

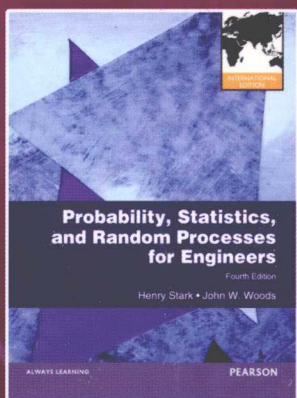
国外电子与通信教材系列

英文版

PEARSON

# 概率、统计与随机过程 (第四版)

Probability, Statistics,  
and Random Processes for Engineers  
Fourth Edition



[美] Henry Stark 著  
John W. Woods



电子工业出版社  
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

<http://www.phei.com.cn>

国外电子与通信教材系列

# 概率、统计与随机过程

( 第四版 ) ( 英文版 )

Probability, Statistics, and Random  
Processes for Engineers

Fourth Edition

[ 美 ] Henry Stark 著  
John W. Woods

电子工业出版社

Publishing House of Electronics Industry

北京 · BEIJING

## 内 容 简 介

本书从工程应用的角度, 全面阐述概率、统计与随机过程的基本理论及其应用。全书共 11 章(其中第 10 章和第 11 章为网上资源), 首先简单介绍概率论, 然后各章分别讨论随机变量、随机变量的函数、均值与矩、随机矢量、统计(包括参数估计和假设检验)、随机序列、随机过程基础知识和深入探讨, 最后讨论了统计信号处理中的相关应用。书中给出了大量电子和信息系统相关实例, 每章给出了丰富的习题。

本书适合作为电子信息类专业本科生和研究生的“随机信号分析”或“随机过程及其应用”课程的双语教学教材, 也可供从事相关技术领域研究的科技人员参考。

Original edition, entitled **Probability, Statistics, and Random Processes for Engineers, Fourth Edition**, 9780273752288 by Henry Stark, John W. Woods, published by Pearson Education, Inc., publishing as Pearson International, Copyright © 2012 Pearson Education Limited.

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage retrieval system, without permission from Pearson Education, Inc. China edition published by PEARSON EDUCATION ASIA LTD., and PUBLISHING HOUSE OF ELECTRONICS INDUSTRY, Copyright © 2012.

This edition is manufactured in the People's Republic of China, and is authorized for sale and distribution in the mainland of China exclusively (except Taiwan, Hong Kong SAR and Macau SAR).

本书英文版专有出版权由 Pearson Education (培生教育出版集团) 授予电子工业出版社。未经出版者预先书面许可, 不得以任何方式复制或抄袭本书的任何部分。

本书在中国大陆地区生产, 仅限在中国大陆发行。

本书贴有 Pearson Education (培生教育出版集团) 激光防伪标签, 无标签者不得销售。

版权贸易合同登记号 图字: 01-2012-5644

### 图书在版编目(CIP)数据

概率、统计与随机过程: 第 4 版 = Probability, Statistics, and Random Processes for Engineers: 英文 / (美) 斯塔克 (Stark, H.), (美) 伍兹 (Woods, J. W.) 著. — 北京: 电子工业出版社, 2012.8

国外电子与通信教材系列

ISBN 978-7-121-17668-5

I. ①概… II. ①斯… ②伍… III. ①概率论 - 高等学校 - 教材 - 英文 ②数理统计 - 高等学校 - 教材 - 英文 ③随机过程 - 高等学校 - 教材 - 英文 IV. ①O21

中国版本图书馆 CIP 数据核字 (2012) 第 161886 号

策划编辑: 马 岚

责任编辑: 马 岚

印 刷: 北京京师印务有限公司

装 订:

出版发行: 电子工业出版社

北京市海淀区万寿路 173 信箱 邮编: 100036

开 本: 787 × 980 1/16 印张: 44 字数: 1281 千字

印 次: 2012 年 8 月第 1 次印刷

定 价: 89.00 元

凡所购买电子工业出版社的图书有缺损问题, 请向购买书店调换; 若书店售缺, 请与本社发行部联系。联系及邮购电话: (010) 88254888。

质量投诉请发邮件至 zlt@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线: (010) 88258888。



