distribution takes the value 13 when the mean of the distribution is 13.35?'' or ''What is the probability that the random variable of the Poisson distribution takes the value 14 when the mean of the distribution is 13.35?'' Correspondingly, the graphical representation of the Poisson distribution (or the Bernoulli distribution, or the binomial distribution) must be a set of discrete points, not a continuous curve; a continuous curve can be used only in a graphical representation of a continuous distribution, such as the normal distribution. A histogram can be used to depict a discrete distribution, but it must be remembered that such a histogram is a substitute for a set of discrete points.

A random variable X is said to have a Bernoulli distribution characterized by a parameter p ($0 < p < 1$) if the probability distribution function for the random variable can be written

$$P(X = x) \equiv P(x) = p^x(1 - p)^{1-x} \tag{17}$$

for x = 0,1. The Bernoulli probability law governs a random variable which can take one of only two possible values (the two possible values carry the nondescript labels ''0'' and ''1'' here), one value occurring with probability p and the other with probability $1 - p$. The traditional example given of a random variable which has a Bernoulli distribution is the outcome of a coin toss; in the toss of a coin there are just two possible outcomes, heads or tails (designate heads as ''0'' and tails as ''1''), and each outcome occurs with equal probability, that is, $p = 1 - p = 1/2$ (these two probabilities would differ if the coin were not perfect, but would always be such that their sum was unity). With a little thought, the reader can conceive of a multitude of other possible situations where the Bernoulli probability law is appropriate.

The mean and variance of the Bernoulli distribution are calculated as follows:

$$\begin{aligned} E\{X\} &= \sum_x xP(x) \\ &= \sum_{x=0}^{1} xp^x(1 - p)^{1-x} \\ &= p \end{aligned} \tag{18}$$

$$\begin{aligned} \mathrm{Var}\{X\} &= E\{X^2\} - [E\{X\}]^2 \\ &= \sum_x x^2P(x) - p^2 \\ &= \sum_{x=0}^{1} x^2p^x(1 - p)^{1-x} - p^2 \\ &= p(1 - p) \end{aligned} \tag{19}$$

The coefficient of variation (or the relative standard deviation) of the Bernoulli distribution is then

$$\begin{aligned} c_v\{X\} &\equiv \frac{\sqrt{\mathrm{Var}\{X\}}}{E\{X\}} \\ &= \sqrt{(1 - p)/p} \end{aligned} \tag{20}$$

Consider a sample size of n, $X_1, \ldots, X_n$, with each datum, $X_i$, independently generated by the same Bernoulli process (such independently generated samples are called random samples). The sum of this sequence of random variables is itself a random variable, $X \equiv X_1 + \ldots + X_n$, distributed as

$$P(x) = \frac{n!}{x!(n - x)!} p^x(1 - p)^{n-x} \tag{21}$$

where x can be any integer such that $0 \leq x \leq n$. This is a binomial distribution characterized by the parameters n and p (a derivation of the binomial distribution can be found in any textbook on probability; see, for example, Reference 15, pages 30 to 31). An obvious application of the binomial distribution is in answering the question: ''In n tosses of a coin, what is the probability of obtaining x heads ($x \leq n$), given that the probability of obtaining a head on a single toss is p?'' Or, the binomial distribution can be used to answer the

此为试读，需要完整PDF请访问：www.ertongbook.com