# Contents

<b>Preface</b>	<b>11</b>
<b>1 Introduction to Probability</b>	<b>13</b>
1.1 Introduction: Why Study Probability?	13
1.2 The Different Kinds of Probability	14
Probability as Intuition	14
Probability as the Ratio of Favorable to Total Outcomes (Classical Theory)	15
Probability as a Measure of Frequency of Occurrence	16
Probability Based on an Axiomatic Theory	17
1.3 Misuses, Miscalculations, and Paradoxes in Probability	19
1.4 Sets, Fields, and Events	20
Examples of Sample Spaces	20
1.5 Axiomatic Definition of Probability	27
1.6 Joint, Conditional, and Total Probabilities; Independence	32
Compound Experiments	35
1.7 Bayes' Theorem and Applications	47
1.8 Combinatorics	50
Occupancy Problems	54
Extensions and Applications	58
1.9 Bernoulli Trials—Binomial and Multinomial Probability Laws	60
Multinomial Probability Law	66
1.10 Asymptotic Behavior of the Binomial Law: The Poisson Law	69
1.11 Normal Approximation to the Binomial Law	75
Summary	77
Problems	78
References	89

<b>2</b>	<b>Random Variables</b>	<b>91</b>
2.1	Introduction	91
2.2	Definition of a Random Variable	92
2.3	Cumulative Distribution Function	95
	Properties of $F_X(x)$	96
	Computation of $F_X(x)$	97
2.4	Probability Density Function (pdf)	100
	Four Other Common Density Functions	107
	More Advanced Density Functions	109
2.5	Continuous, Discrete, and Mixed Random Variables	112
	Some Common Discrete Random Variables	114
2.6	Conditional and Joint Distributions and Densities	119
	Properties of Joint CDF $F_{XY}(x, y)$	130
2.7	Failure Rates	149
	Summary	153
	Problems	153
	References	161
	Additional Reading	161
<b>3</b>	<b>Functions of Random Variables</b>	<b>163</b>
3.1	Introduction	163
	Functions of a Random Variable (FRV): Several Views	166
3.2	Solving Problems of the Type $Y = g(X)$	167
	General Formula of Determining the pdf of $Y = g(X)$	178
3.3	Solving Problems of the Type $Z = g(X, Y)$	183
3.4	Solving Problems of the Type $V = g(X, Y), W = h(X, Y)$	205
	Fundamental Problem	205
	Obtaining $f_{VW}$ Directly from $f_{XY}$	208
3.5	Additional Examples	212
	Summary	217
	Problems	218
	References	226
	Additional Reading	226
<b>4</b>	<b>Expectation and Moments</b>	<b>227</b>
4.1	Expected Value of a Random Variable	227
	On the Validity of Equation 4.1-8	230
4.2	Conditional Expectations	244
	Conditional Expectation as a Random Variable	251
4.3	Moments of Random Variables	254
	Joint Moments	258
	Properties of Uncorrelated Random Variables	260
	Jointly Gaussian Random Variables	263
4.4	Chebyshev and Schwarz Inequalities	267

Markov Inequality	269
The Schwarz Inequality	270
4.5 Moment-Generating Functions	273
4.6 Chernoff Bound	276
4.7 Characteristic Functions	278
Joint Characteristic Functions	285
The Central Limit Theorem	288
4.8 Additional Examples	293
Summary	295
Problems	296
References	305
Additional Reading	306
<b>5 Random Vectors</b>	<b>307</b>
5.1 Joint Distribution and Densities	307
5.2 Multiple Transformation of Random Variables	311
5.3 Ordered Random Variables	314
5.4 Expectation Vectors and Covariance Matrices	323
5.5 Properties of Covariance Matrices	326
Whitening Transformation	330
5.6 The Multidimensional Gaussian (Normal) Law	331
5.7 Characteristic Functions of Random Vectors	340
Properties of CF of Random Vectors	342
The Characteristic Function of the Gaussian (Normal) Law	343
Summary	344
Problems	345
References	351
Additional Reading	351
<b>6 Statistics: Part 1 Parameter Estimation</b>	<b>352</b>
6.1 Introduction	352
Independent, Identically, Observations	353
Estimation of Probabilities	355
6.2 Estimators	358
6.3 Estimation of the Mean	360
Properties of the Mean-Estimator Function (MEF)	361
Procedure for Getting a $\delta$ -confidence Interval on the Mean of a Normal Random Variable When $\sigma_X$ Is Known	364
Confidence Interval for the Mean of a Normal Distribution When $\sigma_X$ Is Not Known	364
Procedure for Getting a $\delta$ -Confidence Interval Based on $n$ Observations on the Mean of a Normal Random Variable when $\sigma_X$ Is Not Known	367
Interpretation of the Confidence Interval	367

6.4	Estimation of the Variance and Covariance	367
	Confidence Interval for the Variance of a Normal Random variable	369
	Estimating the Standard Deviation Directly	371
	Estimating the covariance	372
6.5	Simultaneous Estimation of Mean and Variance	373
6.6	Estimation of Non-Gaussian Parameters from Large Samples	375
6.7	Maximum Likelihood Estimators	377
6.8	Ordering, more on Percentiles, Parametric Versus Nonparametric Statistics	381
	The Median of a Population Versus Its Mean	383
	Parametric versus Nonparametric Statistics	384
	Confidence Interval on the Percentile	385
	Confidence Interval for the Median When $n$ Is Large	387
6.9	Estimation of Vector Means and Covariance Matrices	388
	Estimation of $\mu$	389
	Estimation of the covariance $K$	390
6.10	Linear Estimation of Vector Parameters	392
	Summary	396
	Problems	396
	References	400
	Additional Reading	401
<b>7</b>	<b>Statistics: Part 2 Hypothesis Testing</b>	<b>402</b>
7.1	Bayesian Decision Theory	403
7.2	Likelihood Ratio Test	408
7.3	Composite Hypotheses	414
	Generalized Likelihood Ratio Test (GLRT)	415
	How Do We Test for the Equality of Means of Two Populations?	420
	Testing for the Equality of Variances for Normal Populations:	
	The F-test	424
	Testing Whether the Variance of a Normal Population Has a Predetermined Value:	428
7.4	Goodness of Fit	429
7.5	Ordering, Percentiles, and Rank	435
	How Ordering is Useful in Estimating Percentiles and the Median	437
	Confidence Interval for the Median When $n$ Is Large	440
	Distribution-free Hypothesis Testing: Testing If Two Population are the Same Using <i>Runs</i>	441
	Ranking Test for Sameness of Two Populations	444
	Summary	445
	Problems	445
	References	451

---

<b>8</b>	<b>Random Sequences</b>	<b>453</b>
8.1	Basic Concepts	454
	Infinite-length Bernoulli Trials	459
	Continuity of Probability Measure	464
	Statistical Specification of a Random Sequence	466
8.2	Basic Principles of Discrete-Time Linear Systems	483
8.3	Random Sequences and Linear Systems	489
8.4	WSS Random Sequences	498
	Power Spectral Density	501
	Interpretation of the psd	502
	Synthesis of Random Sequences and Discrete-Time Simulation	505
	Decimation	508
	Interpolation	509
8.5	Markov Random Sequences	512
	ARMA Models	515
	Markov Chains	516
8.6	Vector Random Sequences and State Equations	523
8.7	Convergence of Random Sequences	525
8.8	Laws of Large Numbers	533
	Summary	538
	Problems	538
	References	553
<b>9</b>	<b>Random Processes</b>	<b>555</b>
9.1	Basic Definitions	556
9.2	Some Important Random Processes	560
	Asynchronous Binary Signaling	560
	Poisson Counting Process	562
	Alternative Derivation of Poisson Process	567
	Random Telegraph Signal	569
	Digital Modulation Using Phase-Shift Keying	570
	Wiener Process or Brownian Motion	572
	Markov Random Processes	575
	Birth–Death Markov Chains	579
	Chapman–Kolmogorov Equations	583
	Random Process Generated from Random Sequences	584
9.3	Continuous-Time Linear Systems with Random Inputs	584
	White Noise	589
9.4	Some Useful Classifications of Random Processes	590
	Stationarity	591
9.5	Wide-Sense Stationary Processes and LSI Systems	593
	Wide-Sense Stationary Case	594
	Power Spectral Density	596
	An Interpretation of the psd	598



More on White Noise	602
Stationary Processes and Differential Equations	608
9.6 Periodic and Cyclostationary Processes	612
9.7 Vector Processes and State Equations	618
State Equations	620
Summary	623
Problems	623
References	645

登录华信教育资源网 (<http://www.hxedu.com.cn>) 可下载第 10 章和第 11 章。

采用本书作为教材的教师可获得本书配套教辅, 请联系 010-88254555 或发邮件至 [te\\_services@phei.com.cn](mailto:te_services@phei.com.cn)

<b>10 Advanced Topics in Random Processes</b>	<b>647</b>
10.1 Mean-Square (m.s.) Calculus	647
Stochastic Continuity and Derivatives [10-1]	647
Further Results on m.s. Convergence [10-1]	657
10.2 Mean-Square Stochastic Integrals	662
10.3 Mean-Square Stochastic Differential Equations	665
10.4 Ergodicity [10-3]	670
10.5 Karhunen-Loève Expansion [10-5]	677
10.6 Representation of Bandlimited and Periodic Processes	683
Bandlimited Processes	683
Bandpass Random Processes	686
WSS Periodic Processes	689
Fourier Series for WSS Processes	692
Summary	694
Appendix: Integral Equations	694
Existence Theorem	695
Problems	698
References	711
<b>11 Applications to Statistical Signal Processing</b>	<b>712</b>
11.1 Estimation of Random Variables and Vectors	712
More on the Conditional Mean	718
Orthogonality and Linear Estimation	720
Some Properties of the Operator $\hat{E}$	728
11.2 Innovation Sequences and Kalman Filtering	730
Predicting Gaussian Random Sequences	734
Kalman Predictor and Filter	736
Error-Covariance Equations	741
11.3 Wiener Filters for Random Sequences	745
Unrealizable Case (Smoothing)	746
Causal Wiener Filter	748

11.4	Expectation-Maximization Algorithm	750
	Log-likelihood for the Linear Transformation	752
	Summary of the E-M algorithm	754
	E-M Algorithm for Exponential Probability Functions	755
	Application to Emission Tomography	756
	Log-likelihood Function of Complete Data	758
	E-step	759
	M-step	760
11.5	Hidden Markov Models (HMM)	761
	Specification of an HMM	763
	Application to Speech Processing	765
	Efficient Computation of $P[E M]$ with a Recursive Algorithm	766
	Viterbi Algorithm and the Most Likely State Sequence for the Observations	768
11.6	Spectral Estimation	771
	The Periodogram	772
	Bartlett's Procedure—Averaging Periodograms	774
	Parametric Spectral Estimate	779
	Maximum Entropy Spectral Density	781
11.7	Simulated Annealing	784
	Gibbs Sampler	785
	Noncausal Gauss–Markov Models	786
	Compound Markov Models	790
	Gibbs Line Sequence	791
	Summary	795
	Problems	795
	References	800

## Appendix A Review of Relevant Mathematics

A.1	Basic Mathematics	A-1
	Sequences	A-1
	Convergence	A-2
	Summations	A-3
	Z-Transform	A-3
A.2	Continuous Mathematics	A-4
	Definite and Indefinite Integrals	A-5
	Differentiation of Integrals	A-6
	Integration by Parts	A-7
	Completing the Square	A-7
	Double Integration	A-8
	Functions	A-8

---

A.3 Residue Method for Inverse Fourier Transformation	A-10
Fact	A-11
Inverse Fourier Transform for psd of Random Sequence	A-13
A.4 Mathematical Induction	A-17
References	A-17
<b>Appendix B Gamma and Delta Functions</b>	<b>B-1</b>
B.1 Gamma Function	B-1
B.2 Incomplete Gamma Function	B-2
B.3 Dirac Delta Function	B-2
References	B-5
<b>Appendix C Functional Transformations and Jacobians</b>	<b>C-1</b>
C.1 Introduction	C-1
C.2 Jacobians for $n = 2$	C-2
C.3 Jacobian for General $n$	C-4
<b>Appendix D Measure and Probability</b>	<b>D-1</b>
D.1 Introduction and Basic Ideas	D-1
Measurable Mappings and Functions	D-3
D.2 Application of Measure Theory to Probability	D-3
Distribution Measure	D-4
<b>Appendix E Sampled Analog Waveforms and Discrete-time Signals</b>	<b>E-1</b>
<b>Appendix F Independence of Sample Mean and Variance for Normal Random Variables</b>	<b>F-1</b>
<b>Appendix G Tables of Cumulative Distribution Functions: the Normal, Student t, Chi-square, and F</b>	<b>G-1</b>
<b>Index</b>	<b>I-1</b>



# Preface

While significant changes have been made in the current edition from its predecessor, the authors have tried to keep the discussion at the same level of accessibility, that is, less mathematical than the measure theory approach but more rigorous than formula and recipe manuals.

It has been said that *probability* is hard to understand, not so much because of its mathematical underpinnings but because it produces many results that are counter intuitive. Among practically oriented students, *Probability* has many critics. Foremost among these are the ones who ask, “What do we need it for?” This criticism is easy to answer because future engineers and scientists will come to realize that almost every human endeavor involves making decisions in an uncertain or probabilistic environment. This is true for entire fields such as insurance, meteorology, urban planning, pharmaceuticals, and many more. Another, possibly more potent, criticism is, “What good is probability if the answers it furnishes are not certainties but just inferences and likelihoods?” The answer here is that an immense amount of good planning and accurate predictions can be done even in the realm of uncertainty. Moreover, applied probability—often called *statistics*—does provide near certainties: witness the enormous success of political polling and prediction.

In previous editions, we have treaded lightly in the area of statistics and more heavily in the area of random processes and signal processing. In the electronic version of this book, graduate-level signal processing and advanced discussions of random processes are retained, along with new material on statistics. In the hard copy version of the book, we have dropped the chapters on applications to statistical signal processing and advanced topics in random processes, as well as some introductory material on pattern recognition.

The present edition makes a greater effort to reach students with more expository examples and more detailed discussion. We have minimized the use of phrases such as,

“it is easy to show...”, “it can be shown...”, “it is easy to see...,” and the like. Also, we have tried to furnish examples from real-world issues such as the efficacy of drugs, the likelihood of contagion, and the odds of winning at gambling, as well as from digital communications, networks, and signals.

The other major change is the addition of two chapters on elementary statistics and its applications to real-world problems. The first of these deals with parameter estimation and the second with hypothesis testing. Many activities in engineering involve estimating parameters, for example, from estimating the strength of a new concrete formula to estimating the amount of signal traffic between computers. Likewise many engineering activities involve making decisions in random environments, from deciding whether new drugs are effective to deciding the effectiveness of new teaching methods. The origin and applications of standard statistical tools such as the  $t$ -test, the Chi-square test, and the  $F$ -test are presented and discussed with detailed examples and end-of-chapter problems.

Finally, many self-test multiple-choice exams are now available for students at the book Web site. These exams were administered to senior undergraduate and graduate students at the Illinois Institute of Technology during the tenure of one of the authors who taught there from 1988 to 2006. The Web site also includes an extensive set of small MATLAB programs that illustrate the concepts of probability.

In summary then, readers familiar with the 3<sup>rd</sup> edition will see the following significant changes:

- A new chapter on a branch of statistics called *parameter estimation* with many illustrative examples;
- A new chapter on a branch of statistics called *hypothesis testing* with many illustrative examples;
- A large number of new homework problems of varying degrees of difficulty to test the student’s mastery of the principles of statistics;
- A large number of self-test, multiple-choice, exam questions calibrated to the material in various chapters available on the Companion Web site.
- Many additional illustrative examples drawn from real-world situations where the principles of probability and statistics have useful applications;
- A greater involvement of computers as teaching/learning aids such as (i) graphical displays of probabilistic phenomena; (ii) MATLAB programs to illustrate probabilistic concepts; (iii) homework problems requiring the use of MATLAB/ Excel to realize probability and statistical theory;
- Numerous revised discussions—based on student feedback—meant to facilitate the understanding of difficult concepts.

Henry Stark, IIT  
Professor Emeritus

John W. Woods, Rensselaer  
Professor

The publishers would like to thank Dr Murari Mitra and Dr Tamaghna Acharya of Bengal Engineering and Science University for reviewing content for the International Edition.

# *Introduction to Probability*

## 1.1 INTRODUCTION: WHY STUDY PROBABILITY?

One of the most frequent questions posed by beginning students of probability is, “Is anything truly random and if so how does one differentiate between the truly random and that which, because of a lack of information, is treated as random but really isn’t?” First, regarding the question of truly random phenomena, “Do such things exist?” As we look with telescopes out into the universe, we see vast arrays of galaxies, stars, and planets in apparently random order and position.

At the other extreme from the cosmic scale is what happens at the atomic level. Our friends the physicists speak of such things as the *probability* of an atomic system being in a certain state. The uncertainty principle says that, try as we might, there is a limit to the accuracy with which the position and momentum can be simultaneously ascribed to a particle. Both quantities are fuzzy and indeterminate.

Many, including some of our most famous physicists, believe in an essential randomness of nature. Eugen Merzbacher in his well-known textbook on quantum mechanics [1-1] writes,

The probability doctrine of quantum mechanics asserts that the indetermination, of which we have just given an example, is a property inherent in nature and not merely a profession of our temporary ignorance from which we expect to be relieved by a future better and more complete theory. The conventional interpretation thus denies the possibility of an ideal theory which would encompass the present quantum mechanics

but would be free of its supposed defects, the most notorious “imperfection” of quantum mechanics being the abandonment of strict classical determinism.

But the issue of determinism versus inherent indeterminism need never even be considered when discussing the validity of the probabilistic approach. The fact remains that there is, quite literally, a nearly uncountable number of situations where we cannot make any categorical deterministic assertion regarding a phenomenon because we cannot measure all the contributing elements. Take, for example, predicting the value of the noise current  $i(t)$  produced by a thermally excited resistor  $R$ . Conceivably, we might accurately predict  $i(t)$  at some instant  $t$  in the future if we could keep track, say, of the  $10^{23}$  or so excited electrons moving in each other’s magnetic fields and setting up local field pulses that eventually all contribute to producing  $i(t)$ . Such a calculation is quite inconceivable, however, and therefore we use a probabilistic model rather than Maxwell’s equations to deal with resistor noise. Similar arguments can be made for predicting the weather, the outcome of tossing a real physical coin, the time to failure of a computer, dark current in a CMOS imager, and many other situations. Thus, we conclude: Regardless of which position one takes, that is, determinism versus indeterminism, we are forced to use probabilistic models in the real world because we do not know, cannot calculate, or cannot measure all the forces contributing to an effect. The forces may be too complicated, too numerous, or too faint.

Probability is a mathematical model to help us study physical systems in an *average sense*. We have to be able to repeat the experiment many times under the same conditions. Probability then tells us how often to expect the various outcomes. Thus, we cannot use probability in any meaningful sense to answer questions such as “What is the probability that a comet will strike the earth tomorrow?” or “What is the probability that there is life on other planets?” The problem here is that we have no data from similar “experiments” in the past.

R. A. Fisher and R. Von Mises, in the first third of the twentieth century, were largely responsible for developing the groundwork of modern probability theory. The modern axiomatic treatment upon which this book is based is largely the result of the work by Andrei N. Kolmogorov [1-2].

## 1.2 THE DIFFERENT KINDS OF PROBABILITY

There are essentially four kinds of probability. We briefly discuss them here.

### Probability as Intuition

This kind of probability deals with judgments based on intuition. Thus, “She will probably marry him” and “He probably drove too fast” are in this category. Intuitive probability can lead to contradictory behavior. Joe is still likely to buy an imported Itsibitsi, world famous for its reliability, even though his neighbor Frank has a 19-year-old Buick that has never broken down and Joe’s other neighbor, Bill, has his Itsibitsi in the repair shop. Here Joe may be behaving “rationally,” going by the statistics and ignoring, so-to-speak, his personal observation. On the other hand, Joe will be wary about letting his nine-year-old

daughter Jane swim in the local pond, if Frank reports that Bill thought that he might have seen an alligator in it. This despite the fact that no one has ever reported seeing an alligator in this pond, and countless people have enjoyed swimming in it without ever having been bitten by an alligator. To give this example some credibility, assume that the pond is in Florida. Here Joe is ignoring the statistics and reacting to, what is essentially, a rumor. Why? Possibly because the *cost* to Joe “just-in-case” there is an alligator in the pond would be too high [1-3].

People buying lottery tickets intuitively believe that certain number combinations like month/day/year of their grandson’s birthday are more likely to win than say, 06-06-06. How many people will bet even odds that a coin that, heretofore has behaved “fairly,” that is, in an unbiased fashion, will come up heads on the next toss, *if in the last seven tosses it has come up heads?* Many of us share the belief that the coin has some sort of memory and that, after seven heads, that coin must “make things right” by coming up with more tails.

A mathematical theory dealing with intuitive probability was developed by B. O. Koopman [1-4]. However, we shall not discuss this subject in this book.

**Probability as the Ratio of Favorable to Total Outcomes  
(Classical Theory)**

In this approach, which is not experimental, the probability of an event is computed *a priori*<sup>†</sup> by counting the number of ways  $n_E$  that  $E$  can occur and forming the ratio  $n_E/n$ , where  $n$  is the number of all possible outcomes, that is, the number of all alternatives to  $E$  plus  $n_E$ . An important notion here is that all outcomes are equally likely. Since equally likely is really a way of saying equally probable, the reasoning is somewhat circular. Suppose we throw a pair of unbiased six-sided dice<sup>‡</sup> and ask what is the probability of getting a 7. We partition the outcome space into 36 equally likely outcomes as shown in Table 1.2-1, where each entry is the sum of the numbers on the two dice.

**Table 1.2-1 Outcomes of Throwing Two Dice**

2nd die	1st die					
	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

<sup>†</sup> *A priori* means relating to reasoning from self-evident propositions or prior experience. The related phrase, *a posteriori* means relating to reasoning from observed facts.

<sup>‡</sup> We will always assume that our dice have six sides.



The total number of outcomes is 36 if we keep the dice distinct. The number of ways of getting a 7 is  $n_7 = 6$ . Hence

$$P[\text{getting a 7}] = \frac{6}{36} = \frac{1}{6}.$$

### Example 1.2-1

(*toss a fair coin twice*) The possible outcomes are HH, HT, TH, and TT. The probability of getting at least one tail T is computed as follows: With  $E$  denoting the event of getting at least one tail, the event  $E$  is the set of outcomes

$$E = \{\text{HT, TH, TT}\}.$$

Thus, event  $E$  occurs whenever the outcome is HT or TH or TT. The number of elements in  $E$  is  $n_E = 3$ ; the number of all outcomes  $N$ , is four. Hence

$$P[\text{at least one T}] = \frac{n_E}{n} = \frac{3}{4}.$$

Note that since no physical experimentation is involved, there is no problem in postulating an ideal “fair coin.” Effectively, in classical probability every experiment is considered “fair.”

The classical theory suffers from at least two significant problems: (1) It cannot deal with outcomes that are not equally likely; and (2) it cannot handle an infinite number of outcomes, that is when  $n = \infty$ . Nevertheless, in those problems where it is impractical to actually determine the outcome probabilities by experimentation and where, because of symmetry considerations, one can indeed argue equally likely outcomes, the classical theory is useful.

Historically, the classical approach was the predecessor of Richard Von Mises' [1-6] relative frequency approach developed in the 1930s, which we consider next.

### Probability as a Measure of Frequency of Occurrence

The relative frequency approach to defining the probability of an event  $E$  is to perform an experiment  $n$  times. The number of times that  $E$  appears is denoted by  $n_E$ . Then it is tempting to define the probability of  $E$  occurring by

$$P[E] = \lim_{n \rightarrow \infty} \frac{n_E}{n}. \quad (1.2-1)$$

Quite clearly since  $n_E \leq n$  we must have  $0 \leq P[E] \leq 1$ . One difficulty with this approach is that we can never perform the experiment an infinite number of times, so we can only estimate  $P[E]$  from a finite number of trials. Secondly, we *postulate* that  $n_E/n$  approaches a limit as  $n$  goes to infinity. But consider flipping a fair coin 1000 times. The likelihood of getting exactly 500 heads is very small; in fact, if we flipped the coin 10,000 times, the likelihood of getting exactly 5000 heads is even smaller. As  $n \rightarrow \infty$ , the event of observing